The analysis of borders effects in intra-African trade

Emilie Kinfack Djoumessi and Alain Pholo Bala

ERSA working paper 701

August 2017

Economic Research Southern Africa (ERSA) is a research programme funded by the National Treasury of South Africa.

The analysis of borders effects in intra-African trade

Emilie Kinfack Djoumessi* Alain Pholo Bala *†

August 23, 2017

Abstract

The study aims to analyze the border effects on intra–African trade through the use of a gravity specification based on the monopolistic competition model of trade introduced by Krugman (1980). The study used CEPII data on trade flows between African countries over the period 1980-2006. We accommodate for the significant number of zero trade flows between several African countries by using the Heckman correction method. The findings suggest that while the extent of market fragmentation is on average very high within the African continent, the border effects within SADC and ECOWAS are more in line with other international estimations. Whereas results indicate that border effects faced by intra-African trade are quite substantial: on average an African country trade 108 times more "with itself" than with another country on the continent. Border effects in SADC and ECOWAS are respectively about 5 and 3 times lower. The inclusion of the infrastructure indices contributes significantly to this result. Considering infrastructure is actually an interesting way to capture the effect of distribution networks which represent, along with imperfect information and localized tastes, relevant but generally omitted sources of resistance.

Keyword: Intra-African Trade, Monopolistic Competition Model, Border Effect

JEL Classification: F12, F14, F15

1 Introduction

The analysis of economic growth in sub-Saharan Africa (SSA) has raised some major concerns in the Development Economics literature (Collier, 2006*a*,*b*; Dollar and Easterly, 1999; Easterly and Levine, 1997; Easterly, 2009*a*,*b*). While this continent has experienced strong economic growth rates lately (Pinkovskiy and Sala-i Martin, 2014), until the 1990s SSA was characterized by low and even sometimes negative

*Department of Economics & Econometrics, University of Johannesburg, APK Campus, PO Box 524, Auckland Park 2006, Johannesburg, South Africa.

†Public and Environmental Economics Research Centre (PEERC), *Corresponding author*. E-mail address: apholo@uj.ac.za.

economic growth (Cinyabuguma and Putterman, 2011; Easterly and Levine, 1997; Sachs and Warner, 1997; Artadi and Sala-i Martin, 2003).

Several impediments specific to the African growth process have been identified. Among them, we may consider that SSA is badly affected by low population density, long distances and high fragmentation (World Bank, 2008). High transport costs add to the high fragmentation of that region. For goods to pass the border in Africa, it takes more than 40 days which is twice the period in Latin America. The constraints are even worse for landlocked countries (Freire et al., 2015).

Collier (2006*a*) asserts that agglomeration economies in SSA are less important than those prevailing in Asia and in OECD countries. Because countries in that region are too small and not integrated enough, many African cities tend to be too small compared to the optimum. As shown by Au and Henderson (2006) in the case of Asia, this may have serious impacts in terms of foregone growth. Research on agglomeration economies and international competitiveness further suggests that latecomers to industrialization, such as Africa's natural resource exporters, face a competitive disadvantage linked to the spatial distribution of the global industry (Page, 2008).

Africa appears as fragmented and poorly integrated. Intra-regional trade in the region is fairly low comparatively to what is noticed in other areas. Therefore, improving regional integration appears as one of the key strategies that might foster economic growth in the African continent. Better regional integration might help SSA achieve greater economies of scale, widen markets, enhance industrial efficiency, and reduce the sub-region's external dependency and vulnerability of its economy (Jones, 2002). The other benefits that may be expected from an increased regional integration are the following: a greater bargaining power vis-à-vis the outside world; the minimization of duplication; thin spreading of resources and wasteful competition; a cheaper and more efficient transportation system; greater division of labor and specialization in production; established and strengthened product value chains and facilitated transfer of technology and knowledge via spillover effects; expansion of trade, incomes and employment due to free movement of goods, services, labor and capital, etc. (Jones, 2002; UNECA, 2010).

Another argument exists for promoting intra-African trade. To take advantage of trade-growth opportunities, SSA is prompted to diversify in several ways: one way would be by reducing its reliance on the stagnating markets of its traditional trading partners from the Organisation for Economic Cooperation and Development (OECD). Concurrently, it needs to become less dependent on the export of commodities vulnerable to price shocks (ITC, 2012). The redirection of SSA exports towards Asia, and China in particular, is already underway. However, since the large majority of SSA products destined for Asia are commodities, such a reorientation to growth markets does not eradicate the vulnerability to commodity price shocks. SSA countries increasingly trade within their own region too. This new trend of SSA trade is likely to achieve sustainability in export revenues by integrating deeper into sector value-chains and in this way increasing the share of value-added products within its exports. The redirection of SSA exports towards Africa may further reduce the risks caused by the volatility of commodity price shocks.

However, despite Africa's commitment to ease trade restrictions in order to create a common market within the framework of Regional Economic Communities (RECs) and the reorientation of its trade, barriers to intra-African trade tend to persist. On average over the past decade, only about 10%-12% of African trade has been with African nations, while 40% of North American trade has been with other North American countries, and 63% of trade in Western Europe is with other Western European nations (Economic Commission for Africa, 2011). Intra-African trade is also low when compared to that of other developing regions; intra Latin America trade amounts to 20 % while trade within Asian developing countries represents 48 % (Ancharaz et al., 2011). In this regard, SSA's trade seems to be extroverted as a substantial share of it is directed to the developed world: in 2011 more than half of SSA's merchandise trade was still done with the OECD partners (ITC, 2012).

The orientation of most of African trade towards developed partners raises the issue of the effectiveness of trade integration policies throughout African countries and RECs. In this paper, we assess the effectiveness of the strategies adopted by African countries and regional groupings to foster trade integration. We do so by evaluating the extent of border effects in intra-African trade. The intensity of border effects helps to characterize the extent of market fragmentation. The term "border effect" refers to the downward impact of national boundaries on the volume of trade, i.e. that two different countries trade far less with each other than do two locations in the same country, after controlling for factors such as income, alternative trading opportunities, distance, languages, and regional trade agreements (Evans, 2003). Markets are considered fragmented when national borders influence the pattern of commercial transactions (Head and Mayer, 2000).

Using border effects to measure the extent of market fragmentation is more robust than using market shares. As pointed out by de Sousa et al. (2012), while they are insightful, the market shares referred to above cannot be sufficient to draw conclusions on the level of market fragmentation faced by African exporters on the intra-regional market. The first problem is that any assessment of market access based on trade flows needs to specify a benchmark of trade patterns, to which actual international exchanges of goods will be compared. Without such a benchmark, which can only be provided by theory, we do not know what to compare those numbers to. This need of a reference for comparison justifies why we will derive an empirically estimable gravity-type equation from a theoretical framework.

A second limitation implied by the use of the import shares to assess market access is that for most products, the large majority of overall demand in a country is met by domestic producers, not foreign. A more relevant index of market access must take into account the market share of foreign producers in the overall demand. This is precisely what the border effects method does. The method considers trade flows within countries as well as among countries and compares imports from foreign countries to "imports from self" of domestic producers. Hence, this approach gives a benchmark based on a situation of the best possible market access, the one faced by domestic producers (de Sousa et al., 2012).

Border effects matter when firms have greater access to domestic consumers than to consumers in other nations. Following Head and Mayer (2000), we measure border effects as the average spread between actual trade and the trade that would be expected in an economy without border-related barriers. To capture border effects we rely on a gravity model. The gravity model offers a powerful empirical framework for analyzing bilateral trade flows (Anderson and van Wincoop, 2003; Behrens et al., 2012). However, the specification of gravity models has improved significantly over time. With an ad hoc specification linking bilateral trade flows to importer and exporter GDP and bilateral distance, McCallum (1995) found an impressive border effect between the Canadian provinces and the United States, i.e. the trade between two provinces is more than 20 times larger than trade between a province and an American state. These findings, which were confirmed by Helliwell (1996) and Hillberry (1998), suggest a very large degree of home bias in trade within North-America . One of the implication of the so-called "border puzzle" would be that international trade flows would remain much less dense than national trading ties even after all tariffs were removed.

However, some serious concerns have been raised on this first wave of studies. Anderson and van Wincoop (2003) pointed out that the lack of theoretical foundation of gravity equations has two adverse consequences. First, estimation results are likely to be biased because of omitted variables. Second, it is not possible to conduct a comparative statics exercise. Therefore, Anderson and van Wincoop (2003) used an estimated general equilibrium model to solve the "border puzzle". Anderson and van Wincoop (2003)'s model modifies McCallum (1995)'s specification by adding "multilateral resistance" variables. Applying their model to 1993 data, Anderson and van Wincoop (2003) found only moderate impact of trade barriers on trade flows: borders reduce trade between USA and Canada by 44%, and trade among other industrialized countries by 29%.

In this paper, we follow Head and Mayer (2000) and Combes et al. (2005) by performing a structural estimation based on the monopolistic competition model of trade proposed by Krugman (1980). That Krugman model derives a relation between the relative amounts consumers spend on foreign and domestic goods and their relative prices. The border effect is captured by the divergence between actual and predicted consumption ratios. The remainder of the paper is organized as follows. Section 2 presents the theoretical model and the implied econometric specifications. The data are described in Section 3. And finally, the econometric results are given in Section 4.

2 Model

Several theory-consistent estimation methods for the gravity model exist (Head and Mayer, 2013*a*). With the data at hand, bilateral trade flow between countries per industry, we can estimate border effects only through the complete odds specification as in Head and Mayer (2000) and Combes et al. (2005).[1] Therefore, we follow Head and Mayer (2000)'s and Combes et al. (2005)'s approach of including border effects into a monopolistic competition model. We use a fairly broad specification of preferences. The utility of the representative consumer in country i depends on the quantity of each variety h consumed from each country j. We assume that all varieties are differentiated from each other but products from the same country are weighted equally in the utility function. Representing the quantity consumed with c and the preference weight with a, the constant elasticity of substitution utility function is given by:

$$U_i = \left(\sum_{j=1}^{N} \sum_{h=1}^{n_j} (a_{ij} c_{ijh})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

while the implied budget constraint is

$$Y_i = \sum_{j=1}^{N} \sum_{h=1}^{n_j} p_{ijh} c_{ijh} \tag{2}$$

We may simplify the consumer optimization problem by assuming that all quantities of the varieties imported by country i from country j are symmetric. Then, denoting m_{ij} as the CIF value of exports from country j to country i ($m_{ij} = n_j c_{ij} p_{ij}$) and $m_i = \sum_{k=1}^{N} m_{ik}$ as expenses on goods from all origins (including the home country), one may obtain the following expression of bilateral imports after some algebra

$$m_{ij} = \frac{n_j p_{ij}^{1-\sigma} a_{ij}^{\sigma-1}}{\sum_{k=1}^{N} n_k p_{ik}^{1-\sigma} a_{ik}^{\sigma-1}} m_i. \tag{3}$$

2.1 Econometric specifications

A gravity equation may be derived from this expression and may lead to different econometric specifications. As the number of symmetric varieties is not observed, we may eliminate n_j and n_k in equation (3) by substituting them from equation (4)

$$\nu_j = q p_j n_j \tag{4}$$

where ν_j denotes the value of production in country j, q represents the quantity produced by each firm and p_j denotes the mill price of each variety. Considering the determination of delivered prices, p_{ij}, and of preferences, a_{ij}, we assume that the price paid by consumers in country i for products of country j is the product of the mill price, p_j, and of transaction costs T_{ij}

$$p_{ij} = p_j T_{ij} \tag{5}$$

We model transaction costs as a multiplicative function of the distance d_{ij}; of the bilateral industry level tariffs τ_{ij}; of the Non Tariffs Barriers (NTBs); of the level of infrastructure of the country i (j) ($\mathrm{IN}_{i(j)}$), and of an exponential function involving a contiguity dummy C_{ij}. Furthermore, we assume constant ad valorem NTBs of ξ for all cross-border trade. Defining B_{ij} as a dummy variable taking a value of one for $i \neq j$, we obtain

$$T_{ij} = d_{ij}^{\delta_1} t_{ij}^{\delta_2} \left[(1+\xi) \, \mathrm{IN}_i{}^{-\delta_3} \mathrm{IN}_j{}^{-\delta_4} \mathrm{e}^{-\delta_5 C_{ij}} \right]^{B_{ij}}, \tag{6}$$

with the tariff factor t_{ij} expressed as $t_{ij} = 1 + \tau_{ij}$.

Consumer preferences consist of a stochastic component, ϵ_{ij}, and a systematic preference for home-produced goods (or aversion to foreign-made goods) of α. Moreover, we hypothesize that several variables may mitigate or accentuate home bias and therefore posit the following equation for preferences:

$$a_{ij} = \exp \left[\epsilon_{ij} - \left(\alpha - \beta_1 L_{ij} - \beta_2 E_{ij} - \beta_3 \mathrm{COL}_{ij} + \beta_4 V_{ij} - R'_{ij} \phi \right) B_{ij} \right] \tag{7}$$

where L_{ij} takes the value of one for pairs of countries that share a common official language, and zero otherwise; E_{ij} takes the value of 1 when a language is spoken by at least 9% of the population in both countries i and j (ethnic language); $\mathrm{COL}_{ij} = 1$ if countries i and j had a common colonizer; and V_{ij} is the volatility of the bilateral exchange rate between countries i and j. R_{ij} depicts a vector composed of indicator variables taking the value of one when both countries i and j belong to the same regional groupings like SADC, ECOWAS, and COMESA or if countries i and j had a common colonizer after 1945; while ϕ is the corresponding vector of parameters. Substituting for a_{ij}, p_{ij} and n_j in (3) and taking logs we obtain the formulation of the gravity equation:

$$\begin{aligned} \ln m_{ij} &= \ln m_i + \ln \nu_j - (\sigma - 1)\, \delta_1 \ln d_{ij} - (\sigma - 1)\, \delta_2 \ln t_{ij} - \sigma \ln p_j - I_i \\ &- (\sigma - 1) \left[\alpha - \beta_1 L_{ij} - \beta_2 E_{ij} - \beta_3 \mathrm{COL}_{ij} + \beta_4 V_{ij} - R'_{ij}\phi + \ln(1+\xi) - \delta_3 \ln \mathrm{IN}_i - \delta_4 \ln \mathrm{IN}_j \right. \\ &- \left. \delta_5 C_{ij} \right] B_{ij} + (\sigma - 1)\, \epsilon_{ij} \end{aligned} \tag{8}$$

where I_i depicts the Head and Mayer (2000) importer's "inclusive value" defined as follows:

$$\begin{aligned} I_i &= \ln \left\{ \sum_{k=1}^{N} \exp \left[\ln \nu_k - \sigma \ln p_k + (\sigma - 1) \left(-\delta_1 \ln d_{ik} - \delta_2 \ln t_{ik} + \epsilon_{ik} \right. \right. \right. \\ &- \left. \left. \left. \left[\alpha - \beta_1 L_{ik} - \beta_2 E_{ik} - \beta_3 \mathrm{COL}_{ik} + \beta_4 V_{ik} - R'_{ik}\phi + \ln(1+\xi) + \delta_3 \ln \mathrm{IN}_i + \delta_4 \ln \mathrm{IN}_k - \delta_5 C_{ik} \right] B_{ik} \right) \right] \right\} \end{aligned}$$

The inclusive value captures the impact of the full range of potential suppliers to a given importer by taking into account their size, distance and relative border effects. There are several problems in estimating the influence of I_i (Head and Mayer, 2000). The most critical is that such an estimation would rely on parameters that are already in the equation to be estimated. There are two ways of sidestepping those problems. First, we may derive from equation (8) a fixed-effects specification fully consistent with theory. An alternative specification is to define a complete odds specification by setting $j = i$ in (8) to obtain an expression for $\ln(m_{ii})$, then to subtract $\ln(m_{ii})$ from $\ln(m_{ij})$.

As in Combes et al. (2005) and in the spirit of Hummels (1999) and Redding and Venables (2004), it is possible to derive from equation (8) a fixed-effects specification fully consistent with our theoretical model. This derivation implies replacing all destination-specific and origin-specific variables by two groups of destination and origin fixed effects. Only variables varying both with origin and destination are left in the regression.

However, this approach has several problems. First, it does not permit the estimation of all structural parameters. In particular, σ the elasticity of substitution between varieties cannot be recovered. Second, as the approach implies dropping internal trade flows, it approach excludes the possibility of estimating border effects. Therefore, to be able to estimate structural parameters and to evaluate border effects we rely on the complete odds specification.

2.2 The complete odds specification

Setting $j = i$ in (8) to obtain an expression for $\ln(m_{ii})$ and then, subtracting $\ln(m_{ii})$ from $\ln(m_{ij})$, we obtain the following complete odds specification (Combes et al., 2005):

$$\begin{aligned}\ln\left(\frac{m_{ij}}{m_{ii}}\right) &= -(\sigma-1)\left[\alpha+\ln(1+\xi)\right]+\ln\left(\frac{\nu_j}{\nu_i}\right)-\sigma\ln\left(\frac{p_j}{p_i}\right)-(\sigma-1)\,\delta_1\ln\left(\frac{d_{ij}}{d_{ii}}\right)\\ &- (\sigma-1)\,\delta_2\ln t_{ij}+(\sigma-1)\,\delta_3\ln \mathrm{IN}_i+(\sigma-1)\,\delta_4\ln \mathrm{IN}_j+(\sigma-1)\,\delta_5 C_{ij}\\ &+ (\sigma-1)\,\beta_1 L_{ij}+(\sigma-1)\,\beta_2 E_{ij}+(\sigma-1)\,\beta_3 \mathrm{COL}_{ij}-(\sigma-1)\,\beta_4 V_{ij}+(\sigma-1)\,R'_{ij}\phi+\varepsilon_{ij}\end{aligned} \tag{9}$$

where $\varepsilon_{ij} = (\sigma-1)(\epsilon_{ij}-\epsilon_{ii})$. This expression of the error term implies that errors are not independently distributed. We account for this correlation in the estimation through a robust procedure, allowing residuals of the same importing region to be correlated.

In other words, according to (9) the log of odds ratios is expected to increase with the log of relative production $\ln\left(\frac{\nu_j}{\nu_i}\right)$, with the fact that both importer and exporter share common languages and had a common colonizer after 1945 and to decrease with the logs of relative distance, $\ln\left(\frac{d_{ij}}{d_{ii}}\right)$, tariffs factor, $\ln(1+\tau_{ij})$, bilateral exchange rate volatility V_{ij} and relative price, $\ln\left(\frac{p_j}{p_i}\right)$.

The intercept in (9) may be described as the border effect. It is the average deviation between actual trade and the trade that would be expected in an economy without border-related barriers. It clearly captures the impact of NTBs (ξ) and home bias (α) as in Head and Mayer (2000).

Equation (9) implies a unit elasticity on relative production; thus it may be rewritten as

$$\begin{aligned}\ln\left(\frac{m_{ij}}{m_{ii}}\right) &- \ln\left(\frac{\nu_j}{\nu_i}\right) = -(\sigma-1)\left[\alpha+\ln(1+\xi)\right]-\sigma\ln\left(\frac{p_j}{p_i}\right)-(\sigma-1)\,\delta_1\ln\left(\frac{d_{ij}}{d_{ii}}\right)\\ &- (\sigma-1)\,\delta_2\ln t_{ij}+(\sigma-1)\,\delta_3\ln \mathrm{IN}_i+(\sigma-1)\,\delta_4\ln \mathrm{IN}_j+(\sigma-1)\,\delta_5 C_{ij}\\ &+ (\sigma-1)\,\beta_1 L_{ij}+(\sigma-1)\,\beta_2 E_{ij}+(\sigma-1)\,\beta_3 \mathrm{COL}_{ij}-(\sigma-1)\,\beta_4 V_{ij}+(\sigma-1)\,R'_{ij}\phi+\varepsilon_{ij}\end{aligned} \tag{10}$$

Imposing a unit elasticity on relative production helps addressing two different econometric issues:

first, output and trade are jointly determined in equilibrium (Harrigan, 1996; Head and Mayer, 2000); which entails an endogeneity problem. Moving relative production in the left-hand side as in (10) is a way to handle this simultaneity issue without resorting to instrumental variables. This strategy may address another problem: measurement error for production. As production information may be inexact especially for developing countries of SSA, we may obtain biased coefficients estimates.

2.3 Heckman correction for sample selection bias

The monopolistic competition trade model assumes positive trade between each country pair in each industry (Head and Mayer, 2000). More generally, most of the structural gravity models do not naturally generate zero flows (Head and Mayer, 2013*a*). However, most trade data sets exhibit substantial fractions of zeros (Head and Mayer, 2013*a*). Helpman et al. (2008) show that, at the country level, country pairs that do not trade with each other or trade in only one direction account for about half the observations. There are two implications of a prevalence of zeros (Head and Mayer, 2013*a*). First, trade models need to be adjusted in order to accommodate zeros. Second, estimation methods should be revised to allow for consistent estimates in the presence of a dependent variable having a high incidence of zeros.

Several approaches can be used to accommodate structural gravity models to a prevalence of zeros (Head and Mayer, 2013*a*). One simple approach is to consider that zeros arise from a data recording issue, that is "there are no "structural zeros" but only "statistical zeros"" (Head and Mayer, 2013*a*, p. 50-51). Some structural models generate zeros endogenously by adding a fixed cost of exporting from i to j (Helpman et al., 2008; Eaton et al., 2012).[2] These models have the common characteristic that zeros are more likely between small and/or distant countries (Baldwin and Harrigan, 2011; Head and Mayer, 2013*a*). As log of zero is not defined, observations without positive bilateral trade are lost, which results in systematic selection bias as shown in Head and Mayer (2013*a*).

To find out which estimators may be consistent when zeros are an endogenous component of the data-generating process, Head and Mayer (2013*a*) compare several estimators with a Monte Carlo Simulation: the Least Squares Dummy Variables (LSDV), the Eaton and Tamura (1994) (ET) Tobit model, the Eaton and Kortum (2001) (EK) Tobit model, the Multinomial Pseudo-Maximum-Likelihood (Multinomial PML), the Poisson PML and the Gamma PML. The main findings of this Monte Carlo simulation exercise is that the choice of the most suitable estimator depends on the process generating the error term. Under the Constant Variance to Mean Ratio (CVMR) assumption, Poisson or Multinomial PML provide the best estimators. However, under the log-normality assumption the EK Tobit method should be preferred.

The Heckman selection model was left out of the Monte Carlo simulation exercise. The reason advanced by Head and Mayer (2013*a*) is that the method used by Helpman et al. (2008), the leading paper using a Heckman based approach of zeros, is designed to uncover a different set of parameters than the other approaches which estimate coefficients that combine extensive and intensive margins. The critical challenge implied by the Heckman-based methods is the difficulty of finding an exclusion restriction. One would preferably use a variable in the probit selection equation which according to the theory can be excluded from the gravity regression equation. Since both equations have exporter and importer invariant variables, this variable should be dyadic in nature (Head and Mayer, 2013*a*).

Given, the poor results obtained with the EK Tobit and the Poisson PML, in this paper we follow Head and Mayer (2000) who used a two-step Heckman selection model for a complete odds specification.[3] For the exclusion restriction we follow Helpman et al. (2008) by using two indicators for high entry barriers in both countries i and j. These entry costs are captured through their effects on the number of days, the number of legal procedures, and the relative cost (as a percentage of GDP per capita) for an entrepreneur to legally start operating a business.

More precisely this indicator takes the value of one for country pairs in which both the importing and exporting countries have entry regulation measures above the cross-country median. The first indicator uses the sum of the number of days and procedures above the median (for both countries) whereas the other indicator uses the sum of the relative costs above the median (again for both countries). By construction, these variables capture regulation costs that should not depend on the exports volume of a particular country and therefore, should comply with the necessary exclusion restrictions.

Our Heckman selection model implies either (9) or (10) as the regression equation. Moreover, we assume that the dependent variable is observed if the following selection equation is verified:

$$X'_{ij}\varphi + \mu_1 \mathrm{IN}_i + \mu_2 \mathrm{IN}_j + \eta_{ij} > 0 \tag{11}$$

where

$$\eta_{ij} \sim \mathcal{N}(0,1)$$
$$\mathrm{corr}(\varepsilon_{ij}, \eta_{ij}) = \rho$$

and X_{ij} represents a vector including all the bilateral regressors of either (9) or (10) in addition to the high entry barriers indicators. ρ is the correlation between unobserved determinants of propensity to export η_{ij} and unobserved determinants of the log of odds ratios ε_{ij}. Therefore, testing whether $\rho = 0$ is equivalent to testing if there is sample selection. Defining the indicator variable IND_{ij} to be equal to 1 when country j exports to i and to 0 when it does not, the probability that j exports to i can be formally expressed by the following probit equation:

$$\Pr\left(\mathrm{IND}_{ij} = 1 | \text{observed variables}\right) = \Phi\left(X'_{ij}\varphi + \mu_1 \mathrm{IN}_i + \mu_2 \mathrm{IN}_j\right) \tag{12}$$

where $\Phi(.)$ is the cumulative distribution function of the unit-normal distribution. The choice of the regressors in the probit equation is consistent with the fact that variables that are commonly used in gravity equations also affect the probability that two countries trade with each other (Helpman et al., 2008).

3 Data requirements

In this paper we use trade and production data from the CEPII's TradeProd database.[4] This database proposes bilateral trade, production and protection figures in a compatible industry classification for developed and developing countries. It covers 28 industrial sectors in the ISIC Revision 2 (International Standard Industrial Classification) from 1980 to 2006. We restrict our analysis to the trade flows between African countries.

The relative prices are captured by the price level of consumption from the Penn World Tables v.7.1. Bilateral information on the prevalence of common languages,[5] contiguity and distances are obtained from CEPII's GeoDist database. A valuable contribution of the GeoDist database is to compute internal and international bilateral distances in a totally consistent way. It is critical to define intranational distances in a manner that is compatible with international distances computations as any overestimate of the internal/external distance ratio will imply a mechanic upward bias in the border effect estimate (Mayer and Zignago, 2011). Therefore, de Sousa et al. (2012) have computed the weighted distances (`distw` and `distwces`) using city-level data to assess the geographic distribution of population inside each nation.[6]

We estimate the volatility of the bilateral exchange rate by the standard deviation of a monthly series of bilateral exchange rate. The bilateral exchange rate is expressed as the number of currency units of country i per currency unit of country j. These monthly series are from the International Financial Statistics of the IMF.

The infrastructure index $\mathrm{IN}_{\mathrm{i(j)}}$ of the country i (j) is built using three variables from the database merged from the infrastructure data set constructed by Canning (1998) and the infrastructure data from the World Bank's World Development Indicators 2006 (World Bank, 2006): the density of roads, of railways, and the number of telephone lines per capita of country i (j), each variable being normalized to have a mean equal to one. An arithmetic average is then calculated over the three variables, for each country and each year (the computation is similar to Limão and Venables (2001) and Carrère (2004, 2006)).[7]

The infrastructure data is reported from 1950 to 2005 but, for most of the countries, data is missing for several years. Data is also missing for several countries.For some countries the missing data can be explained by the merging procedure used by World Bank (2006). Generally, the missing World Bank data is filled in using the adjusted Canning data. However, when the two series disagree substantially World Bank (2006) report only the data set they think is more consistent, or in some cases neither data set, in the merged data set. This is why we have data missing for some countries like Canada, Chile, Denmark, Mexico, Russia etc. North Africa (especially Tunisia and Egypt) and Southern Africa (especially Mauritius and South Africa) have the highest infrastructure indices of the continent. In Western Africa Senegal seems to be leading, while Gabon and Rwanda appears to have the highest infrastructure indices respectively in Central and Eastern Africa.

Finally, data on the indicators for high entry barriers in both countries that serve as excluded variables for the Heckman selection method are obtained from the World Bank's Doing Business dataset (`http://www.doingbusiness.org/data`). Tables 6, 7 and 8 display the descriptive statistics respectively for the whole sample, for the intra-African trade and for SADC. The average bilateral flow for the whole sample (23,802) five times higher than the corresponding figure for intra-African trade (4,561) but lower than the average bilateral trade flow within SADC (27,925). The levels of the infrastructure indices are however in average lower in SADC and in Africa than in the whole sample. We have to note that the data on infrastructure is reported only for a fraction of the dataset (a little bit less than 3,000,000 observations versus 15,459,569 observations for the bilateral trade flows).

4 Analysis of border effects

This section is divided as follows. We first present a general overview of intra-African border effects, by estimating different specifications either with or without the infrastructure index. Then, we assess how these border effects are impacted on by tariffs. Finally, we contrast these results with border effects arising from trade between European exporters and African exporters. This comparison would help us to figure out whether Africa is more open to overseas markets than to foreign markets of their own region.

4.1 Intra-African border effects

4.1.1 Preliminary results

The complete odds specification permits an estimation of structural parameters and an evaluation of the border effects. In Table 1 we compare specifications (9) and (10). We pool the years 1980–2006 so that our estimations impose a common set of coefficients on all the industries in the sample.[8] The first two columns of Table 1 provide results for the regression and the selection equations for the specification (9) which does not impose any restriction on the elasticity of the exporter to the importer production ratio.

These columns provide expected results and significant estimates for the elasticities of the relative production, and for the coefficients of the contiguity, the common languages, the common colonizer dummies and the volatility of the bilateral exchange rate. The results regarding the elasticity of the relative price are more puzzling. While the elasticity is significant at 1%, the results indicate a positive value.

The estimate of the elasticity of relative production appears to be quite lower than one. Theoretically, on the one hand such an estimate could arise because varieties from countries with larger production are produced at a larger scale. This would imply that rises in relative production overstate rises in the number of varieties offered. On the other hand, those results could rather be caused by the endogeneity biases mentioned earlier on Head and Mayer (2000). In this regard specification (10), illustrated in the two last columns of Table 1, presents clear advantages.

Results from these last two columns are broadly similar to the previous results except for the elasticities of the relative distance and of the relative price. The elasticity of the relative distance is higher than in the previous results. In their meta-analysis Disdier and Head (2008) find an average elasticity of 0.9. Therefore, while the new result seems to have improved, it is still lower than what is suggested in the literature. An improvement is also noticed for the relative price elasticity which now becomes negative as one would have expected. However, the results indicate an unreasonably small value.[9] Actually, this result of low price elasticities when using direct proxies for prices is quite frequent in the literature (Head and Mayer, 2000; Erkel-Rousse and Mirza, 2002; de Sousa et al., 2012).

The findings suggest therefore that endogeneity biases are quite critical for these two variables. Since specification (10) is the most appropriate to mitigate endogeneity biases, our further empirical analysis of border effects will be based on it. The results from these last two columns of Table 1 indicate that African countries sharing a common border trade 3 (exp(1.0969)) times more than non-contiguous African countries; those having a common official language trade 1.87 (exp(0.6277)) times more; those sharing a common ethnic language trade 1.45 (exp(0.3750)); times more, and those who had a common colonizer trade 1.21 (exp(0.1883)) times more.

We now focus on the analysis of the border effects. There are several ways to express their magnitude. We opt to follow the dominant trend in the literature by expressing borders effects as the ratio of imports from self to imports from others, holding other things equal (McCallum, 1995; Wei, 1996; Head and Mayer, 2000; Anderson and van Wincoop, 2003; de Sousa et al., 2012). Results suggest a border effect equal to 2,350 (exp(7.7622)). So internal trade flows in African countries are about 2,350 larger than their imports from other African countries. This would support the recurrent claim about the poor integration within the African continent (Collier, 2006*a*; World Bank, 2008). However, trade impediments are not homogeneously distributed throughout Africa. Border effects within SADC and ECOWAS are substantially lower than the sample average: $\exp(7.7622 - 1.8538) \approx 368$ and $\exp(7.7622 - 0.4707) \approx 1,468$; while those within COMESA are quite higher $\exp(7.7622 + 0.1186) \approx 2,646$.[10]

Except for the border effect within SADC, the average intra-African border effect and the border effect within the other RECs are all higher than 6.22, the border effects between Southern exporters and Northern importers as computed by de Sousa et al. (2012). This would indicate that, except within SADC, African countries have a poorer access to other African countries than the average access of Southern exporters to Northern markets. Therefore, this first set of results points out to the poor access of African countries to other markets within Africa. Except within SADC, the integration within the African region seems to be lower than the integration between Southern and Northern markets (de Sousa et al., 2012).

It is usual in the literature to compute the tariff equivalent of the border effect (Head and Mayer, 2000; de Sousa et al., 2012). But, such a computation requires an estimate of the elasticity $(\sigma - 1)$. The coefficient of the relative price would be the designated source for this parameter. The problem is that, as previously explained, this estimate is quite disappointing, with a value much lower than expected. Yet the literature provides estimates of the trade elasticity. While Head and Ries (2001), Eaton and Kortum (2002), and Lai and Trefler (2002) suggest an elasticity around 8 for developed countries in recent years, the consensus in the literature seems to have shifted towards half of that value (Head and Mayer, 2013*b*). Using disaggregated price and trade-flow data for 123 countries in the year 2004 Simonovska and Waugh (2014) also found estimates roughly equal to 4 which implies doubling the welfare gains from international trade.

With the last estimate of $(\sigma - 1)$ we obtain a tariff equivalent of:

- The average intra-African border effect equal exp (7.7622/4)-1=596%;
- The SADC border effect equal to exp (5.9084/4)-1=338%;
- The ECOWAS border effect equal to exp (7.2915/4)-1=519% and;
- The COMESA border effect equal to exp (7.8808/4)-1=617%.[11]

These African tariff equivalent border effects are very high. By contrast, those estimates are much higher than the corresponding EU tariff equivalent of 99% computed from 1984 to 1986 by Head and Mayer (2013*b*). We can further elaborate on the interpretation of those border effects by highlighting the expression of the intercept in equations (9) and (10):

$$\text{border effect} = -(\sigma - 1)\left[\alpha + \ln(1 + \xi)\right] \tag{13}$$

Therefore, the border effect can be decomposed as the product of $-(\sigma - 1)$ the trade elasticity and the sum of two terms: α a term capturing the systematic preference for home-produced goods, and the logarithm of 1 plus the constant ad valorem NTBs. The problem with this formulation of the border effects implied by (13) is that it includes parameters that are not measurable; the home-bias preference parameter α for instance. Moreover, only a portion of the so-called NTBs can be documented.[12]

Following Head and Mayer (2013*b*) we may reformulate the border effect as follows

$$\text{border effect} = -(\sigma - 1)\ln(1 + \psi_c + \psi_d) \tag{14}$$

where ψ_c would represent the ad valorem measurable NTB, and ψ_d would represent any 'dark' cost implied by crossing borders.[13] From this formulation we can conclude that $\psi_c + \psi_d$ would be equal to 596% for intra-African trade, to 338% for trade within SADC, to 519% for trade within ECOWAS and to 617% for trade within COMESA. These figures are considerably higher than what is reported in the literature. From the border effect estimated by Anderson and van Wincoop (2003), Head and Mayer (2013*b*) report an ad-valorem dark cost of 49% five years into the implementation of the Canada-USA Foreign Trade Agreement (FTA). From Head and Mayer (2000)'s estimate of the border coefficient between EU members from 1984 to 1986, they derive an ad-valorem 'dark' cost equal to 99%. The African estimates are way beyond those measures.

Table 1: Gravity and selection equations for intra-African trading relationships. Estimation without tariffs and infrastructure index, 1980-2006 averages with common coefficients for all industries.

	1		**2**	
Variables	$\ln\left(\frac{m_{ij}}{m_{ii}}\right)$	(Probit) IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	(Probit) IND_{ij}
Border/Intercept	-8.4591***	2.0719***	-7.7622***	2.0118***
	(0.0813)	(0.0722)	(0.0836)	(0.0718)
Ln rel. prod	0.6906***	0.1048***		
	(0.0084)	(0.0028)		
Ln rel. price	0.1191***	-0.0702***	-0.1699***	0.0793***
	(0.0268)	(0.1263)	(0.0271)	(0.0120)
Ln rel. dist	-0.1360***	-0.3184***	-0.3313***	-0.2873***
	(0.0239)	(0.0089)	(0.0226)	(0.0086)
Reg. days & proc.		-0.7753***		-0.8706***
		(0.0490)		(0.0485)
Reg. costs		-0.9022***		-0.9404***
		(0.0369)		(0.0377)
Exch. rate volat.	-4.08×10^{-6}***	-0.0002***	-3.28×10^{-6}***	0.0002***
	(1.06×10^{-6})	(0.0000)	(1.05×10^{-6})	(0.0000)
Contiguity	1.3277***	0.3674***	1.0969***	0.3775***
	(0.0628)	(0.0339)	(0.0661)	(0.0330)
Official language	0.5955***	0.1858***	0.6277***	0.1600***
	(0.0618)	(0.0215)	(0.0653)	(0.0212)
Ethnic language	0.4109***	0.6086***	0.3750***	0.6219***
	(0.0565)	(0.0187)	(0.0600)	(0.0185)
Common colonizer	0.2879***	0.1557***	0.1883***	0.1690***
	(0.0466)	(0.0217)	(0.0485)	(0.0185)
SADC	1.5927***	0.6010***	1.8538***	0.5109***
	(0.0675)	(0.0312)	(0.0688)	(0..0298)
ECOWAS	0.9049***	0.9143***	0.4707***	1.0380***
	(0.1129)	(0.0732)	(0.1263)	(0..0737)
UEMOA	-0.3989	-1.9680***	-0.2968	-2.0101***
	(0.4966)	(0.1604)	(0.5802)	(0.1568)
COMESA	0.3364***	-0.3347***	-0.1186***	-0.1314***
	(0.0472)	(0.0175)	(0.0457)	(0.0163)
CEMAC	0.3281	0.0605	-0.1542	0.2938
	(0.2632)	(0.2024)	(0.2530)	(0.2150)
ρ	-0.0288		-0.0775***	
	(0.0304)		(0.0261)	
N	31,892			
σ	2.4376		2.5840	
Number of censored obs	15,627			
Number of uncensored obs	16,265			

*** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.10 level.
(1): Heckman ML Estimation with an unrestricted elasticity on relative production
(2): Heckman ML estimation with unit elasticity on relative production

4.1.2 Accounting for infrastructure

We may need some less conventional sources of resistance to improve the estimates of the border effects arising in intra-African trade.[14] The usual sources of resistance – cross border tariffs or border compliance costs – are not sufficient to explain the high level of border effects within Africa. Quoting Grossman (1998), Head and Mayer (2013*b*) mention three possible explanations: informational impediments to trade, localized and historically determined tastes, and business networks. It is quite difficult to get proxies for these dimensions, especially for intra-African trade. However, by using the infrastructure indices of the importer and the exporter together, we might get an imperfect but useful proxy of the business network which might be insightful for explaining the relatively low border resistance prevailing in SADC.

While one might conjecture that SADC has established more effective institutions to promote regional integration, another plausible explanation may emerge: the high quality of the transport infrastructure of several SADC countries (South Africa, Botswana, Namibia) may favor trade flows within the region. It would be useful to use data on quality of transport-related infrastructure to disentangle the impact of the infrastructure from the border effects.

The last two columns of Table 2 provide results with the infrastructure index. The first two columns of the same table serve as a benchmark as they provide estimations with the same sample as with the last two but without the infrastructure index. With the inclusion of the infrastructure index the sample size drops significantly from 31,892 to 10,136. This drop of the sample size has an impact on the elasticity of relative price and on the coefficient of the ethnic language dummy which now become insignificant. Yet, we obtain a distance elasticity of about -0.8097, which suggests that the inclusion of the infrastructure indices allow the distance elasticity to increase towards the estimate reported by Disdier and Head (2008). Furthermore, with the infrastructure indices we find that contiguous countries trade 1.73 (exp(0.5496)) more, countries sharing a common official language trade 2.04 (exp(0.7110)) more, and countries who had a common colonizer trade 1.73 (exp(0.5504)) more.

The impact of the infrastructure indices on the intra-African border effects is even more remarkable. While the average intra-African border effects arising from the first two columns of Table 2 is about 2,750 (exp(7.9192)), it shrinks to 108 (exp(4.6806)) when the infrastructure indices of the importer and the exporter are taken into account. Considering the infrastructure index implies a sharp decrease in the average intra-African border effects. Actually this figure implies that internal trade flows in African countries are about 108 larger than their imports from other African countries, after controlling for distances, languages, contiguity effects, common colonizer effects, and infrastructure.

It is even more interesting to assess the border effects within the different regional groupings. With respectively 22 (exp(4.6806-1.6053)), 33 (exp(4.6806-1.1861)) and 87 (exp(4.6806-0.2111)), the border effects in SADC and ECOWAS are as before significantly lower than the continental average; while those of COMESA are only slightly smaller. Therefore, after accounting for bilateral tariffs by deducting the simple average world tariff (12.5%) the tariff equivalent of the sum of 'grey' and 'dark' costs becomes:

- exp (4.6806/4)-1-0.125=209.5% for intra-African trade;
- exp (3.0753/4)-1-0.125=103.5% for trade within SADC;
- exp (3.4945/4)-1-0.125=127.5% for trade within ECOWAS; and
- exp (4.4695/4)-1-0.125=193.5% for trade within COMESA.

Except for intra-African trade and COMESA, these 'grey' and 'dark' costs are low comparatively to the border effects tariff equivalent of 99% for trade within the EU (Head and Mayer, 2013*b*). With the inclusion of the infrastructure indices, the measurement of border effects and 'dark' costs for SADC and ECOWAS are more in line with other international estimations. SADC is the only REC that displays a rising trend for the intraregional trade share in GDP (de Melo and Tsikata, 2014). The SADC intraregional trade share in GDP is actually on average one of the highest in the region. The proportion of intraregional rose from 1.4% in 1970 to 12.2% in 2000. Furthermore, while the intra-SADC trade volume increased by only 26% between 1970 and 1980, it markedly increased by 206% between 1980 and 1985 and 75% between 1990 and 2000 (Babarinde, 2003)[1].

ECOWAS is another REC that experiences some success in its integration endeavours. There is evidence of trade creation since the inception of the ECOWAS. Intra-ECOWAS trade volume significantly increased by 705% between 1970 and 1980 (following the launch of ECOWAS), 122% between 1980 and 1990, and 117% between 1990 and 1998 (Babarinde, 2003).[2] Moreover, ECOWAS includes the West African Economic and Monetary Union (WAEMU) members who share a common currency, and have achieved deeper integration (de Melo and Tsikata, 2014).

Table 9 in Appendix B reports an estimation including the bilateral tariffs but not the infrastructure indices. Compared to Table 1, this estimation implies a sharp decrease of the number of observations: from 31,892 to 5,916. The reason is because data on the bilateral tariffs are available only between 1989 and 2001. This shortfall in the number of observations may be an explanation why, surprisingly, the inclusion of bilateral tariffs does not seem to improve the results. First, the elasticity of the tariff factor is only significant at the 10% significance level. But, more important the border effects in the presence of tariffs, 2,914 (exp(7.9772) cfr. the two last columns of Table 9) are even higher than in their absence (cfr. Table 1).

[1] The inauguration of the first post-apartheid government in South Africa (SA) and the accession of SA to the SADC seemed to have contributed substantially to the jump in intraregional trade volume: 303% between 1990 and 1995.

[2] In 1999 the ECOWAS initiated a traveler's check program to facilitate trade and the movement of people within the region. Additional achievements of the ECOWAS comprise a trans-African highway, a trans-African pipeline that supplies Nigeria's natural gas to some member countries, and the repeal of visa requirements for ECOWAS citizens (Babarinde, 2003).

Table 2: Gravity and selection equations for intra-African trading relationships. Estimation with infrastructure indices but without tariffs, 1980-2006 averages with common coefficients for all industries and with unit production elasticity.

	1		**2**	
Variables	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	(Probit) IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	(Probit) IND_{ij}
Border/Intercept	-7.9192***	1.9563***	-4.6806***	2.7636***
	(0.1343)	(0.1321)	(0.2192)	(0.1778)
Ln rel. price	0.0362	0.0024	0.0019	-0.0059
	(0.0321)	(0.0177)	(0.0310)	(0.0181)
Ln rel. dist.	-0.2954***	-0.2059***	-0.8097***	-0.2888***
	(0.0322)	(0.0156)	(0.0377)	(0.0212)
Reg. days & proc.		-0.5173***		-0.4842***
		(0.07543)		(0.0776)
Reg. costs		-0.9383***		-0.6424***
		(0.0813)		(0.0900)
Exch. rate volat.	-4.08×10^{-6}***	6.41×10^{-5}***	-3.51×10^{-6}***	8.45×10^{-5}***
	(1.24×10^{-6})	(1.71×10^{-5})	(1.19×10^{-6})	(0.000022)
Contiguity	0.9570***	0.2900***	0.5496***	0.1655*
	(0.0962)	(0.0968)	(0.0965)	(0.0949)
Official language	0.9349***	0.4310***	0.7110***	0.3639***
	(0.1076)	(0.0516)	(0.1038)	(0.0545)
Ethnic language	0.1031	0.3990***	0.1438	0.4072***
	(0.1024)	(0.0403)	(0.0994)	(0.0407)
Common colonizer	0.1524**	0.1355***	0.5504***	0.2135***
	(0.0650)	(0.0472)	(0.0670)	(0.0500)
Ln infra_index_imp			5.3044***	0.7778***
			(0.2565)	(0.1512)
Ln infra_index_exp			1.0412***	1.4672***
			(0.2237)	(0.1507)
SADC	1.8323***	0.9671***	1.6053***	0.8661***
	(0.1338)	(0.2225)	(0.1255)	(0.2187)
ECOWAS	0.7495***	0.7680***	1.1861***	0.8374***
	(0.2295)	(0.1941)	(0.2192)	(0.2020)
COMESA	-0.0589	0.1794***	0.2111***	0.1746***
	(0.0595)	(0.0321)	(0.0587)	(0.0328)
CEMAC	0.1627	-0.1395	0.4885*	-0.0313
	(0.2811)	(0.2237)	(0.2795)	(0.2259)
ρ	-0.2741***		0.1208***	
	(0.0614)		(0.0345)	
N	10,136			
σ	2.4626		2.3724	
Number of censored obs	2,288			
Number of uncensored obs	7,848			

*** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.10 level.
Heckman ML Estimations with an unrestricted elasticity on relative production

4.1.3 The impact of bilateral tariffs

The joint inclusion of bilateral tariffs and infrastructure indices (cfr. Table 10 in Appendix B) does not bring much of an improvement. This joint inclusion reduces the sample size to 1,616 observations. While the implied border effects 1,582 (exp(7.3662)), are lower than in Table 9, they are still much higher than in the results derived from the specification with infrastructure indices but without tariffs (displayed in Table 2).

The evolution of the border effects through time is analyzed next. Whether we include infrastructure indices or not, two opposite outcomes emerge from this analysis. In Table 3 we can see the evolution of border effects without accounting for infrastructure indices. The first five columns of Table 3 shows the evolution when tariffs are not accounted for. We can notice there that border effects decreased from 5,041 (exp(8.5254)) between 1980-1988 to 2,243 (exp(7.7156)) between 1989-1997, and finally to 1,458 (exp(7.2846)). For SADC the border effects also decreased 479 (exp(7.7156-1.5445)) between 1989-1997 to 203 (exp(7.2846-1.9727)) between 1998-2006. Regarding ECOWAS, border effects decrease from 2,598 (exp(8.5254-0.6630)) between 1980-1988 to 389 (exp(7.7156-1.7512)) between 1989-1997 but then increased slightly to 454 (exp(7.2846-1.1671)) between 1998-2006.

When we take tariffs into consideration, as in the last four columns of Table 3, border effects also diminished from 5,998 (exp(8.6992)) between 1989-1997 to 1,395 (exp(7.2403)) between 1998-2006. For SADC they also decreased from 469 (exp(8.6992-2.5477)) between 1989-1997 to 166 (exp(7.2403-2.1271)). Let us now focus on the results with infrastructure indices. In doing so we disregard the impact of bilateral tariffs. In this configuration, border effects evolve in the opposite direction. Table 4 shows the evolution of border effects when the infrastructure indices are included as regressors. Border effects rise from 23 (exp(3.1396)) between 1980-1988 to 187 (exp(5.2314)) and to 1,914 (exp(7.5567)). For SADC border effects increase from 37 (exp(5.2314-1.6227)) between 1989-1997 to 190 (exp(7.5567-2.3113)) between 1998-2006.

The fact that tables 3 and 4 display evolutions in opposite direction might give ground to the hypothesis that the decline of border effects is mostly driven by the improvement of transport and communication infrastructure than by the reduction of NTBs like quotas, export restraints and s forth. However, more information especially on NTBs is needed to confirm this suggestion.

Table 3: Evolution of intra-African border effects over time. Estimations without infrastructure indexes.

	1980-1988	1989-1997		1998-2006		1989-1997		1998-2006	
	OLS	Heckman		Heckman		Heckman		Heckman	
Variables	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	IND_{ij}
Border/Intercept	-8.5254***	-7.7156***	1.7130***	-7.2846***	1.9672***	-8.6992***	2.8092***	-7.2403***	2.1391***
	(0.1349)	(0.1347)	(0.1071)	(0.1533)	(0.1042)	(0.4969)	(0.2608)	(0.3302)	(0.2287)
Ln rel. price	-0.0461	-0.0815**	-0.0554***	-1.0232***	0.6053**	-0.4984***	0.2128***	-0.8970***	0.3815***
	(0.0420)	(0.0384)	(0.0161)	(0.1374)	(0.0342)	(0.1361)	(0.0543)	(0.3091)	(0.0845)
Ln tariff factor						-0.1733	-0.2148	-1.8252*	-0.5500***
						(0.5262)	(0.1587)	(1.0206)	(0.2018)
Ln rel. dist.	-0.2807***	-0.3020***	-0.2867***	-0.4592***	-0.2578***	0.0238	-0.5958***	-0.4211***	-0.2446***
	(0.0349)	(0.0386)	(0.0134)	(0.0420)	(0.0153)	(0.1861)	(0.0582)	(0.0819)	(0.0282)
Reg. days & proc.			-1.0938***		-0.7630***		-1.3334***		-0.9127***
			(0.0625)		(0.0821)		(0.1248)		(0.1824)
Reg. costs			-0.6330***		-1.1078***		-0.6530***		-1.0886
			(0.0658)		(0.0493)		(0.1287)		(0.0921)
Exch. rate volat.	-2.38×10^{-6}**	-7.45×10^{-5}*	22.21×10^{-5}***	-0.0441***	0.0114***	25.13×10^{-5}	-31.2×10^{-5}***	-0.0766***	0.0309***
	(1.07×10^{-6})	(4.03×10^{-5})	(2.55×10^{-5})	(0.0073)	(0.0027)	(33.65×10^{-5})	(8.48×10^{-5})	(0.0153)	(0.0074)
Contiguity	0.9032***	1.3205***	0.6383***	1.3752***	0.1917***	1.9645***	0.0386	1.1004***	0.4117***
	(0.1156)	(0.0988)	(0.0556)	(0.1617)	(0.0526)	(0.3050)	(0.1474)	(0.3151)	(0.1125)
Official language	0.6003***	0.6212***	0.0252	0.4110***	0.3132***	0.1988	-0.2061***	0.6475***	0.3136***
	(0.1685)	(0.0969)	(0.0336)	(0.1042)	(0.0297)	(0.2086)	(0.0704)	(0.2269)	(0.0635)
Ethnic language	0.7666***	0.1232	0.7326***	0.7322***	0.3125***	0.4707**	0.7294***	0.7383***	0.3821***
	(0.1402)	(0.0949)	(0.0278)	(0.0994)	(0.0285)	(0.2146)	(0.0640)	(0.2176)	(0.0618)
Common colonizer	0.0105	0.2677***	0.2889***	0.6177***	-0.1246***	0.8680***	0.3902***	0.5838***	-0.0372
	(0.0925)	(0.0749)	(0.0353)	(0.0943)	(0.0314)	(0.1988)	(0.0827)	(0.1822)	(0.0613)
SADC		1.5445***	0.4002***	1.9727***	0.9220***	2.5477***	0.5416***	2.1271***	0.5613***
		(0.0890)	(0.0419)	(0.1313)	(0.0475)	(0.3011)	(0.1276)	(0.2448)	(0.0761)
ECOWAS	0.6630***	1.7512***	0.2818**	1.1671***	0.9696***	1.8858	-0.7046*	1.5618	1.0870***
	(0.1395)	(0.3946)	(0.1299)	(0.4455)	(0.1702)	(1.6900)	(0.4022)	(1.1523)	(0.3836)
UMEOA		-2.3719***	-1.2296***	-1.3904	-1.4117***			-3.6930**	-1.5325***
		(0.5417)	(0.2727)	(0.8952)	(0.2401)			(1.5588)	(0.5606)
COMESA	0.1531*	-0.0793	0.0003	-0.1962**	-0.2766***	-0.5044**	-0.0070	-0.9261***	-0.2776***
	(0.0848)	(0.0700)	(0.0254)	(0.0983)	(0.0289)	(0.2207)	(0.0738)	(0.2021)	(0.0584)
CEMAC	0.5988*	-0.2757	-0.2318	0.9629		-0.4527	-0.6176*		
	(0.3499)	(0.3733)	(0.2555)	(0.6444)		(0.6112)	(0.3212)		
ρ		-0.1484***		-0.0120		-0.0764		0.1679***	
		(0.0399)		(0.0401)		(0.0766)		(0.0810)	
N	4,219	13,867		13,822		2,698		3,218	
σ	2.4344	2.4883		2.7058		2.5398		2.6479	
Number of censored obs		7,300		8,327		1,442		1,871	
Number of uncensored obs		6,567		5,495		1,256		1,347	

*** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.10 level.
Heckman ML Estimations with a unit elasticity on relative production

Table 4: Evolution of intra-African border effects over time. Estimations with infrastructure indexes but without tariffs.

	1980-1988	1989-1997		1998-2006	
	OLS	Heckman		Heckman	
Variables	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	IND_{ij}
Border/Intercept	-3.1396***	-5.2314***	4.0812***	-7.5567***	4.2167***
	(0.3610)	(0.4032)	(0.2449)	(0.9362)	(0.5222)
Ln rel. price	-0.0755*	0.0092	-0.0714***	-0.6250	0.4284*
	(0.0451)	(0.0453)	(0.0240)	(0.4741)	(0.2417)
Ln rel. dist.	-0.8843***	-0.7723***	-0.5587***	0.1662	-0.4532***
	(0.0538)	(0.0756)	(0.0360)	(0.2534)	(0.1291)
Reg. days & proc.			-0.7673***		0.1428
			(0.0939)		(0.2046)
Reg. costs			-0.2290*		-0.1042
			(0.1207)		(0.2225)
Exch. rate volat.	-4.64×10^{-6}***	14.17×10^{-5}**	16.25×10^{-5}***	0.2049**	-0.0379
	(1.29×10^{-6})	(5.96×10^{-5})	(3×10^{-5})	(0.1004)	(0.0527)
Contiguity	0.8279***	0.4116***	-0.1195		
	(0.1592)	(0.1361)	(0.1111)		
Official language	-0.2458	0.9330***	0.1216*	2.0450***	0.2008
	(0.2206)	(0.1295)	(0.0672)	(0.3512)	(0.1659)
Ethnic language	0.3688*	0.1448	0.3517***	0.4095	0.4529***
	(0.2066)	(0.1267)	(0.0480)	(0.3798)	(0.1592)
Common colonizer	0.8137***	0.5061***	0.3027***	1.1005***	-0.0389
	(0.1098)	(0.0946)	(0.0619)	(0.2289)	(0.1324)
Ln infra_index_imp	6.1594***	5.7437***	1.9083***	9.1719***	3.0751***
	(0.4154)	(0.4564)	(0.2426)	(1.1543)	(0.6259)
Ln infra_index_exp	1.8971***	0.1492	2.6717***	1.2162	5.9379***
	(0.2849)	(0.4789)	(0.2681)	(1.2141)	(0.6947)
SADC		1.6227***	1.0435***	2.3113***	0.3841
		(0.1431)	(0.2669)	(0.5360)	(0.4640)
ECOWAS	0.5602**	1.1958**	0.0338		
	(0.2533)	(0.5487)	(0.2721)		
COMESA	-0.0269	0.3291***	0.3023***	1.9417***	-0.1574
	(0.0874)	(0.0933)	(0.0453)	(0.3036)	(0.1610)
CEMAC	0.4440	0.3043	-0.1291		
	(0.3665)	(0.4241)	(0.2656)		
ρ		0.1366**		0.1411	
		(0.0549)		(0.1589)	
N	3,277	5,863		1,000	
σ	2.3669	2.3618		2.1995	
Number of censored obs		1,878		410	
Number of uncensored obs		3,985		590	

*** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.10 level.
Heckman ML Estimations with a unit elasticity on relative production

4.2 Intra-African border effects vs Europe-Africa border effects

Table 5 provides results of the estimation of the specification (10) for exports from Europe to Africa. The first two columns report results without the infrastructure indices, while the last two columns display results accounting for them. Before commenting on the border effects that apply to European exports in Africa, let us discuss some of the other results. Few of them are actually noteworthy: countries previously involved in a colonial relationship and those sharing a language spoken by at least 5% of the population trade 3 (respectively exp(1.1116) and exp(1.0935)) times more than the sample average.

Surprisingly, memberships to free trade agreements linking African to European countries do not seem to pay off. The coefficient of the dummy for the European Free Trade Association (EFTA)[15] - the Morocco free trade agreement - is not significant in any of the specifications of Table 5. The results for Africa Caribbean Pacific (ACP)[16] - European Union (EU) preferential trade agreement are even more surprising: the dummy coefficient is negative. This is quite counter-intuitive as one would expect that such a preferential trade agreement would boost trade rather than impede it.

According to the last two columns of Table 5 (with infrastructure indices), border effects faced by European exporters in African countries can be estimated to be about 34 (exp(3.5130)). They are higher than the average intra-African border effects (108) but lower than the SADC and the ECOWAS border effects. Hence, while on average African countries seem more open to overseas trade flows than to those coming from their regional partners, regional economic groupings like SADC and ECOWAS seem to be more open to their intra-regional trade flows than to European exports. This finding is consistent with earlier findings showing that, after accounting for the impact of infrastructure, SADC and ECOWAS border effects get closer to international estimations.

Table 5: Gravity and selection equations for trading relationships between European exporters and African importers. Estimation with infrastructure indexes but without tariffs, 1980-2006 averages with common coefficients for all industries and with unit production elasticity.

	1		2	
Variables	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	(Probit) IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	(Probit) IND_{ij}
Border/Intercept	-5.9305***	1.8329***	-3.5130***	1.2862***
	(0.0532)	(0.0416)	(0.0793)	(0.0671)
Ln rel. price	-0.0493*	0.3323***	-0.0569***	0.1064***
	(0.0258)	(0.0122)	(0.0179)	(0.0130)
Ln rel. dist.	-0.9245***	-0.0338***	-1.2948***	-0.1152***
	(0.0143)	(0.0077)	(0.0145)	(0.0114)
Colonial relationship	1.5657***	1.0951***	1.1116***	1.0571***
	(0.0594)	(0.1828)	(0.0534)	(0.1948)
Col. relat. post 1945	-0.5757***	0.2312	1.76×10^{-5}	0.2089
	(0.0690)	(0.2449)	(0.0658)	(0.2946)
Reg. days & proc.		-0.6434***		-0.9358***
		(0.0379)		(0.2031)
Reg. costs		-0.1241***		-0.8478***
		(0.0201)		(0.2217)
Official language	0.1728**	-0.4498***	-0.5480***	0.0155
	(0.0785)	(0.0847)	(0.0768)	(0.1075)
Ethnic language	0.4114***	0.4974***	1.0935***	0.2242**
	(0.0734)	(0.0775)	(0.0719)	(0.1000)
Ln infra_index_imp			5.5967***	1.1019***
			(0.1005)	(0.0951)
Ln infra_index_exp			0.5887***	1.8828***
			(0.0545)	(0.0407)
EFTA_MOROCCO	-0.0254	-1.0497***	-0.4455	-0.8298**
	(0.5404)	(0.3338)	(0.4986)	(0.3533)
UE_ACP	-0.7418***	0.4686***	-0.1584***	0.0074
	(0.0313)	(0.0218)	(0.0286)	(0.0218)
ρ	-0.6822***		0.0048	
	(0.0485)		(0.0192)	
N	42,936			
σ	2.3829		2.1496	
Number of censored obs	4,697			
Number of uncensored obs	38,239			

*** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.10 level.
Heckman ML Estimations with an unrestricted elasticity on relative production

5 Conclusion

The results show that on average the African continent is poorly integrated and more open to overseas export than to intra-African trade flows. However, this negative picture is contrasted by the fact that two African RECs, SADC and ECOWAS have border effects that are more in line with international estimations. Hence, there is evidence that these two African RECs are effective in promoting intraregional trade. RECs are expected to increase trade between their members via three channels. The first is a reduction in tariffs between members; the second is a reduction in NTBs; the third is via the two components of 'trade facilitation': a 'hard' component related to tangible transport and telecommunications infrastructure; and a 'soft' component related to transparency, the business environment, customs management, and other intangible institutional aspects that may facilitate trading (de Melo and Tsikata, 2014).

The first two channels are the outcomes of measures that are easier to implement and have generally been undertaken even by the African RECs that only manage to achieve 'shallow' integration (de Melo and Tsikata, 2014). However, the measures implied by the two components of 'trade facilitation' are much more difficult to put in place. This is especially true for the infrastructure dimension. Our model explicitly account for that 'hard' component. Considering infrastructure indices is an interesting way to capture the effect of distribution networks which represent, along with imperfect information and localized tastes, sources of resistance that are generally omitted and can possibly explain the counter-intuitive high border effects (Head and Mayer, 2013*b*).

The elasticity of the infrastructure indices are high and significant, suggesting that improving transport and telecommunications infrastructures will go a long way in promoting inter-African trade. So far only Northern and Southern Africa are well endowed in this regard, and it is clear that building and improving infrastructure in other parts of the continent will generate additional trade opportunities. The 'soft' component of trade facilitation is likely to be the factor explaining the discrepancies between the different African RECs in terms of border effects. In this regard SADC and ECOWAS seem to have set institutions that are more effective in achieving trade integration and allow them to be one step ahead comparatively to other African RECs.

Yet, even these two RECs are still behind Regional Trade Agreements (RTA) in other continents in terms of the importance of intra-regional trade share in GDP. While the SADC intra-regional trade share in GDP – on average one of the highest in the continent – rose from 6% to 11% from 1992 to 2013, the share of intra-RTA trade worldwide, excluding the EU, increased from 18% in 1990 to 34% in 2008 (from 28% to 51% if EU included). So there is still room for improvement in African RECs regarding this 'soft' component of trade facilitation. Actually, regional integration in Africa was based on the 'linear model' of integration, with a stepwise integration of goods, labour, and capital markets, as well as eventual monetary and fiscal integration. Most of African countries overlooked the importance of tackling 'behind-the-border' impediments to trade, yet this is crucial in global environment characterized by the reduction in trade costs and the subsequent fragmentation of production (de Melo and Tsikata, 2014).

Based on our results, two propositions may help reduce border effects and increase inter-regional trade in Africa. First is the development of large pan-African infrastructure projects which can contribute to defragmenting Africa by decreasing transport costs. Second is the implementation of the African free trade zone that might bring free trade among the members by: removing tariffs and NTBs and implementing trade facilitation, by applying the subsidiarity principle to infrastructure to enhance the transport network, and by fostering industrial development (de Melo and Tsikata, 2014). These measures will allow African countries to achieve a 'deep' rather than a 'shallow integration'.

Finally, we have to acknowledge the limitations of the study. The main challenges are the quality and the availability of the African data. Ideally, anyone would have opted for a more direct way to capture the impact of distribution networks. Combes et al. (2005) for instance examine the impact of business and social networks on trade between French regions by using the financial structure and location of French firms as well as the bilateral stocks of migrants. Moreover, Combes et al. (2005) also use data on transport costs which allow them to separate the effects of transport infrastructure and administrative border effects.

Another issue that we are facing is that we rely on data on trade flows between different countries. This implies that in order to estimate border effects we cannot use LSDV estimation; we rely instead on the complete odds specification which requires a measure of the trade of a country with itself ("trade with self"). Generally, this "trade with self" is proxied by using production minus total exports. While, the complete odds specification has been used by several papers (Head and Mayer, 2000; de Sousa et al., 2012), the measure of "trade with self" may cause some measurement errors as the procedure may generate some negative observations for some countries (Head and Mayer, 2013*a*). The ideal would be to use databases on commodity flows like in Wolf (2000), Anderson and van Wincoop (2003), Hillberry and Hummels (2003), and Combes et al. (2005) which include inter- and intraregional trade flows.

However, it can be argued that this kind of datasets represents the exception rather than the norm. Moreover, in the current state of the data collection on African countries, it will be challenging to get reliable data on the missing sources of resistance. Getting such data is yet the price to pay to improve on the analysis of border effects in intra-African trade.

Acknowledgements. Alain Pholo Bala gratefully acknowledges financial support from Economic Research Southern Africa (ERSA), South Africa. The views expressed in this paper and any remaining errors are ours.

Notes

[1] With data on inter-provincial trade flows, one can estimate the border effects with a fixed effects specification (Feenstra, 2004).

[2] Head and Mayer (2013*a*) offers a more in-depth discussion of those models.

[3]These results might be explained by the fact that in specifications (9) and (10), we are dealing with ratios of trade flow from j to i to trade flow of i from "self" rather than trade flows per se.

[4]Available through the following download page: `http://www.cepii.fr/anglaisgraph/bdd/TradeProd.htm`. In this webpage it is recommended to cite de Sousa et al. (2012)' reference as the source of the data.

[5]One of the file of the GeoDist database, the dist_cepii dataset contains 2 variables indicating whether two countries, origin and destination, share a common official language, or a common ethnic language, i.e. a language that is spoken by at least 9% of the population in both countries.

[6]Details on the formulas of `distw` and `distwces` are given in Mayer and Zignago (2011, p. 11).

[7]Infrastructure data may be found at the following link:
`http://thedata.harvard.edu/dvn/dv/pep/faces/study/StudyPage.xhtml?globalId=hdl:1902.1/11953`.

[8]It would have been useful to estimate the gravity model separately for different industries as in Head and Mayer (2000) and de Sousa et al. (2012), but then we would have ended up with small numbers of observations for some industries.

[9]Head and Mayer (2000) suggest that those low estimates are due to the unobserved variation in relative product quality which is correlated with relative product price.

[10]Border effects within other African RECs are not statistically different from the sample average.

[11]While for the second estimate, we would obtain a tariff equivalent respectively equal to exp (7.7622/8)-1=164% for intra-African trade, exp (5.9084/8)-1=109% for SADC, exp (7.2915/8)-1=149% for ECOWAS andexp (7.8808/8)-1=168% for COMESA. With the same value of the elasticity of trade with respect to trade costs, de Sousa et al. (2012) found a tariff equivalent of 118% between Southern exporters and Northern importers. Not surprisingly this finding confirms the previous finding that SADC is the only REC within Africa, where members have easier access to the markets of their partners than to overseas markets.

[12]While NTBs like quotas, voluntary export restraints and non-automatic import authorizations often have much more significant effects, only a few databases provide information on non-tariff barriers (e.g. the Doing Business database, the Trade Facilitation Database, the Logistic Performance Index, the UNCTAD/TRAINS NTM and the CEPII NTM-MAP databases). The cause of the rarity of databases on NTBs is largely due to the difficulty in collecting the data and in assembling a consistent cross-country dataset. Unlike tariffs, NTB data are not merely numbers; the relevant information is often hidden in legal and regulatory documents. Furthermore, these documents are generally not centralized but often reside in different regulatory agencies. All these issues make the collection of NTB data a very resource-intensive task. The CEPII NTM-MAP database has been designed to address the shortcomings of the previous databases, yet its index is only provided for 63 countries for one year over the period 2010-2012 at the country level and two different product disaggregation levels. Moreover, only 15 African countries are included in the database (Gourdon, 2014).

[13]The concept of 'dark' costs results from an analogy with the cosmology concept of dark energy or dark matter. Neither dark energy nor dark matter can be observed directly, but their presence is inferred to be tremendous. Therefore, 'dark' costs represent unobserved or difficult-to-observe trade costs which imply a significant resistance to trade flows. More information about the astrophysics analogy between 'dark' cost and dark energy or dark matter is available in Head and Mayer (2013*b*).

[14]The 'dark' costs computed by Head and Mayer (2013*b*) include everything other than tariffs. So they may capture border compliance costs which may be evaluated (the so-called 'grey' costs ψ_c). Head and Mayer (2013*b*) acknowledge that their presentation of the concept of dark costs missed that distinction. However, since they report custom clearance costs which seem to be negligible (1.8% for EU, and 2.0%-3.4% for the Canada-USA frontier), the 'dark' costs seem to be dominant. Eventually, while the specifications (9) and (10) account for bilateral tariffs, Table 1 does not. So we may still need to deduct the bilateral tariffs from the estimated 'dark' costs. If we use the simple average world tariff (12.5%) as an approximation (Head and Mayer, 2013*b*), the intra-African trade dark costs are only slightly affected.

[15]The EFTA is an intergovernmental organization set up for the promotion of free trade and economic integration to

the benefit of its four Member States: Iceland, Liechtenstein, Norway, Switzerland (http://www.efta.int).

[16] "The African, Caribbean and Pacific Group of States (ACP) is an organization created by the Georgetown Agreement in 1975. It is composed of 79 African, Caribbean and Pacific states, with all of them, save Cuba, signatories to the Cotonou Agreement, also known as the "ACP-EC Partnership Agreement" which binds them to the European Union"(http://www.acp.int/content/secretariat-acp).

References

Ancharaz, V., Mbekeani, K. and Brixiova, Z. (2011), 'Impediments to Regional Trade Integration in Africa', *Africa Economic Brief* **2**(11).

Anderson, J. E. and van Wincoop, E. (2003), 'Gravity with Gravitas: A Solution to the Border Puzzle', *American Economic Review* **93**(1), 170–192.

Artadi, E. V. and Sala-i Martin, X. (2003), The Economic Tragedy of the XXth Century: Growth in Africa, NBER Working Papers 9865, National Bureau of Economic Research, Inc.

Au, C.-C. and Henderson, J. V. (2006), 'Are Chinese Cities Too Small?', *Review of Economic Studies* **73**(3), 549–576.

Babarinde, F. (2003), Regionalism and economic development, *in* E. U. Nnadozie, ed., 'African economic development', Amsterdam : London : Academic, book; book/illustrated 19, pp. 473–497.

Baldwin, R. and Harrigan, J. (2011), 'Zeros, Quality, and Space: Trade Theory and Trade Evidence', *American Economic Journal: Microeconomics* **3**(2), 60–88.

Behrens, K., Ertur, C. and Koch, W. (2012), 'Dual Gravity: Using Spatial Econometrics To Control For Multilateral Resistance', *Journal of Applied Econometrics* **27**(5), 773–794.

Canning, D. (1998), 'A database of world stocks of infrastructure, 195095', *The World Bank Economic Review* **12**(3), 529–547.

Carrère, C. (2004), 'African regional agreements: Impact on trade with or without currency unions', *Journal of African Economies* **13**(2), 199–239.

Carrère, C. (2006), 'Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model', *European Economic Review* **50**(2), 223–247.

Cinyabuguma, M. M. and Putterman, L. (2011), 'Sub-Saharan Growth Surprises: Being Heterogeneous, Inland and Close to the Equator Does not Slow Growth Within Africa-super- ', *Journal of African Economies* **20**(2), 217–262.

Collier, P. (2006*a*), 'Africa : geography and growth', *Proceedings - Economic Policy Symposium - Jackson Hole* pp. 235–252.

Collier, P. (2006*b*), 'African Growth: Why a 'Big Push'?', *Journal of African Economies* **15**(2), 188–211.

Combes, P.-P., Lafourcade, M. and Mayer, T. (2005), 'The trade-creating effects of business and social networks: evidence from france', *Journal of International Economics* **66**(1), 1 – 29.

de Melo, J. and Tsikata, Y. (2014), Regional integration in Africa: Challenges and prospects, Working Papers P93, FERDI.

de Sousa, J., Mayer, T. and Zignago, S. (2012), 'Market access in global and regional trade', *Regional Science and Urban Economics* **42**(6), 1037–1052.

Disdier, A.-C. and Head, K. (2008), 'The Puzzling Persistence of the Distance Effect on Bilateral Trade', *The Review of Economics and Statistics* **90**(1), 37–48.

Dollar, D. and Easterly, W. (1999), 'The Search for the Key: Aid, Investment and Policies in Africa', *Journal of African Economies* **8**(4), 546–77.

Easterly, W. (2009*a*), 'Can the West Save Africa?', *Journal of Economic Literature* **47**(2), 373–447.

Easterly, W. (2009*b*), 'How the Millennium Development Goals are Unfair to Africa', *World Development* **37**(1), 26–35.

Easterly, W. and Levine, R. (1997), 'Africa's Growth Tragedy: Policies and Ethnic Divisions', *The Quarterly Journal of Economics* **112**(4), 1203–50.

Eaton, J. and Kortum, S. (2001), 'Trade in capital goods', *European Economic Review* **45**(7), 1195–1235.

Eaton, J. and Kortum, S. (2002), 'Technology, Geography, and Trade', *Econometrica* **70**(5), 1741–1779.

Eaton, J., Kortum, S. S. and Sotelo, S. (2012), International Trade: Linking Micro and Macro, NBER Working Papers 17864, National Bureau of Economic Research, Inc.

Eaton, J. and Tamura, A. (1994), 'Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns', *Journal of the Japanese and International Economies* **8**(4), 478–510.

Economic Commission for Africa (2011), "Boosting Intra-African Trade". Preparations for the 7th AU Trade Ministers Conference and the 18th Ordinary Session of the AU Assembly of Heads of State and Government, Issues paper, Economic Commission for Africa.

Erkel-Rousse, H. and Mirza, D. (2002), 'Import price elasticities: reconsidering the evidence', *Canadian Journal of Economics/Revue canadienne d'économique* **35**(2), 282–306.

Evans, C. L. (2003), 'The Economic Significance of National Border Effects', *American Economic Review* **93**(4), 1291–1312.

Feenstra, R. C. (2004), *Advanced international trade: theory and evidence*, Princeton University Press.

Freire, M. E. M., Lall, S. and Leipziger, D. (2015), Africa's urbanization: Challenges and opportunities, *in* C. Monga and J. Y. Lin, eds, 'The Oxford Handbook of Africa and Economics', Vol. 1: Context and Concepts, Oxford University Press, chapter 31, pp. 584–602.

Gourdon, J. (2014), CEPII NTM-MAP: A Tool for Assessing the Economic Impact of Non-Tariff Measures, Working Papers 2014-24, CEPII research center.

Grossman, G. M. (1998), Comment on: Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?, *in* J. A. Frankel, ed., 'The Regionalization of the World Economy', University of Chicago Press, pp. 29–31.

Harrigan, J. (1996), 'Openness to trade in manufactures in the OECD', *Journal of International Economics* **40**(1-2), 23–39.

Head, K. and Mayer, T. (2000), 'Non-Europe: The magnitude and causes of market fragmentation in the EU', *Review of World Economics (Weltwirtschaftliches Archiv)* **136**(2), 284–314.

Head, K. and Mayer, T. (2013*a*), Gravity Equations: Workhorse,Toolkit, and Cookbook, Working Papers 2013-27, CEPII research center.

Head, K. and Mayer, T. (2013*b*), 'What separates us? sources of resistance to globalization', *Canadian Journal of Economics/Revue canadienne d'conomique* **46**(4), 1196–1231.

Head, K. and Ries, J. (2001), 'Increasing returns versus national product differentiation as an explanation for the pattern of u.s.-canada trade', *American Economic Review* **91**(4), 858–876.

Helliwell, J. F. (1996), 'Do National Borders Matter for Quebec's Trade?', *Canadian Journal of Economics* **29**(3), 507–22.

Helpman, E., Melitz, M. and Rubinstein, Y. (2008), 'Estimating trade flows: Trading partners and trading volumes', *The Quarterly Journal of Economics* **123**(2), 441–487.

Hillberry, R. (1998), 'Regional trade and the medicine line: The national border effect in u.s. commodity flow data', *Journal of Borderlands Studies* **13**(2), 1–17.

Hillberry, R. and Hummels, D. (2003), 'Intranational Home Bias: Some Explanations', *The Review of Economics and Statistics* **85**(4), 1089–1092.

Hummels, D. (1999), Toward a Geography of Trade Costs, GTAP Working Papers 1162, Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University.

ITC (2012), *Africa's Trade Potential: Export Opportunities in Growth Markets*, International Trade Centre (ITC).

Jones, B. (2002), 'Economic integration and convergence of per capita income in west africa', *African Development Review* **14**(1), 18–47.

Krugman, P. (1980), 'Scale Economies, Product Differentiation, and the Pattern of Trade', *American Economic Review* **70**(5), 950–59.

Lai, H. and Trefler, D. (2002), The Gains from Trade with Monopolistic Competition: Specification, Estimation, and Mis-Specification, NBER Working Papers 9169, National Bureau of Economic Research, Inc.

Limão, N. and Venables, A. J. (2001), 'Infrastructure, geographical disadvantage, transport costs, and trade', *The World Bank Economic Review* **15**(3), 451–479.

Mayer, T. and Zignago, S. (2011), Notes on CEPIIs distances measures: The GeoDist database, Working Papers 2011-25, CEPII research center.

McCallum, J. (1995), 'National Borders Matter: Canada-U.S. Regional Trade Patterns', *American Economic Review* **85**(3), 615–23.

Page, J. (2008), *Rowing Against the Current: The Diversification Challenge in Africa's Resource-Rich Economies*, Brookings Global Economy & Development Working Papers No. 28 December 2008.

Pinkovskiy, M. and Sala-i Martin, X. (2014), 'Africa is on time', *Journal of Economic Growth* **19**(3), 311–338.

Redding, S. and Venables, A. J. (2004), 'Economic geography and international inequality', *Journal of International Economics* **62**(1), 53–82.

Sachs, J. D. and Warner, A. M. (1997), 'Sources of Slow Growth in African Economies', *Journal of African Economies* **6**(3), 335–76.

Simonovska, I. and Waugh, M. E. (2014), 'The elasticity of trade: Estimates and evidence', *Journal of International Economics* **92**(1), 34 – 50.

UNECA (2010), *Assessing Regional Integration in Africa IV: Enhancing Intra-African Trade*, United Nations Economic Commission for Africa.

Wei, S.-J. (1996), Intra-National versus International Trade: How Stubborn are Nations in Global Integration?, NBER Working Papers 5531, National Bureau of Economic Research, Inc.

Wolf, H. C. (2000), 'Intranational home bias in trade', *The Review of Economics and Statistics* **82**(4), 555–563.

World Bank (2006), *World Development Indicators 2006*, number 8151 *in* 'World Bank Publications', The World Bank.

World Bank (2008), *World Development Report 2009: Reshaping Economic Geography*, World Bank.

Appendix A: Descriptive statistics of continuous variables

Table 6: Descriptive statistics (All sample).

Variable	Description	N	Min	Max	Mean	S.D.
Variables in level						
flow	Bilateral trade flow	15,459,569	0	5.85×10^8	23,802.32	1,383,990
prodimp	Production importer	6,496,323	0	6.39×10^8	7,987,597	3.22×10^7
prodexp	Production exporter	6,814,040	0	6.39×10^8	8,689,362	3.31×10^7
intflow	Internal production	5,284,601	0.1547	5.82×10^8	7,208,616	2.83×10^7
distw	Weighted distance ($\theta = 1$)	13,832,099	8.4497	19,781.39	8,140.105	4,671.91
distw_int	Internal weighted distance ($\theta = 1$)	14,379,902	0.9951	1,853.802	258.2357	290.4453
pcexp	Consumption price index exporter	13,038,951	4.6107	10,557.49	63.1952	49.9450
pcimp	Consumption price index importer	13,023,079	4.6107	10,557.49	63.0925	65.4607
tariff	Tariff (%)	295,988	0	1,738.801	22.6809	57.3162
tariff_factor	1+tariff/100	295,988	1	18.3880	1.2268	0.5732
vol_bilat_exch	Volatility of bilateral exchange rate	59,166	0	562,060.9	358.0921	11,791.4
infraindex_imp	Infrastructure index importer	2,843,938	-0.3986	8.1165	0.3643	1.0591
infraindex_exp	Infrastructure index importer	2,911,319	-0.3986	8.1165	0.4565	1.0330
Ratios						
rel_flow	Relative flow: flow/intflow	5,284,601	0	24,846.42	0.0820	15.7467
rel_prod	Relative production: prodexp/prodimp	2,971,810	0	3.25×10^8	2,534.246	254,362.9
rel_price	Relative price: pcexp/pcimp	10,853,016	0.0017	602.609	1.2392	1.2182
rel_distw	Relative distance: distw/distw_int	13,315,039	0.6124	19,743.68	174.3733	726.5353
Variables in logs						
lflow	log of trade flows	50,465,640	-7.032	20.1865	5.1512	3.3706
lntariff_factor	log of tariff factor	295,988	0	2.9117	0.1818	0.1624
lrel_flow	log of relative flow	2,511,776	-26.2741	10.1205	-7.7154	3.9299
lrel_prod	log of relative production	2,927,976	-19.5994	19.5994	0.3009	3.9509
dlrel_flow	log of relative flow - log of relative prod	1,840,917	-24.7973	9.7522	-7.4274	3.3306
lrel_price	log of relative production	10,853,016	-6.4013	6.4013	0.0072	0.6404
lrel_distw	log of relative distance	10,853,016	-6.4013	6.4013	0.0072	0.6404

Table 7: Descriptive statistics (Intra-African trade).

Variable	Description	N	Min	Max	Mean	S.D.
Variables in level						
flow	Bilateral trade flow	736,490	0	4.65×10^7	4,560.501	157,386.7
prodimp	Production importer	172,866	0	4.67×10^7	612,613	2,226,487
prodexp	Production exporter	179,974	0	4.67×10^7	610,487.1	2,187,505
intflow	Internal production	139,280	1.1893	4.65×10^7	593,543.8	2,189,699
distw	Weighted distance ($\theta = 1$)	736,490	18.9908	962.9211	3,548.438	1,962.716
distw_int	Internal weighted distance ($\theta = 1$)	736,490	18.9908	962.9211	250.3121	178.9875
pcexp	Consumption price index exporter	716,149	21.5687	858.5782	63.1952	49.9450
pcimp	Consumption price index importer	716,332	21.5687	858.5782	57.8241	34.3039
tariff	Tariff (%)	68,537	0	1,738.801	20.8420	56.5255
tariff_factor	1+tariff/100	68,537	1	18.3880	1.2084	0.5653
vol_bilat_exch	Volatility of bilateral exchange rate	59,166	0	562,060.9	358.0921	11,791.4
infraindex_imp	Infrastructure index importer	118,634	-0.3986	0.0945	-0.3215	0.0820
infraindex_exp	Infrastructure index importer	123,703	-0.3986	0.0945	-0.3208	0.0793
Ratios						
rel_flow	Relative flow: flow/intflow	139,280	0	3,219.248	0.1212	9.0058
rel_prod	Relative production: prodexp/prodimp	49,730	0	1,429,653	241.3158	7,425.665
rel_price	Relative price: pcexp/pcimp	696,001	0.0443	17.6309	1.1361	0.7138
rel_distw	Relative distance: distw/distw_int	736,490	0.9269	514.5665	32.1975	60.4278
Variables in logs						
lflow	log of trade flows	168,439	-6.9695	17.6553	3.9221	3.1385
lntariff_factor	log of tariff factor	68,537	0	2.9117	0.1670	0.1624
lrel_flow	log of relative flow	47,208	-23.3191	8.0769	-6.5959	4.3440
lrel_prod	log of relative production	48,348	-14.1729	14.1729	0.1351	3.0901
dlrel_flow	log of relative flow - log of relative prod	25,648	-19.0141	5.7305	-5.3161	4.1811
lrel_price	log of relative production	696,001	-3.1157	2.8697	0.0016	0.4877
lrel_distw	log of relative distance	736,490	-0.0760	6.2433	2.7272	1.1291

Table 8: Descriptive statistics (Trade within SADC).

Variable	Description	N	Min	Max	Mean	S.D.
Variables in level						
flow	Bilateral trade flow	34,439	0	2.68×10^7	27,924.76	445,407.2
prodimp	Production importer	11,035	0	2.94×10^7	1,102,419	2,980,809
prodexp	Production exporter	11,116	0	2.94×10^7	1,062,612	2,947,129
intflow	Internal production	8,854	1.1893	2.68×10^7	1,045,166	2,693,852
distw	Weighted distance ($\theta = 1$)	34,439	18.9908	4,828.119	2,172.012	1,060.675
distw_int	Internal weighted distance ($\theta = 1$)	34,439	18.9908	802.9943	343.1269	240.4223
pcexp	Consumption price index exporter	34,439	30.6960	762.7068	72.1984	52.9030
pcimp	Consumption price index importer	34,439	30.6960	762.7068	71.0756	51.8350
tariff	Tariff (%)	4,955	0	200	14.1709	17.7327
tariff_factor	1+tariff/100	4,955	1	3	1.1417	0.1773
vol_bilat_exch	Volatility of bilateral exchange rate	4,244	2.99×10^{-7}	3,655.614	52.2227	348.9442
infraindex_imp	Infrastructure index importer	3,574	-0.3777	0.0945	-0.2778	0.1150
infraindex_exp	Infrastructure index importer	3,997	-0.3777	0.0945	-0.2913	0.1074
Ratios						
rel_flow	Relative flow: flow/intflow	8,854	0	462.4796	0.4180	6.5818
rel_prod	Relative production: prodexp/prodimp	4,058	0	111,402.3	336.019	3,938.076
rel_price	Relative price: pcexp/pcimp	34,439	0.1345	9.3570	1.2923	1.0829
rel_distw	Relative distance: distw/distw_int	34,439	1	254.2339	31.8623	58.6936
Variables in logs						
lflow	log of trade flows	17,364	-6.9078	17.1045	5.2206	3.4257
lntariff_factor	log of tariff factor	4,955	0	1.0986	0.1234	0.1277
lrel_flow	log of relative flow	5,713	-22.4508	6.1366	-5.7621	4.5720
lrel_prod	log of relative production	4,047	-11.6209	11.6209	-0.0181	3.2858
dlrel_flow	log of relative flow - log of relative prod	2,807	-13.9850	3.4003	-3.5035	3.3977
lrel_price	log of relative production	34,439	-2.0062	2.2361	0.0104	0.6905
lrel_distw	log of relative distance	34,439	0	5.5383	2.1638	1.4619

Appendix B: Estimation of gravity models for intra-African trade with bilateral tariffs

Table 9: Gravity and selection equations for intra-African trading relationships. Estimation with tariffs and without infrastructure index, 1989-2001 averages with common coefficients for all industries.

	1		**2**	
Variables	$\ln\left(\frac{m_{ij}}{m_{ii}}\right)$	(Probit) IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	(Probit) IND_{ij}
Border/Intercept	-8.5816***	2.2190***	-7.9772***	2.2613***
	(0.2440)	(0.1476)	(0.2473)	(0.1511)
Ln rel. prod	0.6461***	0.0609***		
	(0.0195)	(0.0067)		
Ln rel. price	-0.0882	0.1303***	-0.5558***	0.2095***
	(0.1127)	(0.0415)	(0.1145)	(0.0408)
Ln tariff factor	-0.8559*	-0.3275***	-0.9054*	-0.3568***
	(0.4869)	(0.1175)	(0.5305)	(0.1216)
Ln rel. dist	-0.0206***	-0.3168***	-0.2691***	-0.3200***
	(0.0705)	(0.0238)	(0.0694)	(0.0239)
Reg. days & proc.		-0.9952***		-1.0716***
		(0.1021)		(0.1017)
Reg. costs		-0.9001***		-0.9702***
		(0.0750)		(0.0754)
Exch. rate volat.	-1.78×10^{-4}	-0.0003***	12.9×10^{-6}	-0.0003***
	(3.13×10^{-4})	(0.0001)	(0.0003)	(7.71×10^{-5})
Contiguity	1.6407***	0.3235***	1.4390***	0.2918***
	(0.1730)	(0.0851)	(0.1857)	(0.0848)
Official language	0.3156**	0.1332***	0.5197***	0.1118***
	(0.1412)	(0.0456)	(0.1547)	(0.0454)
Ethnic language	0.5644***	0.5125***	0.5827***	0.5464***
	(0.1344)	(0.0435)	(0.1450)	(0.0435)
Common colonizer	0.8713***	0.0988**	0.7406***	0.1061**
	(0.1250)	(0.0492)	(0.1323)	(0.0489)
SADC	1.6632***	0.6911***	2.3010***	0.6150***
	(0.1661)	(0.0631)	(0.1718)	(0.0619)
ECOWAS	1.4051*	0.5726**	1.7619*	0.5159*
	(0.7630)	(0.2724)	(0.9457)	(0.2972)
UEMOA	-2.4782**	-1.0719***	-3.1513**	-0.9704*
	(1.1107)	(0.5026)	(1.3963)	(0.5043)
COMESA	0.0485	-0.2497***	-0.6574***	-0.0865**
	(0.1301)	(0.0446)	(0.1275)	(0.0405)
CEMAC	-0.2564	-0.4826*	-0.7485	-0.3737
	(0.6295)	(0.2876)	(0.5958)	(0.2920)
ρ	-0.0606		0.0743***	
	(0.0682)		(0.0542)	
N	5,916			
σ	2.4326		2.6244	
Number of censored obs	3,313			
Number of uncensored obs	2,603			

*** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.10 level.
(1): Heckman ML Estimation with an unrestricted elasticity on relative production
(2): Heckman ML estimation with unit elasticity on relative production

Table 10: Gravity and selection equations for intra-African trading relationships. Estimation with tariffs and infrastructure indexes, 1989-2001 averages with common coefficients for all industries and with unit production elasticity.

	1		2	
		(Probit)		(Probit)
Variables	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	IND_{ij}	$\ln\left(\frac{m_{ij}/\nu_j}{m_{ii}/\nu_i}\right)$	IND_{ij}
Border/Intercept	-10.0156***	4.3931***	-7.3662***	5.8431***
	(0.6282)	(0.4434)	(0.8101)	(0.5250)
Ln rel. price	-0.5822***	0.1014	-0.2939	0.1904**
	(0.1949)	(0.0741)	(0.1977)	(0.0741)
Ln tariff factor	-0.2078	-0.2125	-0.0862	-0.1915
	(0.6786)	(0.1872)	(0.6315)	(0.1865)
Ln rel. dist.	0.2902	-0.8164***	-0.1287	-0.9708***
	(0.2180)	(0.0912)	(0.2203)	(0.1041)
Reg. days & proc.		-1.1890***		-0.9358***
		(0.1773)		(0.2031)
Reg. costs		-1.4378***		-0.8478***
		(0.1800)		(0.2217)
Exch. rate volat.	59.48×10^{-5}*	-42.93×10^{-5}***	29.56×10^{-5}	-36.01×10^{-5}***
	(35.95×10^{-5})	(10.11×10^{-5})	(38.85×10^{-5})	(10.19×10^{-5})
Contiguity	0.5842	-1.2244***	0.4469	-1.0667***
	(0.3623)	(0.2463)	(0.3615)	(0.2628)
Official language	0.8918***	0.0511	0.7690***	-0.0682
	(0.2554)	(0.1368)	(0.2485)	(0.1413)
Ethnic language	0.7263***	0.2823***	0.9016***	0.3698***
	(0.2618)	(0.0947)	(0.2624)	(0.0982)
Common colonizer	1.1841***	-0.0329	1.3262***	0.1335
	(0.2144)	(0.1269)	(0.2110)	(0.1290)
Ln infra_index_imp			6.3605***	2.8383***
			(1.0024)	(0.5303)
Ln infra_index_exp			0.2365	2.2539***
			(0.9632)	(0.5995)
SADC	4.6359***	1.3726***	3.4186***	0.5711
	(0.5118)	(0.5210)	(0.5199)	(0.5016)
ECOWAS	-0.0582	-0.9550*	0.2049	-0.8797
	(1.4666)	(0.5436)	(1.4568)	(0.5425)
COMESA	-0.4794*	1.0570***	0.4258	1.1697***
	(0.2727)	(0.1061)	(0.3101)	(0.1169)
CEMAC	0.7351	0.5337	0.9629	0.3585
	(0.6363)	(0.3659)	(0.6444)	(0.3730)
ρ	-0.1064		0.1740***	
	(0.0859)		(0.0985)	
N	1,616			
σ	2.4176		2.3747	
Number of censored obs	673			
Number of uncensored obs	943			

*** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.10 level.
Heckman ML Estimations with an unrestricted elasticity on relative production