Developing countries integration in international trade: measurements and determinants

Mariana VIJIL a, b
a, INRA, UMR1302 SMART, F-35000 Rennes, France
b, Agrocampus Ouest, UMR1302 SMART, F-35000 Rennes, France
mariana.vijil@rennes.inra.fr

Contributed Paper prepared for presentation at the “ETSG 2011” conference, Copenhagen Denmark, September 8-10, 2011

Work in progress, please do not cite without permission of the author

Abstract

In the empirical literature it has been demonstrated that trade can be a powerful engine to enhance economic development and poverty reduction. But international market participation is a multidimensional concept that can be apprehended through different outputs: a greater openness to trade, an increase in the extensive margin of exports and imports, an enhancement of the value-added exported or an increase in export competitiveness or import performance. A new openness indicator is proposed that takes into account these different dimensions of trade integration. We try then to reveal the main obstacles to trade that influence it and that can be addressed by existing development instruments, like aid for trade and preferential market access. We find strong evidence that further investment in infrastructure and additional market access is needed to increase the participation of developing countries in international trade.

Keywords: Developing countries, International trade, Trade costs, Ordered probit, Cluster analysis

Copyright by Mariana Vijil. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
I- Introduction:

In the empirical literature it has been demonstrated that trade can be a powerful engine to enhance economic development and poverty reduction (Sachs and Warner, 1995; Frankel and Romer, 1999; Dollar and Kraay, 2004; Winters et al., 2004). Thus, outward-oriented growth has been a popular development strategy within low income countries since structural adjustments plans. Nevertheless, as Brun et al. (2005) noted, evidence is consistent with the claim that poor countries have been marginalized by the recent wave of globalization. Also, the share of the poorest developing countries in global trade has not increased.

Therefore, the development dimension has become the main issue within the Doha Round multilateral negotiations at the World Trade Organisation (WTO). The aim is to re-equilibrate trade rules in favour of developing countries needs, to reinforce their participation in international trade and to make them fully benefit from trade liberalisation. Considering its impacts on development, these questions are particularly important as regards the achievement of the anti-poverty Millenium Development Goals by 2015.

Thus, there is a need to increase developing countries trade integration. In this respect, two development instruments have been widely used by the international community: preferential trade schemes accorded by industrialized partners and aid for trade. Nowadays, evidence suggests that the combination of these two policies is essential. Indeed, market access seems not enough for some countries facing internal obstacles to trade, as a lack of knowledge, excessive red tape, insufficient financing and poor infrastructure (Hoekman and Nicita, 2008; Portugal-Perez and Wilson, 2008; Huchet-Bourdon et al., 2009). However, before analyzing the impact of these two development instruments on trade integration, there is a need to define the exact objective that these instruments aim to achieve.

Actually, international market participation is a multidimensional concept that can translate into different outputs: a greater openness to trade, an increase in the intensive or the extensive margin of exports and imports, a rise in the value-added exported or an increase in export competitiveness or import performance. We are not considering here other features more related to a broader concept of openness, like distortive economic policies in the vein of trade-related policy variables. Indeed, we believe that the latter are instruments for trade integration and not a measure of it. Thus, we will focus exclusively on quantitative trade outputs.

This objective follows the evidence that international trade raises income (Frankel and Romer, 1999; Irwin and Tervio, 2002; Noguer and Siscart, 2005; Feenstra and Kee, 2008; Freund and Bolaki, 2008; Kee et al. 2009) and that different kind of improvements in international trade indicators will have different impacts on development. Indeed, an increase in the extensive margin of trade does not have the same effect as an increase in the volume of existing flows. Our goal is then to build within a unique framework a snapshot of the position of each country in each of these dimensions of international trade in order to have a broader concept of trade integration.

In that matter, we rely on the well known openness ratio, on an export concentration index, on the extensive margin of trade, on an export quality indicator and on a trade performance measure. As far as we know, the study of all these dimensions of trade integration within a single framework has never been done. Furthermore, some important aspects on this concept,

---

1 Note that this question differs from the well know debate on trade-policy measures. We do not attempt to analyse countries in terms of their political-induced trade integration. In that matter, there is an abundant debate
like export and import performance using constant market shares techniques, have been largely under-considered by researchers.

An analysis of the different dimensions of trade integration presented above in a single framework can easily reveal country priorities. In fact, the extensive empirical literature on trade costs (Wilson et al., 2003) offers us a variety of obstacles potentially explaining the level of trade integration observed by each country group. Thus, a classification of developing countries using the indicators cited above can reveal homogeneous groups that also share similar obstacles to trade, like a distortive trade and economic policy, bad geographic characteristics or a lack of institutions and infrastructures. Because a major part of these costs can be addressed by development instruments, it will be interesting to test what kind of trade cost reduction allows a country to graduate from one cluster to another. Therefore, this analysis can be a guide to improve the allocation of aid for trade and preferential trade agreements between developing economies.

To do so, we will start by using clustering techniques in order to construct cluster of countries that share similar values in their trade indicators. It turns out that these cluster can be ordered, representing some increasing degree of trade integration. Thereafter, using an ordered probit approach we will be able to test the impact of aid for trade and preferential schemes on the probability that a country belongs to the highest or the lowest integrated group of countries. Finally, marginal effects derived from the ordered probit will allow us to quantify how much is needed for a country to graduate from his cluster.

In order to measure de multidimensionality of trade for developing countries, we start with a literature review that justifies the choice of indicators in term of their proximity with trade integration and their impact on economic development. We then use statistical classification techniques to construct clusters of countries using the trade output measures justified above, providing thus a new indicator of trade integration. Finally, we try to reveal the main trade costs and internal characteristics that explain the fact that a country belongs to a particular cluster using an ordered probit. This original approach allows us to use an integrated framework to evaluate the impact of diverse trade costs on a multidimensional trade integration measure.

that has been summarized in Rodriguez and Rodrik (2001), which refers to the work of Sachs and Warner (1995), Harrison (1996), Pritchett (1996) and Edwards (1998) attempts to confront diverse trade-policy indicators in order to asses if lower policy-induced barriers to international trade are desirables.
II- Literature review

The abundant debate on trade integration and growth has been alimented by a large variety of openness indicators that can be classified in three branches: widely used GDP-related openness ratios (Frankel and Romer, 1999; Frankel and Rose, 2002, Dollar and Kraay, 2004), trade-policy indicators (Leamer, 1988; Harrison, 1996; Pritchett, 1996; Guillaumont, 2001; Kee et al. 2009) and economic policies measures that include broader aspects related to trade (Sachs and Warner, 1995). Because we want to be apart from the debate on trade-policy barriers (Frankel and Romer, 1999; Rodriguez and Rodrik, 2001), we will focus on the trade to GDP ratio as a measure of dependence of the economy on international trade.

Indeed, Frankel and Romer (1999) evidence suggest that countries that are more open tend to be richer. In a cross-country regression over 105 countries in 1985, they find that a greater trade over GDP ratio enhances income per capita, controlling for the neoclassical determinants of growth and for the endogeneity problem. This result is corroborated by Irwin and Terviö (2002) over a larger period and the robustness of the results are also proven by Noguer and Suscart (2005). Frankel and Rose (2002) also use a similar specification over 210 countries between 1960 and 1996. They find that openness to trade increases the growth rate of per capita income, even after controlling for neoclassical determinants, institutions and geography. Using regressions of changes in decadal growth rates on instrumented changes in trade and institutional quality, Dollar and Kraay (2003) tend to confirm the previous results and highlight the complementarities between trade and institutions. More recently, using broader databases, cross-section and panel-data estimations, Freund and Bolaky (2009) and Chang et al. (2009) also suggest that the positive effect of openness on income is enhanced by policy complementarities.

Moreover, even if the debate is still open, the natural resource curse empirical literature review highlights the negative impact of export concentration in primary products on growth (Auty, 2000; Sachs and Warner, 1999; Lederman and Maloney, 2008). This trend in the literature claims that export diversification away from natural resources is one of the most important economic policies for developing countries. Indeed, considering the deterioration of terms of trade observed the last decades (Prebisch, 1950; Singer, 1950; Harvey et al. 2010), a concentration in primary commodities exports worsen the trade balance deficit (Guillaumont, 1980). Also, the price volatility that characterizes this kind of products tends to increase macroeconomic instability, resulting in economic and political fragility, underinvestment, short-time planning and inflation, between other consequences (Guillaumont, 1987). Finally, a concentration in a limited range of trade partners also tends to intensify the business cycles synchronisation (Calderon et al., 2007). This phenomenon can deepen macroeconomic instability and then affect pro-cyclical variables like social spending, investment, credit and productivity (Fatas, 2002).

Nevertheless, there is evidence that export diversification follows the pattern of development. Indeed, Klinger and Lederman (2004, 2006) and Cadot et al. (2011) find that exports tend to diversify and then re-concentrate with income per capita. Moreover, it seems that this evolution mainly comes from the extensive margin (Cadot et al., 2011)\(^2\). This subject has become a matter of careful analysis in recent years. Papers like the one of Hummels and

---

\(^2\) The intensive margin refers to the variation in export values among existing lines (existing exports). Nevertheless, the definition of extensive margin differs among authors. In the case of Cadot et al. (2011), the extensive margin reflects variation in the number of new lines (new products exported or existing products exported to new partners), whereas in Hummels and Klenow (2005) it refers to export of new products.
Klenow (2005) and Cadot et al. (2011) decompose cross-country trade variation in intensive and extensive margins and study the contribution of each one on trade growth. Using a cross section of countries in 1995, Hummels and Klenow (2005) find that about two-third of the differences in the amount of trade between countries is explained by import and export extensive margins. Cadot et al. (2011) also find that diversification of exports in developing countries occurs mainly at the extensive margin, especially in their first stages of development.

As Hummels and Klenow (2005) state, these results “can be extremely important in determining the welfare consequences of access to trade”. Indeed, because of the theoretical implications of an increase in the extensive margin of trade, also called the variety of products traded, this issue appears as an important policy indicator and a key element on the pathway to economic development (Cadot et al., 2011). In fact, there are three main sources of gain from trade related to product variety, predicted by monopolistic competition models (Feenstra, 2010): first, an increase in consumer welfare arising from the rise in the variety of available products; second, and improvement in industry or country productivity due to an increase in the variety of imported inputs; and finally, an increase in industry or country productivity arising from an increase in export variety.

The first prediction is that trade will allow consumers to access new imported varieties of differentiated products. And because consumers have a “love of variety”, this will induce an increase in welfare. Those gains have recently been measured by Broda and Weinstein (2006) for the United States over the period 1972 to 2001, using the methodologies from Feenstra (1994). These authors find that the gain from trade due to the expansion of import varieties amounts to 2.6 percent of GDP in 2001. Feenstra (2010) also finds that the gain from import variety in 1996 over all the 146 countries used in the sample vary between 9.4 and 15.4 percent of world GDP, depending on the value of elasticity of substitution used. Nevertheless, Arkolakis et al. (2008) find very small gains from increased import variety in the case of Costa Rica between 1986 and 1992.

The second prediction of the monopolistic competition model derives from the endogenous growth model like the one of Grossman and Helpman (1991), were research and development are considered as determinants of growth. In this framework, trade increases the growth rate if it allows the diffusion of international spillover of knowledge via imports of intermediated inputs that increase the variety of inputs available to producers (Feenstra, 2004; 2010). Indeed, an increase in the variety of differentiated inputs is expected to enhance the efficiency of the firm or the industry. Feenstra et al. (1999) provide strong evidences for this using the export variety of South Korea and Taiwan. They find that the variation on this variable between 1975 and 1991 had a positive and significant effect on total factor productivity in 9 out of the 16 sectors studied, principally in industries that rely and produce differentiated manufactures, and thus seems to match with endogenous growth predictions. Funke and Ruhwedel (2001a; b; 2002) obtain similar results for OECD and East-Asian countries.

Finally, the third prediction arises from the monopolistic competition model with heterogeneous firms developed by Melitz (2003), where productivity is endogenous trough the self selection of exporters. Within this framework, exporters are more productive on average than domestic firms. This model predicts that a fall in trade costs will increase the average productivity of the country and raise the number of exporters. Since each exporter produces a differentiated variety, an increase in export variety can be associated with rising average productivity and GDP (Feenstra and Kee, 2008). These authors’ evidence suggest that over 48 countries from 1980 to 2000, the 3,3 % per annum increase in export variety to the
United-States observed in the data is associated with a 3.3% higher GDP for exporters over the period. It should also be noted that benefits from the producer side seem to be greater than benefits from the consumer side. Indeed, as reminded by these authors, welfare benefits for exporting countries due to an increase in export variety are larger than the welfare improvements founded by Broda and Weistein (2006) for the United-States due to an increase in its import variety.

Concerning the trade quality dimension, Hausmann et al. (2007) provide evidences that the productivity level associated with exports influence positively subsequent economic growth. Indeed, by constructing an index of the quality of the basket exported by each country, they find that exporters that specialize in products of better quality benefits from higher economic growth. This evidence has been corroborated by Guerson et al. (2007) for the Argentinean case. Nevertheless, within an industry, developing countries tend to specialize in goods of low value-added (Schott, 2004). Thus, the emergence of new activities of better quality appears to be critical for developing countries.

Finally, considering the positive returns for exports arising from an increase in market shares in world trade, the performance dimension of trade integration seems to be an interesting indicator. Starting from the evolution of growth rates of export market shares, the variance analysis technique allow us to disentangle export performance due to internal country characteristics from effects due to world demand. Using this method, Cheptea et al. (2005) analyse the export performance of a large sample of countries for the period 1995-2002 and find a high heterogeneity between developing countries. It is worth to be noted that the use of constant market shares techniques to evaluate trade performance have been largely under-considered by researchers.

III- Methodology

3.1.Database

For the construction of the indicators we principally use the trade values in current US dollars from the BACI international trade database constructed by the CEPII at a HS6 desegregated level (among 5,000 lines), except for the market share analysis which is done at the HS2 level. This database is a modified version of COMTRADE, for which trade flows have been corrected for the institutional capacity of both the exporter and the importer. This harmonization procedure gives us more accurate trade values and allows us to extend the number of countries for which data is available (mostly developing countries). Thus, we are able to work with around 199 countries representing all regions and every level of development between 2000 and 2007. Intra UE-15 trade has been dropped from the database. We also use the World Development Indicators (WDI) database for the construction of the quality index of Hausmann et al. (2007).

In this section, we compute several indicator of trade integration for each country and each year. These variables are the trade to GDP openness ratio, the Herfindahl concentration index at a product level, the extensive margin of trade, the quality of export basket and the trade performance obtained from a market shares analysis. Indices are then averaged over the 2000-2007 period in order to proceed with the clustering.
3.2. Construction of indicators

Measure of trade openness

In order to measure openness to trade, we rely on the trade over GDP ratio that accounts for the fact that some countries like Singapore are re-export platforms. Indeed, generally the openness ratio is the ratio of the sum of exports and imports to GDP, or $\frac{X+M}{GDP}$.

Nevertheless, as Guillaumont (2001) and Combes et al. (2002) noted, some countries can appear with an openness ratio higher than 100% because they are re-export platforms and add very low value-added to the imports they re-exports. Indeed, the openness ratio is a ratio of turnover to value-added and to be more accurate we should measure the value-added included in exports. Lacking this, we rather use as proxy for openness the ratio of exports and imports to total available resources: $\frac{X+M}{GDP+M}$. We compute this indicator in a yearly base and we average it over the 2000-2007 period in order to use it as an input in the clustering.

Measure of export concentration / diversification of products

To measure the concentration of exports at a product level, we compute the Herfindahl index, a variable easy to understand and that has been widely used by the literature on trade concentration (Jaud et al., 2009; Cadot et al., 2011). This index, normalized to range between zero and one, is

$$H^* = \frac{\sum_k s_k^2 - \frac{1}{n}}{1 - \frac{1}{n}},$$

where $s_k = x_k / \sum_{k=1}^n x_k$ is the share of export line $k$ in total exports, and $n$ is the number of export lines that could be exported\(^3\). Temporal and individual indices have been omitted for convenience. This index measures the degree of concentration of the export basket, and varies between 0 (totally de-concentrated) to 1 (totally concentrated). One of the advantages of this measure is that it can be read as a percentage of concentration. We compute a yearly Herfindahl index over the 2000-2007 period and we average it over years in order to use it as an input in the clustering.

Measure of export / import variety

In order to allow comparability of the index between countries and time, the export variety (or extensive margin of exports) is constructed following a modified version proposed by Feenstra and Kee (2008) of the Hummels and Klenow (2005) index.

Hummels and Klenow (2005) propose a measure of “extensive margin” of trade that is consistent with product variety for a CES function. This indicator can be defined as changes in exports or imports that are due to changes in the number of goods (a change in the variety of products) rather than changes in the amount purchased of each good. Besides the fact that this formula is consistent with trade theory, we choose it among all the definitions of

\(^3\) $n$ varies around 5 000 lines depending on the years.
extensive margin available in the literature review because it takes into account the importance of the traded good instead of roughly counting lines. To see the proof of this measure please refer to Feenstra and Kee (2008) and Feenstra (2010), we will present here only the final formula. The construction of the import extensive margin measure is symmetric to the export one.

The construction of the indicator is based on the idea that exports from countries \( h \) and \( F \) differ but have some products varieties in common. This common set is denoted by \( J \equiv (J_h h \cap J_F F) \neq \emptyset \). An inverse measure of export variety from country \( h \) will be defined by

\[
\lambda^h(J) = \frac{\sum_{j \in J} p^h(j) q_i^h(j)}{\sum_{j \in J} p^h(j) q_i^h(j)}.
\]  

(2)

Therefore, the ratio \( \frac{\lambda_h^J(J)}{\lambda_h^F(J)} \) measures the export variety of country \( h \) relative to country \( F \). And it increases with the variety exported from country \( h \), and decreases with the variety exported from country \( F \). Thus, to be measured, this indicator needs a consistent comparison country \( F \).

Feenstra and Kee (2008) use the worldwide exports from all countries to the United States (US) as benchmark. Indeed, US appear as the major partner in terms of imported variety (US imports almost 99% of all the varieties existing) and provides highly disaggregated trade databases (until 10 digit codes). Nevertheless, as Feenstra and Kee (2008) noted, it would be preferable to use countries’ worldwide imports instead of US imports. Indeed, this restriction makes the measure dependent to the import structure of the US. And for countries that export goods that have a small value in the import structure of this partner, the magnitude of their export variety will appear under-evaluated. As an example, considering the incentives accorded by the US government, there is an important domestic production of meat. Thus, we expect this country to import relatively small values of this product. This introduces a downward bias to the export variety of countries like Argentina, which exports relatively high values of meat. Also, countries that do not export to the US some kind of varieties (mostly developing countries) will also see their index under-valuated. Thus, in order to correct for these effects we prefer to work with the entire world as the benchmark \( F \), as in Hummels and Klenox (2005), even if this forces us to use only HS-6 desegregated trade data.

Moreover, we need a benchmark \( F \) that doesn’t change thought time, in order to associate any variation in the indicator to a variation in the export variety of the country \( h \). So, following Feenstra and Kee (2008) we take the union of all products sold in the world market in any year over the period 2000-2007, and we average real exports sales of each product over years. In this way, \( J_F F \equiv \cup_{h,F} J_h F \) is the total set of varieties imported by the entire world in sector \( i \) over all years, and \( p^F_i (j) q_i^F (j) \) is the average real value of world imports for product \( j \) (summed over all source countries and averaged across years). Then, comparing country \( h \) to the world \( (F) \) allows us to set \( \lambda^h(J) = 1 \) and the export variety by country \( h \) takes the form:
Thus, export variety only changes due to variations in the numerator, and thus, due to changes in the set of goods sold by the country \( h \). This allows us to do comparison of export varieties across countries and over time. Moreover, this indicator goes beyond a simple count of trade lines, because it takes into account the relevance of the sector \( i \) (HS-6 line) in world trade. This is the methodology we use for the construction of our export and import variety. We compute a yearly indicator and we average it over the period in order to use it as an input in the clustering.

Measure of export quality

The quality of the export basket is constructed following Hausmann et al. (2007). First, they propose an index called PRODY that attributes a level of productivity to each \( k \) (HS-6) line. The total exports for a country \( i \) is,

\[
X_i = \sum_{k=1}^{n} x_{ik}
\]

And the level of productivity \( PRODY_k \) associated to each \( k \) (HS-6 line) is constructed as

\[
PRODY_k = \frac{\sum X_{ik}}{\sum x_{ik}} Y_i,
\]

where \( Y_i \) is the GDP per capita in Purchasing Power Parity of each country \( i \). This index is a variant of the Balassa’s index of revealed comparative advantage, weighted by the level of development of exporters. This way, exports from developed countries are considered as more productive that the ones coming from developing economies.

Finally, the level of productivity associated to the export basket of each country \( i \) is,

\[
EXPY_i = \sum_k \left( \frac{x_{ik}}{X_i} \right) PRODY_k.
\]

Thus, it depends on the degree of concentration of the export basket, weighted by the quality of the products exported. The underlying idea behind this indicator is that diversifying its exports basket away from products of low productivity may accelerate subsequent growth. We compute a yearly \( EXPY_i \) indicator and we average it over the 2000-2007 period in order to use it as an input in the clustering.

Measure of export / import performance

Following Berzeg (1978), Jayet (1993) and Cheptea et al. (2005), trade performance indicators are derived from a constant market share technique obtained by weighted variance analysis. Starting from the evolution of growth rates of export market shares, the variance analysis technique allows us to disentangle the export performance due to internal country characteristics, from effects due to partner demand and product composition of the export basket. This framework also has the additional advantage of providing standard errors for the estimated effects. Moreover, the import performance measure can also be obtained from the same specification.
The average growth rate of sectoral bilateral exports $x_{ijk}$ can be disentangled in three different effects:

$$
\sum X_{ijk}^0 = \sum X_{ij}^0 \alpha_i = \sum X_{ik}^0 \beta_j = \sum X_{jk}^0 \gamma_k = 0
$$

(6)

Thus, the average growth rate of exports from country $i$ to partner $j$ in sector $k$ over the period 2000-20007, $x_{ijk}$, can be regressed over three dummies which correspond to the export performance $\alpha_i$, the partner import dynamism $\beta_j$ and the sectoral world demand $\gamma_k$. The constant $m$ roughly represents the average growth rate of world exports. Equation (6) is estimated using a weighted variance analysis technique, weighted by the value of exports at the beginning of the period $X_{ijk}^0$ in order to solve the heteroscedasticity problem (Jayet, 1993). It should be noted that instead of working with annual export growth rates as in Cheptea et al. (2005), we decided to work with an average export growth rate over the period. Indeed, results seem very sensitive to the high volatility of exports of highly concentrated developing countries. In fact, some developing countries (mostly island states or countries in conflict) face high volatility in the value of their exports, and when we perform the weighted variance analysis on a year to year basis, we find that this instability is absorbed by the export performance fixed effect. Thus, working with average export growth rates allows us to smooth this instability.

Once equation (6) has been estimated, another weighting is necessary to give more importance to countries that play a prominent role in world markets. Indeed, we need to correct for the fact that China has a higher influence in international trade than Vanuatu, or that agroindustrial trade is more important than the fur one. This is done by weighting each fixed effect coefficient by the country/partner/sector market share in world markets at the beginning of the period. Finally, because of collinearity problems between the constant and the dummies, a country/partner/sector has been automatically dropped from the estimation. Thus, results should be read as deviations from the omitted effect, which is fixed to zero. Instead of this, we rearrange results in order to interpret each dummy coefficient as deviations from the world average. Finally, we run a Fischer test over results of equation (6) in order to evaluate their statistical significance and confirm the good fit of the decomposition (Jayet, 1993). We see that more than half of the fixed effects are significantly different from zero (Table 1, Annexe 1).

To convert the average growth rate of country $i$ exports of good $k$ to partner $j$, $x_{ijk}$, to an average growth rate of total country $i$ exports, $r_i$, we can rewrite equation (6) in the following form:

$$
r_i = \hat{m} + \hat{\alpha}_i + \sum_j \frac{X_{ij}^0}{X_{ij}} \hat{\beta}_j + \sum_k \frac{X_{ik}^0}{X_{ik}} \hat{\gamma}_k ,
$$

(7)
Where $\hat{\alpha}_i$ is the export performance of country $i$ due to its internal characteristics, $\frac{X_i^0}{X_i}$, his export performance due to the import dynamics of his partners, and $\frac{X_{i,k}^0}{X_i}$ his export performance due to the dynamics of the sectors where he exports. Thus, we obtain the decomposition we were looking for: average export growth rates are disentangled in a more accurate export performance, a geographical structure effect and a sectoral structure effect.

Finally, because we want to focus on an export performance indicator related to the growth in exports market shares rather than on exports growth, a further computation is needed. A country’s exports market share growth can be expressed as follows:

$$g_i = \frac{(1 + r_i)}{(1 + \hat{m})} - 1.$$  
(8)

Thus, following the equations (6) and (7), the market share average growth rate can be rewritten as:

$$g_i = \frac{\hat{\alpha}_i}{(\hat{m} + 1)} + \frac{\sum_{j} X_i^0 \hat{\beta}_j}{(\hat{m} + 1)} + \frac{\sum_{k} X_{i,k}^0 \hat{\gamma}_k}{(\hat{m} + 1)}.$$  
(9)

And thus, the export performance indicator explained by country characteristics becomes

$$Performance_i = \frac{\hat{\alpha}_i}{(\hat{m} + 1)}.$$  
(10)

We follow equation (10) to construct the export performance indicator. And because by construction the export, import and sectoral effects are independent between each other in equation (6), the import dynamic for each country can be measured by the $\hat{\beta}_j$ from equation (7) (Cheptea et al., 2005). These are the two trade performance indicators that will be used in the future clustering.

In term of interpretation, considering that by construction the export performance indicator is dependent on the export structure of each country at the beginning of the period (partners and sectors in $t_0=2000$), a positive sign in this variable tells us that the country has been able to increase its competitiveness, or/and that it has been capable to export to more dynamic importing markets or in sectors with a higher world demand. In the case of the import performance indicator, a positive sign tell us that the country has increased its demand for foreign goods, and that this is not explained by a better performance of its exporting partners or by the kind of goods it imports. Finally, now that we have constructed all the trade-related indicators, we can proceed to the clustering analysis.

### 3.3. Clustering analysis

We use the openness ratio, the Herfindahl export concentration index, the export and import variety measures, the quality of exports, and the export and import performance using market
shares, as inputs to construct a classification of countries by implementing hierarchical and non-hierarchical clustering methods.

As Johnson and Wichern (2002) noted, hierarchical agglomerative clustering techniques provide clusters of countries that are similar. Thus, in our case we are able to provide groups of countries that are similar in terms of their degree of trade integration, taking into account the multidimensionality of this concept. This technique also allows us to obtain the optimal number of clusters with respect to the data. Finally, applying a k-means procedure after the hierarchical agglomerative clustering method enables us to enhance the robustness of our results.

We proceed to a clustering of countries based on the value of each of the seven indicators presented above. Note that we have four variables that are related to exports, two to imports and one that address both issues. This means that we weigh more heavily the export performance of countries. Indeed, the international trade models and empirical evidence, and the observation of South-Asian countries tends to indicate that export-led growth is more desirable than further dependence on imports.

The first cluster of countries is obtained using the Ward’s hierarchical agglomerative clustering technique. This method allows for the union of the two clusters where fusion results in a minimum increase in “information loss”, that is, in a minimum increase in the sum-of-squares. Because Ward’s approach is very sensible to outliers, we drop from the sample countries that are too dissimilar. In our case, ten outliers were dropped from our sample. After the Ward’s method, we apply the k-means procedure to validate the clustering, allowing countries to move from one cluster to another in an iterative way, until each country is closer to the members of its cluster that to the members of a neighbor one.

We expect that clusters obtained after the k-means could be ordered in an increasing way, from the highest integrated in world trade to the least integrated one, thus creating a new discrete variable reflecting trade integration. We also think it is likely that countries within a cluster will share similar structural characteristics, like a comparable level of development, a lack of infrastructure, bad quality institutions or a closed trade policy. Indeed, literature review on trade costs argues that these variables are determinants of export volumes and diversification (Limao and Venables, 2001; Wilson et al., 2003; 2005, Anderson and van Wincoop, 2004; Shepherd, 2010, Dennis and Shepherd, 2011). We also anticipate that some groups of countries recognized by the donor community, like the Least Developed Countries (LDCs) or the Small and Vulnerable Economies (SVEs) will appear concentrated in some clusters.

Finally, in order to highlight the determinants of trade integration, we will estimate an ordered probit model over the discrete variable created by the clustering. This will allow us to test the impact of structural and policy-induced country characteristics on the probability that a country belongs to the highest or the lowest trade-integrated group of economies. It is worth reminding that the clustering is only constructed using the seven trade indicators; none of the structural and policy-induced determinants are used to cluster countries. Thus, coherent results with the literature review on trade costs arising from the probit estimation will tend to confirm a good fit of our clustering analysis.

---

4 Angola, China, Germany, Equatorial Guinea, Hong Kong, Ireland, Japan, Palau, Singapore and the USA.
The level of development is approximated using the log of GDP per capita, and data comes from the World Bank’s WDI. This variable allows us to control for a lot of features highly related to the level of development, like education and health. Some geographic characteristics are measured using the log of population, population density, a landlocked dummy, an oil exporter dummy and the FERDI’s remoteness indicator. The quality of institutions is alternatively measured by the “time to export” variable from the Doing Business database (Djankov et al. 2007), by the governance indicators from the World Bank and by the institutional database from the ICRG. Infrastructure quality and quantity are controlled by two indices regrouping by simple mean and Principal Component Analysis (PCA) the following variables: road density and percentage of paved road, and the number of subscribers to mobile and fixed lines (per 100 people).

Finally, we introduce the value added of agriculture in total GDP (in %) and a high-technology exports (% of manufactured exports) variables that measure, respectively, the structure of production of the economy and the content in research and development of the industrial sector. These variables come from the WDI. The trade policy is measured by the trade restrictiveness indices constructed by Kee et al. (2009), by the weighted mean tariff applied and by the percentage of tariff lines with peaks. The results from the ordered probit will allow us to tell which set of factor enhances the probability that a country belongs to the highest or the lowest cluster. We will thus be able to quantify the needs of bad performing countries in order to graduate from the worse integrated clusters.
IV- Results

4.1. Clustering

The Ward’s clustering technique provides us with five very interesting clusters. Indeed, we observe that the oil exporting countries are automatically placed in a unique cluster. Nevertheless, some of these countries are re-organized within other clusters once we perform the k-means method.

The results of the clustering using hierarchical and k-means techniques can be seen in Table 1 and Figure 1. The cluster 1 is the category regrouping the least integrated countries and 5 the more integrated ones. We see that almost all the indicators follow a linear evolution among these ordered clusters, except for the Herfindahl index of exports concentration and for the openness ratio. The first result can be explained by the fact that oil exporting countries, which have an export basket highly concentrated in exports of oil, are mainly concentrated in Cluster 3. For the second fact, as predicted by theoretical models, larger countries tend to be more closed that smaller ones. Thus, a drop in the openness ratio for the 5th cluster may be explained by the high concentration of rich countries in it. This still holds when the clustering is made only within developing economies: oil exporting countries appears in the same cluster, and the Herfindahl index and the openness ratio are still not linear between categories. Thus, it seems that using the openness ratio as a measure of trade integration, like it is usually made in the research on the impact of trade on the level or the growth of GDP, may be misleading.

Finally, a look at the mean of some variables within each cluster provides a snapshot of the characteristics of the countries that share similar levels of trade integration. We see in Table 1 that the lowest cluster is mainly composed with LDCs and SVEs. We also observe that trade integration seems positively correlated with GDP per capita. Finally, we highlight the fact that India, which is a low income country, belongs to the highest integrated cluster. This certainly should be the case of China if it was not considered as an outlier by our statistical technique.

Figure 1: Ward’s and k-means clustering, mean values over the period 2000-2007

Source: author’s calculations, BACI and WDI data

* Angola, China, Germany, Equatorial Guinea, Hong Kong, Ireland, Japan, Palau, Singapore and the USA are considered as outside values by the clustering procedure. Other missing country is due to missing values.
## Table 1: Clustering over all counties (Ward’s technique + K-means)

<table>
<thead>
<tr>
<th>Cluster n°1</th>
<th>Quality index</th>
<th>Herfindahl Index</th>
<th>Export Variety</th>
<th>Import Variety</th>
<th>Import Performance</th>
<th>Export Performance</th>
<th>Openess</th>
<th>LIC</th>
<th>LMIC</th>
<th>UMIC</th>
<th>HIC</th>
<th>LDC</th>
<th>SVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>19842</td>
<td>0.036</td>
<td>0.102</td>
<td>0.405</td>
<td>-2.19E-04</td>
<td>-4.32E-04</td>
<td>0.234</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>76309</td>
<td>0.642</td>
<td>0.638</td>
<td>0.938</td>
<td>-3.08E-05</td>
<td>-1.59E-04</td>
<td>0.938</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAN</td>
<td>47026</td>
<td>0.198</td>
<td>0.378</td>
<td>0.734</td>
<td>-5.68E-05</td>
<td>-4.07E-04</td>
<td>0.519</td>
<td>64%</td>
<td>23%</td>
<td>11%</td>
<td>2%</td>
<td>66%</td>
<td>86%</td>
</tr>
<tr>
<td>STD</td>
<td>14197</td>
<td>0.134</td>
<td>0.132</td>
<td>0.112</td>
<td>2.34E-05</td>
<td>3.47E-05</td>
<td>0.175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cluster n°2</th>
<th>N=30</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>41591</td>
</tr>
<tr>
<td>MAX</td>
<td>92629</td>
</tr>
<tr>
<td>MEAN</td>
<td>62695</td>
</tr>
<tr>
<td>STD</td>
<td>14509</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cluster n°3</th>
<th>N=18</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>62000</td>
</tr>
<tr>
<td>MAX</td>
<td>130899</td>
</tr>
<tr>
<td>MEAN</td>
<td>88983</td>
</tr>
<tr>
<td>STD</td>
<td>16010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cluster n°4</th>
<th>N=38</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>64378</td>
</tr>
<tr>
<td>MAX</td>
<td>142028</td>
</tr>
<tr>
<td>MEAN</td>
<td>101630</td>
</tr>
<tr>
<td>STD</td>
<td>20738</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cluster n°5</th>
<th>N=14</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIN</td>
<td>81964</td>
</tr>
<tr>
<td>MAX</td>
<td>142419</td>
</tr>
<tr>
<td>MEAN</td>
<td>105157</td>
</tr>
<tr>
<td>STD</td>
<td>19336</td>
</tr>
</tbody>
</table>

Source: author’s calculations, BACI and WDI data.

Following the classification provided by the World Bank, LIC=Low Income Country, LMIC=Low Middle Income Country, UMIC=Upper Middle Income Country, HIC=High Income Country, LDC=Least Developed Country and SVE=Small and Vulnerable Economy.
4.2. Ordered probit

We did four different clustering: one with the entire sample of countries and performing only the Ward’s technique, the same one but adding the k-means technique over the Ward’s results (Table 1), and the same procedure for the sample of developing countries only. This will allow us to test the robustness of our results.

When we estimate the ordered probit over the discrete trade integration variable arising from the Ward’s clustering, the oil exporting dummy appears highly significant. This comforts our intuition that these countries are highly concentrated in one cluster. Thus, in order to obtain a linear ordering of clusters that represents an increase in trade integration, we had to drop the cluster N° 3 from the explained variable in equations 1 to 4 (Table 2). We see then that once we control for trade-policy variables, the significance on the fuel dummy disappears (Equation 3 and 4, Table 2). We see in these results that infrastructure seems to be an important determinant explaining the probability that a country belongs to the highest or the lowest category. Indeed, this variable appears with a positive sign and highly significant and tells us that an improvement in infrastructures increase the probability that a country belongs to the highest cluster and decrease its probability to belong to the lowest cluster. Other explanatory variables, like the log of GDP per capita, or geographic variables like the log of population and population density appears significant and with the expected sign. Results indicate that richer and larger countries have a higher probability to be in the highest cluster.

These results still holds when we estimate the ordered probit with the clusters created by the Ward’s and k-means techniques over the entire sample (Equations 5 to 8, Table 2). We observe that previous results still hold, and that the percentage of tariff lines with peaks always appears significant. Finally, the share of agriculture in the economy and the high technology content of exports seems to influence the trade integration. Thus, countries depending heavily on agriculture are more likely to be in the lowest cluster. This is consistent with the idea that countries diversify their production away from agriculture within the development process. Also, countries that invest more in research and development are more likely to be very well integrated in international trade.

We now turn to the specification that uses the clustering made only within the sample of developing countries. We see in equations 9 to 12 (Table 2) that previous results still hold. Indeed, infrastructure always appears as a determinant of trade integration, which is consistent with the extensive empirical evidence that point out the importance of infrastructure as a determinant of trade. Indeed, marginal effects suggest that for a country located at the mean of all the variables used in the estimation (12), Table 2, a unit increase in the index of infrastructure decrease of 42,5% the probability that this country does not belong to the Cluster n°2 and increase of 27% the probability that this economy will be located in Cluster n°3 (Table 3). This is a high economic impact that transit by infrastructure. Considering that it has been demonstrated that aid for trade effectiveness also transit via this channel (Helble et al. 2010; Vijil and Wagner, 2010), results seems to conclude that more investment is needed in this feature in order to increase the participation of developing countries in international trade.

Finally, we see the tariff and non-tariff barriers applied by to the rest of the world (MA-O credible) seems to increase the probability that a developing country belongs to the lowest cluster (equations 11 and 12, Table 2). Considering that this variable didn’t appear significant

---

5 Results under request.
in previous specifications, market access seems to be a matter only for exporter from developing countries. This suggest that preferential schemes should be used as an instrument to increase trade integration. For example, marginal effects suggest that an unit decrease in the MA-OTRI indicator enhance the probability that countries in cluster n°2 move to cluster n°3 of 208%, an decrease the probability that these countries stay in their initial cluster of 329% (Table 3).

V- Conclusion:

Despite the relevance of the trade integration concept on development debates, there is still little consensus on the definition of this subject and its measurements. However, this is particularly important to improve the allocation of aid for trade and preferential market access. Using clustering techniques, this work proposes a new trade indicator that combine within a unique framework the position of each country in different dimensions of overall international trade participation, including concepts like openness, diversification, variety, quality and performance. Then, an ordered probit performed on this new indicator reveals the main obstacles to trade that can be addressed by existing development instruments. We find that investment in infrastructure and further market access for exports from the South are needed. This is an interesting result, considering that evidence suggest that aid for trade effectiveness in term of export performance transits via the infrastructure channel (Vijil and Wagner, 2010). For future research, it seems then pertinent to test the effectiveness of aid for trade and preferential market access on the different dimensions of trade integration highlighted in this analysis. Furthermore, the new indicator proposed here can be used to revisit the research on trade and development.
<table>
<thead>
<tr>
<th>Ordered Probit estimations</th>
<th>All Countries – 4 clusters (“Fuel cluster” dropped) Clustering using Ward’s method</th>
<th>All Countries – 5 clusters Clustering using Ward’s and k-means methods</th>
<th>Developing Countries – 5 clusters Clustering using Ward’s and k-means methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (2) (3) (4)</td>
<td>(5) (6) (7) (8)</td>
<td>(9) (10) (11) (12)</td>
</tr>
<tr>
<td>GDP per capita (log)</td>
<td>1.050 (0.108)***</td>
<td>1.043 (0.121)***</td>
<td>1.117 (0.150)***</td>
</tr>
<tr>
<td>Population (log)</td>
<td>0.555 (0.066)***</td>
<td>0.299 (0.049)***</td>
<td>0.376 (0.061)***</td>
</tr>
<tr>
<td>Fuel dummy</td>
<td>0.761 (0.313)**</td>
<td>0.467 (0.327)***</td>
<td>0.888 (0.419)**</td>
</tr>
<tr>
<td>Landlocked dummy</td>
<td>-0.032 (0.285)</td>
<td>0.186 (0.244)***</td>
<td>0.182 (0.281)</td>
</tr>
<tr>
<td>Density of Population</td>
<td>-0.001 (0.001)**</td>
<td>-0.000 (0.001)**</td>
<td>-0.001 (0.000)**</td>
</tr>
<tr>
<td>Infrastructures</td>
<td>0.450 (0.194)**</td>
<td>1.397 (0.272)*****</td>
<td>1.259 (0.340)****</td>
</tr>
<tr>
<td>Days for export</td>
<td>-0.022 (0.014)</td>
<td>-0.006 (0.012)**</td>
<td>-0.007 (0.013)**</td>
</tr>
<tr>
<td>MA-OCTI</td>
<td>-4.139 (5.001)</td>
<td>-7.103 (4.301)***</td>
<td>-16.097 (4.811)****</td>
</tr>
<tr>
<td>OTRI</td>
<td>3.455 (2.506)</td>
<td>0.742 (1.703)**</td>
<td>1.800 (1.711)**</td>
</tr>
<tr>
<td>Tariff Peaks</td>
<td>-0.007 (0.011)</td>
<td>-0.021 (0.010)**</td>
<td>-0.021 (0.011)**</td>
</tr>
<tr>
<td>Applied Tariff (weighted mean)</td>
<td>-0.062 (0.068) **</td>
<td>0.030 (0.042)</td>
<td>-0.013 (0.045) **</td>
</tr>
<tr>
<td>High-tech Exports</td>
<td>0.004 (0.002)**</td>
<td>0.012 (0.004)**</td>
<td>0.009 (0.002)**</td>
</tr>
<tr>
<td>Remoteness</td>
<td>-0.004 (0.013)</td>
<td>-0.017 (0.013)</td>
<td>-0.014 (0.011)</td>
</tr>
<tr>
<td>Share of Agriculture in GDP</td>
<td>0.001 (0.016)</td>
<td>-0.047 (0.019)**</td>
<td>-0.052 (0.016)****</td>
</tr>
<tr>
<td>Observations</td>
<td>144 129 112 70</td>
<td>157 140 121 78</td>
<td>120 107 92 72</td>
</tr>
<tr>
<td>Pseudo R2</td>
<td>0.43 0.45 0.45 0.52</td>
<td>0.39 0.47 0.55 0.58</td>
<td>0.32 0.33 0.43 0.52</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Source: author’s calculations, BACI and WDI data
Table 3: Marginal effects for the estimation (12) in table 2

<table>
<thead>
<tr>
<th>Marginal Effects (%)</th>
<th>Cluster n°1</th>
<th>Cluster n°2</th>
<th>Cluster n°3</th>
<th>Cluster n°4</th>
<th>Cluster n°5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructures</td>
<td>-0.154</td>
<td>-42.491</td>
<td>26.949</td>
<td>15.695</td>
<td>0.001</td>
</tr>
<tr>
<td>MA-OTRI</td>
<td>1.199</td>
<td>328.705</td>
<td>-208.479</td>
<td>-121.42</td>
<td>-0.004</td>
</tr>
<tr>
<td>Tariff Peaks</td>
<td>0.003</td>
<td>0.957</td>
<td>-0.607</td>
<td>-0.353</td>
<td>-0.001</td>
</tr>
<tr>
<td>High-tech Exports</td>
<td>-0.001</td>
<td>-0.303</td>
<td>0.192</td>
<td>0.112</td>
<td>0.0001</td>
</tr>
<tr>
<td>Share of Agriculture in GDP</td>
<td>0.006</td>
<td>1.739</td>
<td>-1.103</td>
<td>-0.642</td>
<td>-0.0001</td>
</tr>
</tbody>
</table>

Source: author’s calculations, BACI and WDI data

Marginal effects of each independent variable are calculated holding all covariates at their sample mean.
Annex 1

Table 1: Absolute values of the Student test statistics on fixed effects from the weighted variance analysis

<table>
<thead>
<tr>
<th>Source</th>
<th>Exporter</th>
<th>Importer</th>
<th>Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: author’s calculations
References:


