South-South Trade:
Geography Matters

Souleymane Coulibaly*
DEEP-HEC, University of Lausanne and TEAM, University of Paris 1

Lionel Fontagné
CEPII and TEAM, University of Paris 1

May 2004

Acknowledgment: we are indebted to Marius Brülhart, Matthieu Crozet, Guillaume Gaulier, Sébastien Jean, Philippe Martin, Thierry Mayer, Daniel Mirza and the participants of the 2002 ETSG conference in Kiel for their helpful comments. Remaining errors are obviously ours.

*BFSH1 1015 Lausanne-Dorigny (Switzerland), Tel: +41 (0)21 692 3474, Fax: +41 (0)21 692 3365, e-mail: souleymane.coulibaly@hec.unil.ch

Abstract

Intra-subsaharan African trade appears to be very low, an outcome that is often attributed to the size of the exporting and the importing economies. If that were the explanation, there would be no untapped trade potential. We argue instead that the main determinants of this "missing trade" are geographical and infrastructure-related impediments. Being landlocked and poor translates into high trade costs. In this paper, we try to measure the impact of geographical impediments on South-South trade. We focus on the intra and extra-regional trade of the countries belonging to the West African Economic and Monetary Union. We use an Armington-based model in order to evaluate the impact of geographical and infrastructure-related impediments on bilateral trade flows within this region. We find two main results: paving all inter-state roads would increase trade by a factor of 3, and crossing a transit country reduces intra-bilateral trade flows by 6%, ceteris paribus.

Keywords: *South-South trade, landlocked, transport infrastructure, border infrastructure*

J.E.L classification: *F11, F15, O55*

1 The Puzzle

"The road to hell is unpaved", according to a journalist[1] riding a beer truck from Douala to Bertoua, two towns in Cameroon separated by less than 500 km. Indeed, "according to a rather optimistic schedule, it should have taken 20 hours, including overnight rest. It took four days. When the truck arrived, it was carrying only two-thirds of its original load". Hence, the role of decaying roads and police harassment throughout the journey.

How geography and infrastructure affect trade flows among developing countries is not anecdotal. According to Sachs (2001) "since sea-navigable regions are generally richer than landlocked regions, regions that are both temperate and easily accessible to sea-based trade almost everywhere have achieved a very high measure of economic development. Tropical and landlocked regions, by contrast, are among the very poorest in the world". Geographical patterns may explain and keep up inequalities among nations: a glance at the world economy points to developing landlocked countries loosely integrated to international trade, as can be seen in Table 1 below.

Table 1: The disadvantage of landlocked countries
(unit: billion $ US, current value 2000)

		Developing countries				Developed
		Africa	Asia	America	Mideast	countries
Landlocked	Export	0.7	0.5	1.5		69.6
	GDP	3.1	2.1	7.9		149.2
	$\frac{Export}{GDP}$	0.2	0.3	0.2		0.5
Non-Landlocked	Export	4.6	61.4	13.8	22.0	232.8
	GDP	13.22	141.5	66.6	54.7	1178.5
	$\frac{Export}{GDP}$	0.4	0.4	0.2	0.4	0.2

Sources: World Development Indicators 2000 and our calculations.

Landlocked developing countries are less involved in international trade than landlocked developed countries.The export to GDP ratio for developing landlocked countries is 20%, while for developed landlocked countries this ratio is 50% (2000). Turning to non-landlocked countries, the ratio is respectively 40% and 20% for developing and developed countries.

But here, landlocking interfers with economic size; hence the need to sort out these two impacts with a gravity-type methodology. The poor performance of Southern countries is confirmed by a gravity model on a sample of 84 developed and developing exporters (see the list in Table 14 and 15 in Annexes) as can be seen in Table 7 in Annexes. Controlling distance, GDP, GDP per capita, contiguity and common language variables, it appears that European

[1] The Economist print edition 2002.

landlocked countries[2] trade 30% less than other countries in the world, while non-European landlocked countries trade 40% less. Besides, African countries seem to face higher impediments to trade since estimation indicates that an African country is trading 60% on average. Hence, land-locking and more generally geography have no straightforward impact on trade: the explanation might be a combination of geography and other development-related determinants.

Another prominent evidence is the limited benefits of South-South trade agreements so far (Greenaway & Milner, 1990); intra-regional trade (particularly in subsaharan Africa) remains very low. In the West African Economic and Monetary Union (WAEMU henceforth)[3] for instance, the share of intra-regional trade in total trade did not exceed 3% during the 1990s, despite an openness rate of 70%.

These results show the tremendous weakness of South-South trade and raise three issues that will be addressed in this paper:

1) What is the magnitude of untapped trade potential in the South?

2) What responsibility does geography[4] bear?

3) Is the traditional gravity-type methodology a suitable econometric device to sort out these effects?

The difficulty of the first issue is to find the best definition of trade potential. According to Havrylyshyn (1985) a country's optimal level of trade in any geographical direction is that which leads to the greatest gains in its economic welfare. This is an interesting heuristic definition but generally, trade economists focus on residuals of a gravity model to assess trade potentials. Based on the latter approach, Foroutan & Pritchett (1993) claim that there is no untapped potential in subsaharan Africa intra-regional trade. Concerning the second issue, Amjadi & Yeats (1995) find that the relatively low level of subsaharan African exports was essentially due to high transport costs [5]. Between 1990 and 1991, the net freight and insurance bill of this region represented 15% of the value of their exports, compared to less than 6% for all the developing countries. Limao & Venables (2000) suggest a significant impact of transportation infrastructure quality on transport costs and consequently, on trade flows:poor infrastructures account for 40% of predicted transport costs for coastal countries and 60% for landlocked countries. Concerning the third issue, Fontagné, Pajot & Pasteels (2002)stress an heterogeneity problem: using a sample of developed and developing countries they find a strong non-linearity in the impact of income per capita on trade, leading to biased elasticities in a sample obtained with heterogeneous countries.

These quotations deserve credit for giving scientifically-based answers to such contentious questions, but are missing an explicit model taking into ac-

[2] In our sample these countries are: Austria, Switzerland, Czech Republic and Hungary

[3] This Union consists of eight countries: Benin, Burkina Faso, Cote d'Ivoire, Guinea-Bissau, Mali, Niger, Senegal and Togo (See Figure 1).

[4] By geography, we mean physical geography as well as infrastructure endowments. See Henderson, Shalizi & Venables (2001) for review of the literature and Limao & Venables (2000) for an attempt at measuring the impact of infrastructure and geographical location of a country on transport costs.

[5] They examine net freight and insurance payments from IMF balance of payment statistics.

count the geographical and infrastructural features which seem to be sizeable barriers to trade in subsaharan Africa. Addressing this issue properly should permit quantifying the importance of geographical and infrastructural disadvantages. Against this background, this paper aims at assessing the importance of subsaharan African (SSA) countries' geographical and infrastructural disadvantages by focusing on their intra and extra regional trade flows. We limit our investigation to the WAEMU countries for which data on intra-regional trade and infrastructure is available, but we include their trade flows with OECD countries in order to take their impressive openness rate into account.

The remainder of the paper proceeds as follows. Some stylised facts on geographical and infrastructural disadvantages of the WAEMU are detailed in section 2 and an Armington-based model for the determination of trade flows is developed in section 3. In section 4, we discuss the econometric issues raised by the data we use. In section 5, we first estimate a traditional gravity model for the sake of comparison and carry out product-specific estimations, then we estimate the Armington-based model. The last section concludes.

2 Some stylised facts

The long experience in intra-regional cooperation in SSA makes WAEMU a good case study to consider issues related to South-South trade. The WAEMU was before 1994 a monetary union formed in 1963 to consolidate the common currency used within French colonies. During the 1990s, the drastic economic situation faced by these countries encouraged them to reinforce their solidarity in a deeper economic integration. Whether this region is an optimal currency area and how it may impact trade is not examined here (see Bénassy-Quéré and Coupet , 2003).

Figure 1: The West African Economic and Monetary Union

In the following, we will describe the geography of this union.

2.1 Road Infrastructures

The West African Economic and Monetary Union comprises five coastal (Benin, Cote d'Ivoire, Guinea-Bissau, Senegal and Togo) and three landlocked (Burkina Faso, Mali and Niger). More than three quarters of this area is located in the Sahel and two coastal countries (Senegal and Guinea-Bissau) are located at a distance from the other members (see Figure 1 above).

Roads is the main transportation infrastructure used for intra-regional trade (more than 90%[6]). The road network of the Union is 146,352 km long with only 14% paved. This network is unevenly distributed among members and is integrated in the whole West African roads network, which comprises three types of roads: the coastal roads linking coastal countries, the corridors linking landlocked countries to the sea, and the trans-sahel road from the border between Niger and Chad to Senegal. The coastal countries, representing 20% of the Union surface area, concentrate more than 70% of the Union roads. Table 2 below shows the road network distribution within the Union:

Table 2: Roads distribution throughout the WAEMU

Country	Roads	% paved	Density per 100 km^2
BEN	13842	9	10.8
BFA	13117	14	6.7
CIV	68351	8	17.0
MLI	14776	17	2.0
NER	13800	25	2.7
SEN	14358	29	21.1
TGO	8108	20	28.4
Union	146352	14	5.9

Sources: WAEMU commission

Cote d'Ivoire concentrates about half of the whole Union road-network and more than a quarter of paved roads, but Senegal has the second network with a better percentage of pavement (nearly 30%). Togo has the smallest road network but the highest road density (nearly 30 km of road per 100 km^2). The average road density of the Union is about 5.9 km per 100 km^2 and only 14% of the union road network is paved.

The Inter-State[7] roads network is 13,202 km long, of which 80% are paved. Nevertheless, the road linking Senegal to Mali is poorly paved (only 31% of pavement), a situation that practically isolates Senegal and Guinea Bissau from the other members of the Union in terms of land transport.

[6] Estimation of the transport department of WAEMU commission in 2001.

[7] That is highways between countries. Table 8 in Annexes gives a overview of these inter-state roads.

2.2 Border Infrastructures

The Union members have signed two multilateral conventions[8] to regulate and facilitate road transport and transit across borders. Despite these arrangements, limited border infrastructures might hinder the development of intra-regional traffic.

A recent survey [9] funded by the WAEMU Commission provided information on custom offices (suitable or not, joined or not, adjacent or not), weighbridges, radios, documentation on tax rate, typewriters, parking and stocking places.

To give an overview of these border infrastructure endowments of the Union, a score combining all the information available on each border equipment can be calculated. The method is rather crude: at a border, if the two customs offices possesses a given item of equipment or characteristic, the score is 2. If only one has it, the score is 1, and 0 if no one possesses it. These scores then add up to a percentage on a scale of border equipment, presented in Table 3 below:

Table 3: Border equipment and accessibility to some trading partners

Border	Economic Centers	Distance (km)	Road distance	% paved	Borders scores
CIV-BFA	Abidjan-Ouagadougou	832	1,176	100	39
CIV-MLI	Abidjan-Bamako	925	1,184	100	56
BEN-TGO	Cotonou-Lomé	160	189	100	33
TGO-BFA	Lome-Ouagadougou	757	970	100	44
MLI-SEN	Bamako-Dakar	1044	1,486	31	22
BFA-NER	Ouagadougou-Niamey	415	537	100	44
BFA-MLI	Ouagadougou-Bamako	705	610	100	44
NER-BEN	Niamey-Cotonou	785	1,041	100	33

Sources: WAEMU commission and our calculations.

On the basis of this scoring, it appears that only borders between Cote d'Ivoire and Mali, Togo and Burkina Faso, Burkina Faso and Niger and Burkina Faso and Mali are close to the 50% score. In addition, these countries are connected with paved roads. Table 4 also reveals an additional source of remoteness (apart from geographical distances) of Senegal and Guinea-Bissau[10] from the other members of the Union. Indeed, the score of the border between Mali and Senegal is the lowest (22%), and in addition only 31%[11] of the Senegal-Mali inter-state road is paved, a fact that adds to the isolation of Senegal and

[8] Referring to the document "Etude sur la facilitation du transport et du transit routier Inter-Etats" (1998), WAEMU Commission.

[9] "Rapport de synthèse préparatoire à la table ronde des bailleurs de fonds sur les infrastructures et le transport routier des Etats membres de l'UEMOA" (2000), WAEMU Commission.

[10] Note that these two countries are located at the far west of the Union (see Figure 1 above).

[11] Note for the sake of comparison that within the Union, 61% of the inter-state roads are paved on average.

Guinea-Bissau from the rest of the WAEMU countries. This border score is a very useful variable but since we do not have an evaluation for all the borders within the WAEMU, we cannot use it in the empirical estimations.

These facts on road and border infrastructure specify the geography of the WAEMU. The extensive form of the union and low quality of the transport infrastructure (few paved roads and poor border infrastructure) forebode high inland trade costs and thus lower intra-regional trade flows. In the following section, we will build a bilateral trade model and focus on geographical disadvantages in order to analyse the intra and extra trade of these southern countries. The model is derived from the Armington assumption of country-specific product and we obtain a structural model to be tested.

3 The model

Let us consider a two-region world: South and North. South represents a developing region (namely WAEMU countries) and North represents a developed region (namely OECD countries). We focus on the southern countries' import flows from all the trading partners, that is WAEMU and OECD countries. We thus consider South-South and South-North import flows. Southern countries are denoted by i, $i = 1, ..., I$ and Northern countries are denoted by k, $k = 1, ..., K$. According to the Armington assumption goods are differentiated by their origin. We also assume that within each country j, there are N_j representative firms producing the country-specific good. We assume a constant and non-unit elasticity of substitution between all the differentiated goods.

Let consider a southern country i importing from the other southern and northern countries. The representative consumer in this country has the following utility function:

$$U_i = \left(\sum_{j=1}^{I+K} \sum_{s=1}^{N_j} m_{ijs}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{1}$$

where m_{ijs} is the import of country i[12] from firm s in country j and σ is the elasticity of substitution between the traded good. The consumer problem is then to set his import for each differentiated good so as to maximise this utility function under his budget constraint:

$$Y_i = \sum_{j=1}^{I+K} \sum_{s=1}^{N_j} P_{ij} m_{ijs} \tag{2}$$

[12] Summing this quantity over the N_j representative firms yields $M_{ij} = \sum_{s=1}^{N_j} m_{ijs}$, the total import of country i from country j. Here, we only focus on the import flows M_{ij} and do not deal with the internal trade M_{ii} since we aim at describing only bilateral southern trade flows.

where Y_i is the income of the representative consumer in country i, P_{ij} is the price set by country j's firm in country i. $P_{ij} = P_j \tau_{ij}$ where P_j is the production price and τ_{ij} is an iceberg transport cost between countries i and j. This means that a firm producing in country j set a price P_j and the consumer in country i bears this price but also the cost (expressed in term of the imported good) required to ship this good from the production country to the import country. We derive the first order conditions of the maximization problem and combine them to obtain the following equation:

$$P_{ij} M_{ij} = \tau_{ij}^{1-\sigma} \left(\frac{P_j}{\left(\sum_{j=1}^{I+K} \sum_{s=1}^{N_j} \left(P_j \tau_{ij} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}} \right)^{1-\sigma} Y_i N_j \lambda^{-\sigma} \qquad (3)$$

which indicates a gravity type relation: the import value of country i from country j ($P_{ij} M_{ij}$) depends on the trade cost between these countries (τ_{ij}), the income of import country (Y_i) and the production level of the export country captured by the number of active firms (N_j) and a price term including characteristics of country j and those of the other trading partners (notice that λ is the Lagrange multiplier). We can simplify this equation by re-expressing equation (3) relative to a reference country, so as to cancel out the price term[13]. We use here France as reference country because of its historical ties with West African countries. This method will also correct any "colonization effect" in WAEMU imports from OECD countries.

Let us denote as E_{ij} bilateral trade values ($P_{ij} M_{ij} = E_{ij}$). Equation (3) becomes:

$$\frac{E_{ij}}{E_{iFRA}} = \left(\frac{\tau_{ij}}{\tau_{iFRA}} \right)^{1-\sigma} \left(\frac{P_j}{P_{FRA}} \right)^{1-\sigma} \left(\frac{N_j}{N_{FRA}} \right). \qquad (4)$$

The left-hand side of equation (4) represents country i's import from country j relatively to country i's import from France. The right-hand side represents three determinants of the relative bilateral trade: the relative transport costs, the relative production prices and the relative number of active firms in exporting countries. Equation (4) is the structural equation we will estimate. The next step is to define relevant proxies for these determinants of the relative import flows.

How can geographical impediments (captured by the relative transport costs) be measured? Since we are dealing here with intra and extra regional trade, geographical impediments to regional and extra-regional trade flows are specified as follows.

In the regional context, geographical impediments between two trading partners separated by a transit country can be due to four factors:

i) a border factor (extra borders have to be crossed), which can be proxied by the number of borders to be crossed by the shipped good;

[13] This method has been used by Head & Mayer (2000).

ii) a distance factor which can be proxied by the road distance between the two trading partners;

iii) a transit factor[14], which can be approximated by the road distance from the first border to the last border crossed by the imported good;

iv) an infrastructure factor, which can be estimated by the percentage of paved roads between the two trading partners.

The graph below summarises these impediments.

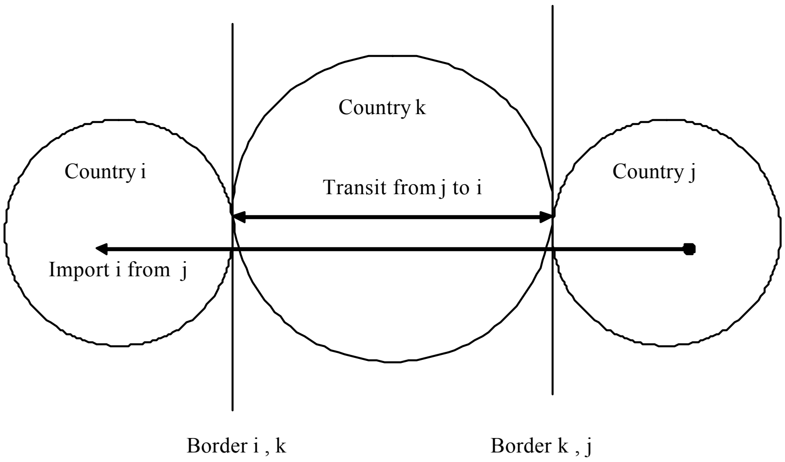

Figure 2: Measuring geographical impediments to intra-regional trade.

In extra-regional context, geographical impediments between a northern (j) and a southern (i) trading partners can be evaluated as follows:

i) the extra-regional distance (to be crossed by the imported good before reaching the developing region), which can be proxied by sea distance (SD_{ij}) for a coastal importer i, and by an average sea distance over all the southern coastal countries (δ_i) for a landlocked importer i;

ii) the inland distance (distance to be crossed by the imported good within the developing region), which is zero for a coastal importer i and can be proxied by the average road distance over all the southern coastal countries (κ_i) for a landlocked importer i. [15]

This computation for landlocked countries is justified by the fact that we do not know, from the database we use, which coastal country is used as transit

[14] For two contiguous southern countries, there is obviously no transit distance factor;

[15] using this average road distance to port but this remains statistically relevant. To recover this dimension, we constructed an adjusted road distance multiplying the average road distance to port by a remoteness coefficient computed as the distance of a landlocked country to a given OECD partner divided by its average distance to all his OECD partners, but this adjusted variable yields non-significant elasticities.

country hence, so that a better way to randomise the transit country being to use average distance over all the possible transit countries.

The following non-linear transport costs function takes the regional and the extra-regional context into account:

$$\tau_{ij} = SD_{ij}^{\alpha_1} \times RD_{ij}^{\alpha_2} \times V_{ij} \times e^{\varepsilon_{ij}}. \tag{5}$$

SD_{ij} denotes sea distance between countries i and j, RD_{ij} denotes road distance between countries i and j, ε_{ij} is a disturbance term taking into account all unobservable sources of trade costs and $V_{ij} = 1$ if the exporter country is an OECD country. For southern exporters, the following specification is adopted:

$$V_{ij} = e^{\beta_1 FRENCH + \beta_2 WAEMU} \times \%PR_{ij}^{\gamma_1} \times TRANSIT_{ij}^{\gamma_2} \times e^{\gamma_3 NBORDER_{ij}} \tag{6}$$

where $FRENCH$ is a dummy variable specifying French speaking partners, $WAEMU$ is a dummy variable specifying intra-regional trade, $\%PR_{ij}$ is the percentage of paved bilateral road, $TRANSIT_{ij}$ is the transit distance, that is the distance from the first border to the last border to be crossed by a shipped good and $NBORDER_{ij}$ is the number of borders to be crossed.

This transport cost function suggests that transport costs are non-linearly affected by sea and road distances between the two trading partners, and also by geographical and infrastructural characteristics of the importer located in the South. To complete the analysis of the impact of geographical impediments on trade flows, and in order to account for non-linearities, we include the squared variable $\%PR_{ij}$ in estimations reported in Table 13 in the Annexes.

How can we measure the relative price and the relative number of active firms in exporting countries? Since we only need an aggregate price reflecting the exporter production price, we use the GDP deflator (labelled as Π henceforth) of the exporter relatively to this price proxy for France, adjusted by a factor η: $\Pi_j^{\eta}/\Pi_{FRA}^{\eta}$. Introducing this factor allows trade elasticity with respect to the price proxy to depend on data rather than being constrained to be equal to unity. This variable is easy to get from the World Development Indicators database and is relevant to approximate an aggregate production price. We proxy the number of active firms in an exporting country (a variable which captures the production level of the exporter) by its GDP adjusted by a factor φ. Here again, the factor φ allows trade elasticity with respect to production level not to be constrained to one. Since we focus here on aggregate import and production flows, it sounds relevant to use the aggregate production of a country (its GDP) as a proxy the number of active firms within this country. Using these different specifications in equation 4 yields:

$$\frac{E_{ij}}{E_{iFRA}} = \left(\frac{SD_{ij}^{\alpha_1}}{SD_{iFRA}^{\alpha_1}} \frac{RD_{ij}^{\alpha_2}}{RD_{iFRA}^{\alpha_2}} \frac{V_{ij}}{V_{iFRA}} \right)^{1-\sigma} \left(\frac{\Pi_j^{\eta}}{\Pi_{FRA}^{\eta}} \right)^{1-\sigma} \left(\frac{GDP_j^{\varphi}}{GDP_{FRA}^{\varphi}} \right) \frac{e^{\varepsilon_{ij}}}{e^{\varepsilon_{iFRA}}}. \tag{7}$$

The relative road distance[16] can be simplified as follows:

i) when i is a coastal WAEMU country and j is an OECD country, $RD_{ij} = 1$, hence $RD_{ij}^{\alpha_2}/RD_{iFRA}^{\alpha_2} = 1$,

ii) when i is a landlocked WAEMU country and j is an OECD country, $RD_{ij}^{\alpha_2}/RD_{iFRA}^{\alpha_2} = \kappa_i^{\alpha_2}$ where κ_i is the average road distance to port defined above,

iii) when i is a coastal country and j is a WAEMU country, $RD_{ij}^{\alpha_2}/RD_{iFRA}^{\alpha_2} = RD_{ij}^{\alpha_2}$,

iv) when i is a landlocked country and j is a WAEMU country, $RD_{ij}^{\alpha_2}/RD_{iFRA}^{\alpha_2} = RD_{ij}^{\alpha_2}$.

Finally, we have to estimate the following version of the structural equation explaining relative bilateral imports by geographical, infrastructural, relative price and GDP variables:

$$\begin{aligned}\ln\left(\frac{E_{ij}}{E_{iFRA}}\right) &= (1-\sigma)\,\alpha_1 \ln\left(\frac{SD_{ij}}{SD_{iFRA}}\right) + (1-\sigma)\,\alpha_2 \ln\frac{RD_{ij}}{RD_{iFRA}} \\ &+ (1-\sigma)\,\beta_1 FRENCH + (1-\sigma)\,\beta_2 WAEMU \\ &+ (1-\sigma)\,\gamma_1 \ln \%PR_{ij} + (1-\sigma)\,\gamma_2 \ln TRANSIT_{ij} \\ &+ (1-\sigma)\,\gamma_3 NBORDER_{ij} + (1-\sigma)\,\eta \ln\left(\frac{\Pi_j}{\Pi_{FRA}}\right) \\ &+ \varphi \ln\left(\frac{GDP_j}{GDP_{FRA}}\right) + \xi_{ij}.\end{aligned}$$

In this equation, $\xi_{ij} = \varepsilon_{ij} - \varepsilon_{iFRA}$ represents the error term taking into account all the disturbance factors. In the following, we deal with econometric issues raised by the empirical implementation of this model.

4 Econometric issues

Firstly, let us mention that we do not use panel data estimations since the time span is short (three years) and the OLS with robust variance estimators estimations yield similar results. In this section we address three econometric issues relevant in our empirical estimations: missing dependent observations, censored regressions and instrumental variables estimations.

4.1 Missing dependent observations

Since we use COMTRADE data for the estimations, one problem arises: only four of the seven WAEMU countries[17] are reporter countries at the UN trade statistics. We can resort to mirror statistics when one of the trading partners is a reporter, but there no mirror statistics for two non-reporter countries. How can we deal with these missing dependent observations?

One approach is to simply ignore the missing observations. Since we have a sample of 596 observations, this will yield consistent estimators. But the ignored

[16] For simplicity, we set by assumption $RD_{iFRA} = 1$ and $V_{iFRA} = 1$.

[17] Benin, Niger, Senegal and Togo.

observations are useful since they concern South-South trade flows this paper is dealing with.

Another approach is to use the intra WAEMU intra-trade data to fill in the missing trade, but this yields a heterogeneity problem, since the observations of these two databases are seemingly different. However, we can combine the two data sources as follows: for the extra-regional trade flows, we use COMTRADE data and for the intra-regional trade we use WAEMU intra-trade data. We thus have a complete data set usable for estimations.

We can also use an econometric device to bypass this problem by trying to recover any helpful information from the incomplete COMTRADE database. Many papers[18] have addressed this topic of missing dependent observations. Greene (1997) discusses this issue, starting from a general econometric model:

$$Y = X\beta + \epsilon. \tag{8}$$

In this model, data are partitioned into two subsets: n_A complete observations and n_B observations for which Y is missing. Let $\hat{Y}_B$ be a predictor of Y_B from X_B. The least squares slope vector is:

$$b_f = \left\{ \binom{X_A}{X_B}' \binom{X_A}{X_B} \right\}^{-1} \binom{X_A}{X_B}' \binom{Y_A}{\hat{Y}_B}$$

This vector can be written as:

$$b_f = \{X_A' X_A + X_B' X_B\}^{-1} \left\{ X_A' Y_A + X_B' \hat{Y}_B \right\}. \tag{9}$$

Let b_A be the least squares slope in a regression that uses only the observations in group A, and define b_B likewise using $\hat{Y}_B$. We may then write:

$$\begin{aligned} b_f &= \{X_A' X_A + X_B' X_B\}^{-1} \{X_A' X_A b_A + X_B' X_B b_B\} \\ &= F b_A + (1 - F)\, b_B. \end{aligned}$$

In the last equation, $F = \{X_A' X_A + X_B' X_B\}^{-1} X_A' X_A$. This equation gives a matrix weighted average of the two least squares estimators, and we have:

$$E(b_f) = F\beta + (1 - F)\, E(b_B). \tag{10}$$

It appears that b_f will be unbiased only if b_B is unbiased as well. What is the best estimate of $\hat{Y}_B$? Kelejian (1969) assessed the efficiency of the so-called "first-order method" which uses $\hat{Y}_B = X_B b_A$, consisting in using the regressors obtained with the complete sample n_A to estimate $\hat{Y}_B$. This method passes the test of unbiasedness and appears to increase efficiency, even if we must account for the additional variation present in the predicted values.

To sum up, in the following empirical estimations, we use three sets of data:

[18] Afifi and Elashoff (1996, 1967), Haitovsky (1968), Anderson (1957), and Kelejian (1969) are a few of the major works.

i) only COMTRADE data,

ii) COMTRADE data for extra-regional trade and WAEMU intra-trade data for intra-regional trade,

iii) COMTRADE data and replace the missing dependent observations using the first-order method described above.

4.2 Censored regressions

In the empirical part of the paper, we try to estimate the determinants of product specific trade flows. A common feature of statistics at this detailed level is that low observations are set equal to zero. In the PC-TAS database (using COMTRADE statistics), the trade value must be at least 50$. This is a typical problem of censored observations and it is easy to prove that OLS are no longer relevant.

Indeed, let us consider the following model to be estimated:

$$Y^* = X\beta + \varepsilon \tag{11}$$

for which we do not observe (Y^*, X) but rather (Y, X) where:

$$Y = \max(0\ ,\ Y^*)\,. \tag{12}$$

It is such that:

$$\begin{aligned} E(Y \mid X) &= E(Y \mid X, Y = 0) \cdot P(Y = 0 \mid X) + E(Y \mid X, Y > 0) \cdot P(Y > 0 \mid X) \\ \Rightarrow E(Y \mid X) &= \{X\beta + E(\varepsilon \mid \varepsilon > -X\beta)\} \cdot P(\varepsilon > -X\beta)\,. \end{aligned}$$

Since $E[Y \mid X]$ is not a linear function of X, we cannot estimate β by OLS. One convenient way to solve such a model is to use maximum likelihood estimation. In STATA, the estimation is straightforward using the TOBIT estimation device. This is the way we estimate the product specific gravity model in Table 5, 9 and 10.

4.3 Instrumental variables estimation

The percentage of paved bilateral roads is designed to measure the quality of the journey between two trading partners. This variable is endogenous in the sense that trading partners with high GDP are likely to have more paved road and thus a higher percentage of paved bilateral road. To correct this endogeneity problem, we can use instrumental variable device. Empirically, adjusting the percentage of paved bilateral roads appears to be relevant in the following way:

$$\ln \%PR_{ij} = \alpha_1 \ln AREA_i + \alpha_2 \ln AREA_j + \alpha_3 \ln INFRA_{ij} + \nu_{ij} \tag{13}$$

where $INFRA_{ij}$ is the total length of paved road within countries i and j plus the lentgh of paved road between these two trading partners, $AREA_i$ and $AREA_j$ being the surface area of countries i and j. The variable $INFRA_{ij}$ is a measure of infrastructure quality between the two partners and reflects their level of GDP. We can then estimate the following system:

$$\begin{aligned}
\ln\left(\frac{E_{ij}}{E_{iFRA}}\right) &= (1-\sigma)\,\alpha_1 \ln\left(\frac{SD_{ij}}{SD_{iFRA}}\right) + (1-\sigma)\,\alpha_2 \ln\frac{RD_{ij}}{RD_{iFRA}} \\
&+ (1-\sigma)\,\beta_1 FRENCH + (1-\sigma)\,\beta_2 WAEMU \\
&+ (1-\sigma)\,\gamma_1 \ln \%PR_{ij} + (1-\sigma)\,\gamma_2 \ln TRANSIT_{ij} \\
&+ (1-\sigma)\,\gamma_3 NBORDER_{ij} + (1-\sigma)\,\eta \ln\left(\frac{\Pi_j}{\Pi_{FRA}}\right) \\
&+\varphi \ln\left(\frac{GDP_j}{GDP_{FRA}}\right) + \xi_{ij}.
\end{aligned}$$

$$\ln \%PR_{ij} = \alpha_1 \ln AREA_i + \alpha_2 \ln AREA_j + \alpha_3 \ln INFRA_{ij} + \nu_{ij}$$

This is a triangular equation system which can be easily estimated by two-stage least squares[19]. In the estimations, we use both OLS and two-stage least squares for the sake of comparison.

5 Estimations and results

In this section we try to quantify the impact of geographical and infrastructural disadvantages on the intra and extra regional trade of the WAEMU. Several data sources are mobilised: COMTRADE statistics, bilateral and internal paved roads from the WAEMU intra-trade and infrastructure statistics database[20]; the World Development Indicators, providing many macroeconomics aggregates and lastly geographical distance from the web site of Jon Haveman[21]. Since foreign trade statistics are missing for Guinea-Bissau, the eighth country of the WAEMU, we did not include this country in the sample. The time horizon is the period 1996-1998.

In the following, we examine how relevant are geographical variables to explain the puzzle referred to above. We start by reconsidering the results of a traditional gravity model, by sake of comparison. Then we proceed to the estimation of the Armington-based model we derived in section 3. The results stress the importance of properly modeling geographic determinants; when the proper specification is adopted, the impact of geography on southern trade flows is confirmed: among southern countries, it is worth paving road and reducing transit costs.

5.1 The traditional gravity model estimations

Table 5 below sticks on the traditional gravity model and tries to assess the role of the geographical variables in explaining the low level trade among WAEMU

[19] Note that since ν_{ij} and ξ_{ij} are assumed to be independent and equation (13) does not depend on the relative import flows, the derivation is straightforward.

[20] Source of these data: WAEMU commission. The database specifies clearly intra-WAEMU trade flows excluding any re-exportation flows.

[21] www.haveman.org. Alternative distance measures are provided on the CEPII website www.cepii.org.

countries. The dependent variable of these estimations is the import[22] of country i from country j ($\ln M_{ij}$). The regressors are the sea distance between countries i and j ($\ln SD_{ij}$, which is 0 if country j is a WAEMU country), the road distance between countries i and j ($\ln RD_{ij}$, which is 0 if country i is a coastal WAEMU country and country j an OECD country), a dummy variable specifying wether country j is a French speaking country[23] ($FRENCH$), a dummy variable specifying the WAEMU intra-regional trade ($WAEMU$), the GDP and GDP per capita of countries i and j ($\ln GDP_i$, $\ln GDP_j$, $\ln GDPPC_i$, $\ln GDPPC_j$), the percentage of paved inter-state road between country i and j ($\ln \%PR_{ij}$, for Paved Road between i and j), the transit distance between country i and j ($\ln TRANSIT_{ij}$[24]) and the number of borders to cross from country i to j ($NBORDER_{ij}$).

We estimate different specifications organised in two ways:

i) according to the database used: specifications 1 and 4 use only COMTRADE data, specifications 2 and 5 use COMTRADE data for extra-regional trade and WAEMU intra-trade data for intra-regional trade, specifications 3 and 6 use the database completed by the first-order method,

ii) according to the estimation method: specifications 1, 2 and 3 use OLS and specifications 4, 5 and 6 use two-stage least squares to correct the endogeneity problem of the percentage of paved bilateral road variable evoked above.

[22] Evaluated in current US $ value.

[23] We consider Switzerland, Belgium and Canada as French speaking countries.

[24] Note that this variable is set equal to 0 if countries i and j are contiguous. If they are not contiguous, this variable is measured as the road distance from the first to the last border to be crossed by the shipped good.

Table 5: The traditional gravity model estimation
Dependant variable: $\ln M_{ij}$
OLS with robust variance estimators

	1	2	3	4	5	6
LnSD_{ij}	-2.53***	-2.53***	-2.53***	-2.54***	-2.03***	-2.54***
	(-10.70)	(-10.68)	(-10.68)	(-10.74)	(-8.09)	(-10.73)
LnRD_{ij}	-0.65**	-0.65***	-0.75***	-0.96***	-0.59***	-0.97***
	(-2.56)	(-3.65)	(-4.14)	(-3.32)	(-2.86)	(-4.93)
FRENCH	1.40***	1.40***	1.40***	1.39***	-0.37	1.39***
	(9.10)	(9.12)	(9.13)	(9.10)	(-1.09)	(9.11)
WAEMU	-16.80***	-16.81***	-16.88***	-17.20***	-11.23***	-17.21***
	(-7.48)	(-7.71)	(-7.74)	(-7.63)	(-4.78)	(-7.88)
LnGDP_i	0.84***	0.84***	0.85***	0.86***	0.83***	0.86***
	(6.01)	(6.38)	(6.39)	(6.05)	(5.25)	(6.44)
LnGDP_j	1.39***	1.39***	1.39***	1.40***	1.08***	1.40***
	(22.07)	(22.39)	(22.24)	(22.29)	(14.31)	(22.59)
LnGDPPC_i	-0.90**	-0.90***	-0.91***	-0.88**	-0.35	-0.88**
	(-2.28)	(-2.59)	(-2.59)	(-2.21)	(-0.84)	(-2.51)
LnGDPPC_j	0.02	0.02	0.03	0.02	0.31***	0.01
	(0.23)	(0.22)	(0.31)	(0.19)	(4.01)	(0.19)
$\text{Ln\%PR}_{ij}$	0.88**	0.88***	1.06***	1.43***	0.99***	1.44***
	(2.16)	(3.02)	(3.58)	(2.96)	(2.88)	(4.34)
LnTRANSIT_{ij}	-0.09	-0.09	-0.10	-0.11	-0.16*	-0.11
	(-0.85)	(-1.04)	(-1.11)	(-0.94)	(-1.67)	(-1.17)
NBORDER_{ij}	0.21	0.21	0.27	0.35	-0.08	0.36
	(0.66)	(0.91)	(1.16)	(1.07)	(-0.26)	(1.48)
CONST	11.95***	11.97***	11.89***	11.66***	4.74	11.65***
	(3.74)	(4.08)	(4.04)	(3.64)	(1.45)	(3.97)
R^2	0.57	0.58	0.58	0.57	0.41	0.58
P-value	0.00	0.00	0.00	0.00	0.00	0.00
N	596	640	640	596	640	640

***represents a 99% level of significance
** represents a 95% level of significance
* represents a 90% level of significance

The estimations are globally significant, with R^2 statistics greater than 40%. The three first specifications are hardly different. The traditional gravity model variables are significant except for the GDP per capita of the exporter. A doubling of sea distance induces a 80%[25] reduction of imports of a coastal importer. For a landlocked WAEMU country, we have to take into account the inland distance crossed by the shipped good and thus a doubling of the total

[25] If we focus only on the sea distance variable, we have $LnM_{ij} = -2.53 \ln SD_{ij}$ which yields $M_{ij} = Dist_{ij}^{-2.53}$, so that a doubling of the distance implies: $M_{ij}^{*} = 2^{-2.53} Dist_{ij}^{-2.53} = 0.17 M_{ij}$, hence about 80% of trade reduction.

distance from an OECD country induces a 90% reduction of import flows. The dummy variables are also significant and bear the expected sign: a common language has a positive impact on trade flows and the intra-regional trade of the WAEMU countries is very low with regard to the extra-regional trade flows. Being a French-speaking exporter induces four times more import demand from WAEMU countries. Since the Armington-based model is supposed to "filter" any "colonization effect" of France, we will assess whether this result vanishes or not.

These specifications are slightly different when we focus on geographical variables: the data set using the first-order method to fill in missing observations shows that the percentage of paved bilateral road has a higher impact. It appears that a 10% increase of this percentage induces a 11%[26] increase in trade. The other geographical variables are not statistically significant.

The three following specifications (4, 5 and 6) instrument[27] the percentage of paved bilateral road with the surface of country i (and j if it is a WAEMU country) and the total paved road in and between countries i and j. These estimations provide two interesting results:

i) the impact of the percentage of paved bilateral road is higher,

ii) the transit distance makes for an additional impediment to trade.

Specification 6 in Table 5 indicates that a 10% increase of paved bilateral road induces a 15% increase in trade flows and specification 5 suggests that transit distance accounts for 6% of trade cost[28].

These estimations resulting from the traditional gravity model indicate a statistically significant effect of the traditional gravity model variables with the expected sign, except for the GDP per capita variable.

How these results are impacted by the nature of the shipped products is an important outcome. The COMTRADE data provide 2-digit trade flow statistics and we can use these disaggregated bilateral imports as the dependent variables. There are 99 2-digit product categories and for most of them, the import flows of WAEMU countries are very low, and cannot therefore yield robust estimations. To bypass this problem, we group these products into 14 industries following Fontagné, Freudenberg & Péridy (1997), as summarised in Table 9 in Annexes. For these product-specific estimations, it is not realistic to use GDP deflator as a price proxy. Besides, we do not fill in the missing observations for the non-reporter countries, because it would be too tricky to guess the specific products they are supposed to exchange with each other. Thus we only perform a gravity model based on the complete observations set (Table 10 and 11 in Annexes), instead of estimating the Armington-based model we derived in section 3. We

[26] If we focus only on the percentage of paved bilateral road, we have $LnM_{ij} = 1.06 \ln \%PR_{ij}$ which yields $M_{ij} = (\%PR_{ij})^{1.06}$, so that a 10% increase of this variable implies: $M_{ij}^{*} = 1.1^{1.06} (\%PR_{ij})^{1.06} = 1.11 M_{ij}$.

[27] When we regress the instrumented variables on all the instruments, the F-statistics is bigger than 10, which indicates that we do not have a "weak instruments" problem, as has been shown by Staiger & Stock (1997).

[28] In this specification, trade costs are due to sea distance (73%), road distance (21%) and transit distance (6%).

use a Tobit[29] estimation to take into account the low trade values censured to zero.

These disagreggated estimations are globally significant except for Other Transport Equipment industry. Focusing on the remaining 13 industries, the estimations provide interesting results illustrating the importance of geography in intra and extra WAEMU trade. The sea-distance reduction effect is unsurprisingly high for heavy products (Agriculture, forestry, mining...), while the road distance parameter estimate seems to capture first of all patterns of comparative advantages. The colonial ties seem to matter for non-agricultural raw materials and machinery. The bilateral geographical variables do not yield interesting results, except for the number of borders crossed by the shipped Leather-Textile goods within the WAEMU. Since this industry is the most concerned for intra-WAEMU trade, this last result seems to reveal that borders hinder trade within the Union.

5.2 The Armington-based model estimations

In this section, we turn to the testable form of the theoretical model derived in section 3, which is the main contribution of this paper. The dependent variable is relative import as described in section 3 and the regressors are those included in this final formulation:

$$\begin{aligned}\ln\left(\frac{E_{ij}}{E_{iFRA}}\right) &= (1-\sigma)\,\alpha_1 \ln\left(\frac{SD_{ij}}{SD_{iFRA}}\right) + (1-\sigma)\,\alpha_2 \ln\frac{RD_{ij}}{RD_{iFRA}} \\ &+ (1-\sigma)\,\beta_1 FRENCH + (1-\sigma)\,\beta_2 WAEMU \\ &+ (1-\sigma)\,\gamma_1 \ln \%PR_{ij} + (1-\sigma)\,\gamma_2 \ln TRANSIT_{ij} \\ &+ (1-\sigma)\,\gamma_3 NBORDER_{ij} + (1-\sigma)\,\eta \ln\left(\frac{\Pi_j}{\Pi_{FRA}}\right) \\ &+ \varphi \ln\left(\frac{GDP_j}{GDP_{FRA}}\right) + \xi_{ij}.\end{aligned}$$

As in the gravity model estimations, the specifications are organised in two ways:

i) according to the database used: specifications 1 and 4 use only COMTRADE data, specifications 2 and 5 use COMTRADE data for extra-regional trade and WAEMU intra-trade data for intra-regional trade, specifications 3 and 6 use the database completed by the first-order method,

ii) according to estimation method: specifications 1, 2 and 3 use OLS and specifications 4,5 and 6 use two-stage least squares to make up the endogeneity problem of the variable percentage of paved bilateral road.

The results are reported in Table 6 below.

Table 6: The Armington-based model estimations,
Dependant variable: $\text{Ln}\frac{E_{ij}}{E_{iFRA}}$
OLS and 2SLS with robust variance estimators

[29] See section 4.2 for theoretical justification.

	1	2	3	4	5	6
$\text{Ln}\frac{SD_{ij}}{SD_{iFR}}$	-2.51***	-2.52***	-2.51***	-2.50***	-2.51***	-2.50***
	(-10.77)	(-10.79)	(-10.80)	(-10.78)	(-10.79)	(-10.80)
$\text{Ln}\frac{RD_{ij}}{RD_{iFRA}}$	-0.64**	-0.26	-0.64***	-0.97***	-0.53***	-0.86***
	(-2.35)	(-1.50)	(-3.37)	(-3.20)	(-2.85)	(-4.25)
$\text{Ln}\frac{\Pi_j}{\Pi_{FR}}$	-0.99***	-0.71**	-0.99***	-1.03***	-0.74***	-1.01***
	(-3.18)	(-2.55)	(-3.23)	(-3.26)	(-2.64)	(-3.29)
$\text{Ln}\frac{GDP_j}{GDP_{FRA}}$	1.33***	1.37***	1.33***	1.34***	1.37***	1.33***
	(23.99)	(24.85)	(24.58)	(24.05)	(24.96)	(24.61)
Ln%PR$_{ij}$	1.20***	0.65**	1.21***	1.78***	1.11***	1.59***
	(2.77)	(2.22)	(3.99)	(3.49)	(3.41)	(4.59)
LnTRANSIT$_{ij}$	-0.18	-0.22**	-0.18*	-0.19	-0.23**	-0.18*
	(-1.53)	(-2.31)	(-1.92)	(-1.58)	(-2.38)	(-1.93)
NBORDER$_{ij}$	0.34	0.05	0.34	0.49	0.15	0.42*
	(1.08)	(0.18)	(1.49)	(1.48)	(0.50)	(1.76)
FRENCH	1.08***	1.12***	1.08***	1.07***	1.11***	1.07***
	(5.84)	(6.05)	(5.84)	(5.81)	(6.03)	(5.83)
WAEMU	-17.28***	-17.12***	-17.28***	-17.58***	-17.33***	-17.45***
	(-7.68)	(-7.77)	(-8.04)	(-7.78)	(-7.89)	(-8.10)
CONST	-1.54***	-1.52***	-1.54***	-1.54***	-1.52***	-1.54***
	(-13.36)	(-13.07)	(-13.54)	(-13.31)	(-13.14)	(-13.57)
R^2	0.52	0.53	0.53	0.52	0.53	0.53
P-value	0.00	0.00	0.00	0.00	0.00	0.00
N	573	617	617	573	617	617

***represents a 99% level of significance

** represents a 95% level of significance

* represents a 90% level of significance

The estimations are globally significant with R^2 statistics greater than 50% and the distance, GDP and WAEMU variables yield coefficients similar to those from the traditional gravity model estimation. Here again, we do not detect a weak instrument problem when performing specifications 4, 5 and 6. The common language effect decreases from four to three times more trade between French speaking partners, indicating a correction of the French colonization effect over these developing countries. Here again, the border variables does not yield significant coefficients. In the following, we will focus on the relative price variable, the percentage of paved bilateral road and the transit distance that yield significant and interesting results.

The first interesting result is the substitution effect captured by the relative GDP deflator Π_j/Π_{FRA}. Indeed, specification 6 indicates that if an exporter price double relatively to price in France, the importer reduces its imports from this country by 70%.

The second result is the positive return of paved bilateral road on trade flows. This is the key variable measuring the quality of roads in this paper since about 90% of the intra-regional trade is by roads. The positive and statistically significant sign of the coefficient of this variable indicates that paving an extra portion of a road between two trading partners increases their bilateral trade flows. For the inter-state roads not totally paved, we can use the elasticity of this variable to compute the extra import flows created when the percentage of pavement is completed to 100%. [30] In addition, we use the elasticity of the variable percentage of paved bilateral road obtained with specification 6 which is econometrically more accurate because correcting for the endogeneity problem of this variable and also using the so called "first-order method" to replace the missing dependent observations. Not surprisingly, the results presented in Table 12 in Annexes indicate that the lower the percentage of paved bilateral road, the higher the impact of this infrastructure improvement on the import flows. The trading partners most concerned are Mali and Senegal. Indeed, for the year 1998, the simulations indicate that improving the inter-state road between these partners from 31% to 100% paving can increase trade between them more than three times. This seems to be a big issue for the Union, because Senegal is the second most dynamic economy after Cote d'Ivoire, and its remoteness from the other members tends to weaken the Union economy. Moreover, this remoteness also affects trade flows between Senegal and Cote d'Ivoire. Indeed, the simulations suggest that a 100% pavement of the road between these two countries could double trade flows between them. If we take into account all the extra trade created by this "100% paving of inter-state roads" infrastructure policy, trade flows in this region are 2.87 times higher, a figure that does make sense if we recall that the intra-regional trade flows for this Union was only of 3% during the 90s.

The third result is the additional cost due to transit distance measured by the distance from the first to the last border to be crossed by the shipped good. The negative and statistically significant effect of this variable confirms the idea that crossing a transit country yields extra trade costs independently from the distance between exporting and importing country. Indeed, doubling this variable induces 15% less trade which correspond to 6% of trade costs estmated by the model, an effect which adds to the traditional distance effect. This variable proves thus to be a good proxy for internal geographical impediments of transit countries.

To complete this analysis, we also consider two additional factors (Table 13 in Annexes): export diversification/concentration and non-linear impact of paved bilateral road. First we analyse the impact of export concentration of WAEMU countries on the low level of trade observed between them. Indeed, if these countries export only agricultural raw materials dedicated to developed countries, their bilateral trade will obviously be low.[31] To assess such effect, we

[30] claim an over-estimation of trade flow when using this elasticity of the variable percentage of paved bilateral road which takes into account extra and intra regional trade flows to simulate intra-regional trade flows. However in the specifications, we include an intra-WAEMU trade dummy variable which captures all effects specific to intra but also extra trade. Thus, using this elasticity for simulation is relevant.

[31] We are indebted to Sébastien Jean and Thierry ayer for suggesting such explanation

add an herfindhal sectorial concentration index of the most exported product of each WAEMU country using ITC[32] Trade Performance Index. The estimations yield non statistically significant coefficients for this variable indicating that export concentration has no statistical impact on intra-WAMEU trade flows. Second, we explore a non-linear impact of the percentage of paved bilateral roads on trade flows using the term $\ln \%PR_{ij} + (\ln \%PR_{ij})^2$. The estimations yields no statistically significant coefficients indicating that introducing only the variable $\ln \%PR_{ij}$ like in Table 6 is a relevant way to assess the impact of this variable on trade flows.

6 Conclusion

In this paper, we aimed at analyzing the impact of geography on South-South trade, starting with the puzzle indicating a global disadvantage faced by landlocked countries, and particularly developing ones. We focused on the integrated countries of the West African Economic and Monetary Union for which suitable data for such an analysis is available.

The traditional gravity model estimates confirmed the statistically significant effect of sea distance, road distance and GDP of the trading partners on trade flows. Being a French-speaking country induces four times more import demand and increasing the percentage of paved bilateral road leads to higher trade flows. Shipping goods through a transit country also proves to yield additional trade costs, accounting for 6% of the trade costs estimated in the model. The industry-specific estimations provide interesting additional results, two of them being most appealing. First, colonial ties seem to matter for non-agricultural raw materials and machinery trade. Second, it appears that the leather and textile industry which mainly concerns intra-regional trade faces a strong border impediment revealing a weakness of the integration process.

The estimations from the Armington-based model provide three interesting results and emphasize the role of geographical determinants. First, the paved bilateral road return on trade flows is confirmed and reinforced. If all the interstate roads were paved, the countries would trade 2.87 times more than what is observed. We can now answer to our initial question, as whether there is an untapped South-South trade potential, given remoteness, economic size and eventually landlocking of the countries in the region. The answer is yes, there is an untapped potential for roads pavement projects. Second, transit distance prove to be an additional impediment to trade, indicating that the internal geography of the transit countries matters.

The main aim of this paper was to estimate to what extent geographical disadvantages are a handicap for South-South trade. We focused on the West African Economic and Monetary Union, but the results could be extended to other southern regions. Two types of disadvantages seem to affect these countries: one due to their location in a poor southern area and one due to the higher impediments when crossing transit zones within this area. Beyond this

[32] International Trade Center UNCTAD/WTO, www.intracen.org.

result, this paper proposes an alternative way of analysing the determinants of trade flows in southern areas by using an Armington-based model and specific definitions of geographical impediments. Applying such methodology to other geographical areas, different databases and proxies of these geographical impediments remains on the research agenda of trade economists.

References

[1] Afifi T., Elashoff R. (1966), "Missing observations in multivariate statistics I", *Journal of the American Statistical Association*, 61, pp. 595-604.

[2] Afifi T., Elashoff R. (1967), "Missing observations in multivariate statistics II", *Journal of the American Statistical Association*, 62, pp. 10-29.

[3] Amjadi A., Yeats A. (1995), "Have transport costs contributed to the relative decline of African exports? Some preliminary evidence", *World bank working paper* n°1559.

[4] Anderson T. (1957), "Maximum likelihood estimates for a normal distribution when some observations are missing", *Journal of the American Statistical Association*, 52, pp. 200-203.

[5] Benassy-Quéré A., Coupet M. (2003), "On the Adequacy of Monetary Arrangements in Subsaharan Africa", *forthcoming in The World Economy.*

[6] Fontagné L., Pajot M., Pasteels J-M. (2002), "Potentiels de commerce entre économies hétérogènes: un petit mode d'emploi des modèles de gravité", *Economie et Prévision.*, (152-153):115-139.

[7] Fontagné L., Freudenberg M., Péridy N. (1997), "Trade patterns inside the single market", *Document de travail du CEPII*, n°97 − 07 avril 1997.

[8] Foroutan F., Pritchett L. (1993), "Intra-subsaharan African Trade: Is it too little?", *Journal of African Economies*; 2(1): 74-105.

[9] Greenaway D., Milner C. (1990), "South-South trade: theory, evidence and policy", *The world Bank Research Observer*, 5(1): 47-68.

[10] Greene W. H. (1997), *Econometric Analysis*, Prentice Hall Press: USA.

[11] Haitovsky Y. (1968), "Missing data in regression analysis", *Journal of the Royal Statistical Society*, Series B, pp. 67-82.

[12] Havrylyshyn O. (1985), "The Direction of Developing Country Trade: Empirical Evidence of Differences Between South-South and South-North Trade", *Journal of Development Economics*, 19(3):255-281

[13] Head K., Mayer T. (2000), "Non-Europe: the magnitude and causes of market fragmentation in the EU", *Weltwirtschaftliches Archiv*, 136(2): 284-314.

[14] Henderson J. V., Shalizi Z., Venables A. J. (2001), "Geography and development", *Journal of Economic Geography*, 1:81-105.

[15] Kelejian H. (1969), "Missing observations in multivariate regression- efficiency of a first order method", *Journal of the American Statistical Association*, 64, pp. 1609-1616.

[16] Limao N., Venables A. J. (2000), "Infrastructure, geographical disadvantage, transport costs and trade ", *World Bank Economic Review*,15(3):451-79.

[17] Redding S., Venables A. J. (2001), "Economic geography and international inequality", *CEPR discussion paper* n° 2568.

[18] Sachs J. (2001), "Tropical underdevelopment", *NBER working paper* n° 8119.

[19] Silverman B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman & Hall.

[20] Staiger D. Stock J. H. (1997), "Instrumental Variable Regression with Weak Instruments", *Econometrica*, 65: 557-586.

[21] The Economist print edition (2002), "The Road to Hell is Unpaved", 19th December 2002, Douala

ANNEXES

Table 7: The disadvantage of landlocked countries: a gravity model approach
Dependent variable: $Ln\,(Export98_{ij})$
TOBIT estimations

	1	2	3	4	5
$LnDIST_{ij}$	-1.21***	-1.26***	-1.24***	-1.24***	-1.23***
	(-29.18)	(-30.25)	(-29.38)	(-29.38)	(-29.31)
$LnGDP_i$	1.16***	1.15***	1.15***	1.15***	1.16***
	(52.27)	(52.34)	(52.38)	(52.35)	(52.82)
$LnGDP_j$	0.85***	0.83***	0.83***	0.83***	0.85***
	(41.47)	(40.98)	(41.01)	(41.00)	(42.14)
$LnGDPPC_i$	0.24***	0.16***	0.17***	0.17***	0.17***
	(8.95)	(5.73)	(5.88)	(-5.86)	(5.89)
$LnGDPPC_j$	0.16***	0.06**	0.07**	0.07**	0.07**
	(6.13)	(2.21)	(2.40)	(2.38)	(2.44)
$CONTIG_{ij}$	1.08***	0.89***	0.85***	0.88***	0.84***
	(5.19)	(4.31)	(4.09)	(4.26)	(4.08)
$LANG_{ij}$	0.85***	0.82***	0.82***	0.81***	0.84***
	(8.62)	(8.40)	(8.24)	(8.29)	(8.59)
1LLE	-0.32***	-0.40***	-0.39***	-0.39***	
	(-2.78)	(-3.46)	(3.42)	(-3.42)	
1LLNE	-0.50***	-0.41***	-0.42***	-0.41***	
	(-5.50)	(-4.58)	(-4.71)	(-4.58)	
2LLNE	1.15*		1.07*		
	(1.81)		(1.70)		
1AFR		-0.93***	-0.93***	-0.94***	-0.95***
		(-12.75)	(-12.79)	(-12.81)	(-12.96)
2AFR			0.16	0.18	0.20
			(1.04)	(1.20)	(1.32)
CONST	-6.26***	-3.79***	-4.01***	-3.99***	-4.47***
	(-13.90)	(-7.84)	(-7.85)	(-7.81)	(-8.84)
Pseudo-R^2	0.15	0.16	0.16	0.16	0.15
P-value	0.00	0.00	0.00	0.00	0.00
N	7,825	7,825	7,825	7,825	7,825

***represents a 99% level of significance

** represents a 95% level of significance

* represents a 90% level of significance

Notes: DIST for geographical distance, GDPPC for GDP per capita, CONTIG for contiguity, LANG for common langauge, 1LLE for one European landlocked partner, 1LLNE for one non-European landlocked partner, 2LLNE for two non-European landlocked partners, 1AFR for one African partner and 2AFR for two African partners.

Table 8: Bilateral paved road within WAEMU (in km)

Partners	Road distance	% paved
BEN-BFA	1,022	55
BEN-CIV	568	100
BEN-MLI	1,552	100
BEN-NER	1,041	100
BEN-SEN	3,038	69
BEN-TGO	189	100
BFA-CIV	1,176	100
BFA-MLI	610	100
BFA-NER	537	100
BFA-SEN	2,016	57
BFA-TGO	970	100
CIV-MLI	1,184	100
CIV-NER	1,609	100
CIV-SEN	2,634	62
CIV-TGO	588	100
MLI-NER	1,423	80
MLI-SEN	1,486	31
MLI-TGO	1,500	100
NER-SEN	2,909	65
NER-TGO	1,507	100
SEN-TGO	2,986	68

Sources: WAEMU commission

Table 9: Aggregating the 2-digit products by industry

Industry	2-digits products	label
AA	01-14	Agriculture, Hunting, Forestry
AB	15-24	Food, Beverages, Tobacco
B	25-27	Mining, Quarrying, Oil
CD	28-40	Chemicals
E	44-49	Wood, Paper, Printing
FD	41-43, 50-67	Textile, Leather
G	68-72	Non-metallic mineral products
HI	73-83	Basic metals and Manufactured metal products
JA	84	Non-electrical machinery
JB	85	Electrical machinery
KA	87	Motor vehicles
KB	86, 88, 89	Other transport equipment
LA	90-92	Professional goods
N	93-99	Other industries

Source: Fontagné, Freudenberg and Péridy (1997)

Table 10: TOBIT estimations of the bilateral imports by industry
Dependant variable: $\ln M_{ij}^{Ind}$

	AA	AB	B	CD	E	FD	G
LnSD_{ij}	-2.36	-4.60***	-3.75***	-1.27***	-0.14	-0.19	-0.65
	(-3.87)	(-8.04)	(-4.77)	(-3.43)	(-0.33)	(-0.53)	(-1.43)
LnRD_{ij}	0.03	0.36	0.54	0.05	0.93***	-0.36	0.74*
	(0.06)	(0.77)	(0.87)	(0.13)	(2.61)	(-1.03)	(1.64)
FRENCH	0.14	-0.56	1.58**	0.20	0.23	0.29	1.40***
	(0.21)	(-1.07)	(2.15)	(0.51)	(0.61)	(0.73)	(2.79)
WAEMU	-15.46***	-31.93***	-22.85***	-8.75	3.19	6.06*	-3.30
	(-2.67)	(-5.91)	(-3.21)	(-2.46)	(0.81)	(1.71)	(-0.76)
LnGDP_i	0.48***	0.32	0.93	1.86***	0.87*	-0.29	1.29**
	(0.63)	(0.50)	(1.10)	(3.86)	(1.83)	(-0.58)	(2.25)
LnGDP_j	1.07	0.76	-0.14	-0.25	-0.17	-1.10	0.09
	(0.85)	(0.73)	(-0.11)	(-0.33)	(-0.22)	(-1.44)	(0.09)
LnGDPPC_i	0.78***	0.98***	0.45**	0.84***	0.12	0.74***	0.29*
	(4.26)	(6.38)	(1.97)	(6.67)	(0.94)	(5.81)	(1.87)
LnGDPPC_j	0.50**	0.23	0.53**	0.45***	0.69***	0.22	0.58***
	(2.30)	(1.29)	(2.01)	(2.85)	(3.81)	(1.52)	(3.34)
$\text{Ln\%PR}_{ij}$	0.66	-0.71	-1.72*	-0.39	-1.74***	-0.38	-0.87
	(0.84)	(-0.91)	(-1.93)	(-0.69)	(-3.49)	(-0.68)	(-1229)
LnTRANSIT_{ij}	-0.39	0.04	0.27	-0.04	0.09	0.57***	-0.30
	(-1.56)	(0.17)	(0.97)	(-0.19)	(0.54)	(3.22)	(-1.12)
NBORDER_{ij}	-0.33	-0.97	-0.87	1.17**	-0.79*	-2.06***	-0.26
	(-0.43)	(-1.43)	(-1.10)	(2.22)	(-1.71)	(-3.83)	(-0.44)
CONST	-4.88	24.76***	16.81***	-5.83	-9.94	3.48	-2.03
	(-0.51)	(3.00)	(1.55)	(-0.96)	(-1.56)	(0.58)	(-0.27)
Pseudo-R^2	0.03	0.06	0.07	0.06	0.05	0.07	0.04
P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	392	401	306	476	375	417	334

***represents a 99% level of significance
** represents a 95% level of significance
* represents a 90% level of significance

Table 11: TOBIT estimations of the bilateral imports by industry (continued)
Dependant variable: $\ln M_{ij}^{Ind}$

	HI	JA	JB	KA	KB	LA	N
LnSD_{ij}	-1.48***	-0.79***	-0.24	-0.27	-0.26	-0.94***	-1.04***
	(-3.41)	(-2.88)	(-0.78)	(-0.79)	(-0.21)	(-2.93)	(-2.63)
LnRD_{ij}	-0.11	0.38	0.50	0.81*	4.98	-3.00	-0.24
	(-0.26)	(0.90)	(1.37)	(1.85)	(0.63)	(-0.30)	(-0.35)
FRENCH	1.18***	0.84***	0.98***	1.10***	0.43	0.03	-0.20
	(2.69)	(2.79)	(3.07)	(3.23)	(0.46)	(-0.11)	(-0.51)
WAEMU	-9.41**	-0.52	-0.90	0.03	-	-	-6.14
	(-2.21)	(-0.17)	(-0.29)	(0.01)	-	-	(-1.55)
LnGDP_i	1.63***	0.64*	0.62	0.90**	0.85	1.57***	0.42
	(3.07)	(1.65)	(1.49)	(2.01)	(0.64)	(3.39)	(0.78)
LnGDP_j	-1.76**	1.06*	0.27	-0.73	1.65	0.04	-0.15
	(-1.99)	(1.65)	(0.40)	(-0.99)	(0.64)	(0.05)	(-0.17)
LnGDPPC_i	0.49***	0.82***	0.57***	0.73***	0.25	0.44***	0.82***
	(3.45)	(8.80)	(5.21)	(6.25)	(0.88)	(4.18)	(6.12)
LnGDPPC_j	0.22	0.58***	0.35**	0.58***	0.45	0.34**	0.65***
	(1.40)	(4.63)	(2.44)	(4.02)	(0.86)	(2.17)	(3.36)
$\text{Ln\%PR}_{ij}$	-0.48	-1.13	-0.73	-1.16*	3.13	-4.22	1.08
	(-0.78)	(-1.60)	(-1.34)	(-1.79)	(0.34)	(-0.54)	(0.97)
LnTRANSIT_{ij}	0.38*	0.25	-0.15	-0.28	-6.70	5.36	-0.39
	(1.86)	(1.04)	(-0.85)	(-1.20)	(-0.51)	(0.38)	(-1.30)
NBORDER_{ij}	-0.36	-1.89***	0.43	-0.19	-4.10	-2.14	0.08
	(-0.65)	(-2.94)	(0.92)	(-0.24)	(-0.57)	(-0.35)	(0.13)
CONST	11.61*	-17.68***	-12.74**	11.72***	-14.42	-3.33	-6.60
	(1.69)	(-3.67)	(-2.43)	(-6.92)	(-0.67)	(-0.56)	(-0.91)
Peseudo-R^2	0.06	0.12	0.05	0.07	0.01	0.06	0.07
P-value	0.00	0.00	0.00	0.00	0.55	0.00	0.00
N	378	462	433	408	188	318	371

***represents a 99% level of significance
** represents a 95% level of significance
* represents a 90% level of significance

Table 12: Extra 1998 import flows when the % of paved bilateral roads is raised to 100[33] (units: 1,000$)

Country i	Country j	$\%PR_{ij}$	ΔM_{ij}	M_{ij}	$\frac{\Delta M_{ij}}{M_{ij}}$ (%)
BEN	BFA	55	1,218	936	130
BEN	SEN	69	6,247	8,745	71
BFA	BEN	55	226	174	130
BFA	SEN	57	4,033	3,362	120
CIV	SEN	62	13,855	14,217	97
MLI	NER	80	212	533	40
MLI	SEN	31	105,290	29,751	354
NER	MLI	80	1053	1,081	97
NER	SEN	65	735	858	86
SEN	BEN	69	103	144	71
SEN	BFA	57	71	59	120
SEN	CIV	62	23,932	24,558	97
SEN	MLI	31	10,713	3,027	354
SEN	NER	65	9	11	86
SEN	TGO	68	89	119	75
TGO	SEN	68	2,508	3,352	75
Total			170,293	90,927	187

Sources: WAEMU Commission and our calculations.

[33] In fact, we have $\Delta M_{ij} = 1.59 \times \frac{\Delta \% PR_{ij}}{\% PR_{ij}} \times M_{ij}$, using the estimated coefficient of $\%PR_{ij}$ in specification 6 of Table 10 which is econometrically well specified.

Table 13: The Armington-based model estimations,
Dependent variable: $\text{Ln}\frac{E_{ij}}{E_{iFRA}}$
2SLS with robust variance estimators

	1	2	3	4
$\text{Ln}\frac{SD_{ij}}{SD_{iFR}}$	-2.50***	-2.51***	-2.52***	-2.52***
	(-10.74)	(-10.76)	(-10.77)	(-10.77)
$\text{Ln}\frac{RD_{ij}}{RD_{iFRA}}$	-0.97***	-0.36	-0.53***	0.40***
	(-3.19)	(-0.41)	(-2.87)	(0.64)
$\text{Ln}\frac{\Pi_j}{\Pi_{FR}}$	-1.03***	-1.03***	-0.74***	-0.73***
	(-3.26)	(-3.25)	(-2.64)	(-2.61)
$\text{Ln}\frac{GDP_j}{GDP_{FRA}}$	1.34***	1.34***	1.37***	1.38***
	(23.93)	(23.53)	(24.79)	(24.98)
Ln%PR$_{ij}$	1.78***	-2.20	1.14***	-5.04
	(3.47)	(-0.42)	(3.49)	(-1.29)
$(Ln\%PR_{ij})^2$		0.68		1.05
		(0.75)		(1.59)
LnTRANSIT$_{ij}$	-0.18	-0.20	-0.24**	-0.24**
	(-1.53)	(-1.54)	(-2.46)	(-2.35)
NBORDER$_{ij}$	0.48	0.37	0.17	-0.06
	(1.46)	(1.03)	(0.56)	(-0.18)
LnHerfindhal	0.03	0.03	-0.08	-0.08
	(0.28)	(0.25)	(-0.86)	(-0.81)
FRENCH	1.07***	1.07***	1.11***	1.11***
	(5.83)	(5.91)	(5.95)	(7.56)
WAEMU	-17.28***	-17.34***	-17.43***	-16.94***
	(-7.66)	(-7.60)	(-7.85)	(-7.56)
CONST	-1.51***	-1.50***	-1.59***	-1.58***
	(-11.03)	(-10.84)	(-11.56)	(-11.36)
R^2	0.52	0.52	0.53	0.51
P-value	0.00	0.00	0.00	0.00
N	573	573	573	573

***represents a 99% level of significance
** represents a 95% level of significance
* represents a 90% level of significance

Note: For all these specifications, $\ln\%PR_{ij}$ is instrumented by $\ln area_i$, $\ln area_j$ and $\ln INFRA_{ij}$. Specifications 1 and 2 use only COMTRADE data and specifications 3 and 4 use COMTRADE data for extra-regional trade and WAEMU intra-trade data for intra-regional trade. The non-linear term $(\ln\%PR_{ij})^2$ is instrumented by the same instruments than $\ln\%PR_{ij}$.

Table 14: Exporters sample for the estimations in table 7

Code	Country	Code	Country
ARG	Argentina	MDG	Madagascar
AUS	Australia	MEX	Mexico
AUT	Austria	MLI	Mali
BAR	Barbados	MUS	Mauritius
BEN	Benin	MYS	Malaysia
BFA	Burkina Faso	NER	Niger
BGD	Bangladesh	NGA	Nigeria
BOL	Bolivia	NIC	Nicaragua
BRA	Brazil	NLD	Netherlands
CAN	Canada	NOR	Norway
CHL	Chile	NPL	Nepal
CIV	Cote d'Ivoire	NZL	New Zealand
COL	Colombia	OMN	Oman
CRI	Costa Rica	PAN	Panama
CYP	Cyprus	PER	Peru
CZE	Czech Rp	PHL	Philippines
DEU	Germany	POL	Poland
DNK	Denmark	KOR	Korea Rp
DZA	Algeria	PRT	Portugal
ECU	Ecuador	PRY	Paraguay
EGY	Egypt	ROM	Romania
ESP	Spain	RUS	Russia
FIN	Finland	SAU	Saudi Arabia
FRA	France	SDN	Sudan
GBR	UK	SEN	Senegal
GHA	Ghana	SGP	Singapore
GHA	Ghana	SLV	El Salvador
GMB	Gambia	SWE	Sweden
GNB	Guinea-Bissau	TGO	Togo
GRC	Greece	THA	Thailand
GTM	Guatemala	TTO	Trinidad Tbg
HKG	Hong kong	TUN	Tunisia
HND	Honduras	TUR	Turkey
HUN	Hungary	TZA	Tanzania
ICE	Iceland	UGA	Uganda
IDN	Indonesia	URY	Uruguay
IND	India	USA	USA
IRL	Ireland	VEN	Venezuela
ISR	Israel	YUG	Yugoslavia
ITA	Italy	ZAF	South Africa
JOR	Jordan	MAR	Morocco
JPN	Japan		
MAL	Malta		

Table 15: Importers sample for the estimations in table 7

Code	Country	Code	Country	Code	Country
AFG	Afghanistan	GHA	Ghana	OMN	Oman
AGO	Angola	GIN	Guinea	PAK	Pakistan
ARE	Untd Arab Em	GMB	Gambia	PAN	Panama
ARG	Argentina	GNB	Guinea-Bissau	PER	Peru
AUS	Australia	GRC	Greece	PHL	Philippines
AUT	Austria	GTM	Guatemala	PNG	Papua NG
BAH	Bahamas	GUY	Guyana	POL	Poland
BAR	Barbados	HKG	Hong kong	KOR	Korea Rp
BDI	Burundi	HND	Honduras	PRT	Portugal
BEL	Belgium-Lux	HTI	Haiti	PRY	Paraguay
BEN	Benin	HUN	Hungary	QAT	Qatar
BFA	Burkina Faso	ICE	Iceland	REU	Reunion
BGD	Bangladesh	IDN	Indonesia	ROM	Romania
BGR	Bulgaria	IND	India	RUS	Russia
BHA	Bahrain	IRL	Ireland	RWA	Rwanda
BHU	Bhutan	IRN	Iran	SAU	Saudi A
BLZ	Belize	IRQ	Iraq	SDN	Sudan
BOL	Bolivia	ISR	Israel	SEN	Senegal
BRA	Brazil	ITA	Italy	SEY	Seychelles
CAF	Central A R	JAM	Jamaica	SGP	Singapore
CAN	Canada	JOR	Jordan	SLE	Sierra Leone
CHE	Switzerland	JPN	Japan	SLI	Solomon I
CHL	Chile	KEN	Kenya	SLV	El Salvador
CHN	China	KWT	Kuwait	SOM	Somalia
CIV	Cote d'Ivoire	LAO	Laos	SUR	Suriname
CMR	Cameroon	LBR	Liberia	SWE	Sweden
COG	Congo	LKA	Sri Lanka	SYR	Syrn A R
COL	Colombia	MAL	Malta	TCD	Chad
COM	Comoros	MAR	Morocco	TGO	Togo
CRI	Costa Rica	MDG	Madagascar	THA	Thailand
CYP	Cyprus	MEX	Mexico	TTO	Trinidad
CZE	Czech Rp	MLI	Mali	TUN	Tunisia
DEU	Germany	MMR	Myanmar	TUR	Turkey
DJI	Djibouti	MNG	Mongolia	TWN	Taiwan
DNK	Denmark	MOZ	Mozambique	TZA	Tanzania
DOM	Dominican	MRT	Mauritania	UGA	Uganda
DZA	Algeria	MUS	Mauritius	URY	Uruguay
ECU	Ecuador	MWI	Malawi	USA	USA
EGY	Egypt	MYS	Malaysia	VEN	Venezuela
ESP	Spain	NER	Niger	YEM	Yemen
ETH	Ethiopia	NGA	Nigeria	YUG	Yugoslavia
FIJ	Fiji	NIC	Nicaragua	ZAF	South Africa
FIN	Finland	NLD	Netherlands	ZAR	Congo D R

FRA	France	NOR	Norway	ZMB	Zambia
GAB	Gabon	NPL	Nepal	ZWE	Zimbabwe
GBR	UK	NZL	New Zealand		

Table 16: OECD countries included in the database

AUSTRALIA
AUSTRIA
BELGIUM-LUX
CANADA
CZECH REP
DENMARK
FINLAND
FRANCE
GERMANY
GREECE
HUNGARY
ICELAND
IRELAND
ITALY
JAPAN
KOREA REP.
MEXICO
NETHERLANDS
NEW ZEALAND
NORWAY
POLAND
PORTUGAL
SPAIN
SWEDEN
SWITZ.LIECHT
TURKEY
UNTD KINGDOM
USA