New panel data evidence on Sub-Saharan trade integration - Prospects for the COMESA-EAC-SADC Tripartite*

Jana Riedel Anja Slany

June 19, 2015

Abstract

In the year 2008 the member states of the three major trading blocs in southern and eastern Africa agreed on establishing a common free trade area (FTA). This so-called COMESA-EAC-SADC Tripartite is supposed to be an important milestone towards Africa's continental trade integration. This study analyzes the impact of regional integration among the Tripartite countries on their bilateral exports and evaluates the latest integration efforts. We estimate an extended gravity model on a panel data set using yearly observations from 1995 to 2010. Specifically, we apply two approaches to proxy limited market access and effectively applied tariff rates. Therefore, we combine Sub-Saharan- and country-average import and export tariff rates and indicator variables for the membership in regional FTAs to isolate distinct effects on real exports. We find a robust and significantly negative effect of tariff barriers with respect to the rest of the world. Interestingly, an FTA status does not show any export enhancing effect. From a methodological point of view, we detect a bias in our loglinear specifications when we compare the estimates with those we obtain from Poisson pseudo-maximum likelihood estimation.

Keywords: trade union, Africa, Hausman-Taylor, panel data, Poisson pseudo-maximum likelihood, tariff barrier

JEL classification: F13, F14, F15, C23, C26

---

*Jana Riedel: Westfälische Wilhelms-Universität Münster, Wirtschaftswissenschaftliche Fakultät, Institut für Internationale Ökonomie, Universitätsstraße 14-16, 48143 Münster, Germany, email: jana.riedel@wiwi.uni-muenster.de; Anja Slany: Ruhr-Universität Bochum, Lehrstuhl für Internationale Wirtschaftsbeziehungen, Universitätstraße 150, 44801 Bochum, Germany, email: anja.slany@rub.de. We thank Matthias Busse, Bernd Kempa, and Philipp Dybowski, and the participants of the ETSG 16th Annual Conference 2014, the EABEW 1st Annual Conference 2015, the International Conference on Globalization and Development 2015, and the Annual Conference of the Research Group on Development Economics of the Verein für Socialpolitik 2015 for valuable comments on an earlier version of this paper.

1 Introduction

Over the last two decades world merchandise trade has more than tripled, accelerated by a large increase of South-South and North-South trade. Today, developing countries make up 45 percent of world trade, generated to a large extent by Asian countries and Latin America. Despite an increasing volume of African exports and imports, the share of African trade in world trade in 2013 is still low at 3.2 percent (UNCTAD (2014)). Aside from this, many African economies increasingly focus on regional integration as a strategy to effectively promote economic independence. The process of African regional integration received a big impulse in October 2008 when three of the trading blocs, the Common Market for Eastern and Southern Africa (COMESA), the East African Community (EAC) and the Southern African Development Community (SADC), agreed on forming a trading bloc free of tariffs, quotas and exemptions that combines the already existing free trade agreements. By now the 26 member states of this so-called Tripartite already make up more than 50 percent of the total GDP of the African Union. In 2011 a declaration that launches the negotiations on the Tripartite FTA was signed by 24 member states. In June 2015 the establishment of the FTA became reality for the goods market. 15 member states already signed the political declaration, all other countries are expected to sign within the next 12 months. They also agreed on a road map on negotiating outstanding issues such as trade in services and co-operation in trade and development (COMESA (2015)). The agreement promises an increase in intra-regional trade by generating a larger market and overcoming the problems of multi-membership.

However, the three blocs themselves are at different stages of their integration process. The COMESA was formed in 1994 as a successor organization of the so-called Preferential Trade Area and by now consists of 18 countries.[1] The organization aims at realizing a large economic trading bloc in order to overcome individual countries' barriers to trade. After the implementation of an FTA in nine member states in 2000, a customs union was launched in 2009, though a list of sensitive products has been subject to a so-called 'common external tariff' within a transition period. This common tariff is already harmonized with the tariff rate of the EAC (Othieno & Shinyekwa (2011)). The EAC was founded in 2000. In 2005, the member states Tanzania, Kenya and Uganda formed a customs union that was transformed into a common market in 2010; Rwanda and Burundi joined in 2007. The members have fully liberalized the goods sector and only face remaining tariffs in a few service sectors. The third regional economic community, the SADC, exists since 1992 and today consists of 15 countries. In 2008 an FTA was established including the Southern African Customs Union members who allow tariff-free imports from the SADC members. Full tariff liberalization is only provided on 85 percent of intra-SADC trade.[2]

The intra-Tripartite trade share of total Tripartite exports is still low at 15 percent compared to other developing regions. However, the total bilateral exports between Tripartite countries experienced an increase of 30 percent between 2000 and 2010 after a slight downward trend at the end of the 1990s. In contrast, total exports to the rest of the world increased by only 3.7 percent between 2000 and 2010. Figures 1 and 2 in the appendix provide detailed information. Within the

[1] See Table 5 for a list of current and past member states of the three regional economic communities and the corresponding FTAs.

[2] Sandrey (2013) offers an overview on remaining tariffs. Angola, Democratic Republic of the Congo and Seychelles refused to participate in the FTA.

last decade, the increase in intra-Tripartite trade was mainly driven by manufactured goods trade, which holds the potential for further product development and the creation and strengthening of regional value chains. Thus, the need to quantify the effects of regional trading blocs' formation and the reduction of tariff barriers on Sub-Saharan intra-regional trade naturally arises.

Regional trade agreements (RTAs) have received a lot of attention in the literature during the last decade. Significantly positive effects of certain trade agreements are revealed for example in Baier & Bergstrand (2007), Baier, Bergstrand, Egger & McLaughlin (2008), Egger, Larch, Staub & Winkelmann (2011) and Egger (2004a), and Santos Silva & Tenreyro (2006). This literature focuses on North-North or North-South trade. Empirical findings on different African RTAs vary from a skeptical view (Yang & Gupta (2005), Longo & Sekkat (2004), and Kirkpatrick & Watanabe (2005)) to a rather optimistic one (Korinek & Melatos (2009), Musila (2005), and Cernat (2001)). Recently Afesorgbor & van Bergeijk (2014) analyze multi-membership in the ECOWAS and the SADC from 1980 to 2006 and show that competing membership hampers the effectiveness of trade agreements. They argue that the Tripartite FTA would resolve this. Moreover, Afesorgbor (2013) conducts a meta-analysis in order to summarize empirical studies on African regional integration and concludes that there might be a general upward estimation bias of RTA effects in standard panel regressions. In addition, the author compares the trade effects of the five major recognized regional economic communities, among them the COMESA and the SADC, and finds a significantly positive effect on trade within SADC countries using different estimators. The results on the role of a COMESA membership are inconclusive.

Previous studies on African trade rarely include tariff measures, mainly due to the poor data availability. We address this shortcoming and use a simple proxy for general market barriers that a country faces and empirically show that this average tariff rate still maintains important information about existing barriers in regional bilateral trade. Furthermore, we address multi-membership with a second proxy for bilateral tariffs. We apply the Sub-Saharan average tariff rate that a country has to pay on its imports from the world, whenever the two trading partners are not in the same FTA. To the best of our knowledge the effect of FTA membership has not been examined yet, although potential gains from integration via tariff eliminations or reductions are obtained within an FTA and not necessarily by the official membership status in an RTA. We will close this gap. In addition, the panel dimension is still largely unexplored. Afesorgbor & van Bergeijk (2014) and Afesorgbor (2013) are two exceptions, however, in contrast to their analyses and to most studies on this region, we control for additional variables that may determine trade flows, such as a measure of the level of education and an indicator for corruption in the trading economies. We analyze a panel of 24 out of 26 member states of the Tripartite on 16 years of bilateral export data precisely reported in 1000 US-Dollar from the UN COMTRADE Statistics. This allows us to consider cross-country data as well as the time dimension. To observe the effects of regional integration over a long horizon, the panel consists of annual data from 1995 to 2010. This includes the important period of tariff liberalization in east and south Africa at the end of the 1990s and the beginning of the 2000s.

Our analysis of intra-bloc Tripartite trade is based on a sound estimation strategy including different estimators and specification tests. This is due to a high sensitivity of results with respect to the underlying estimation method in the literature. We deal with heterogeneity and zero trade flows. The latter especially occurs when considering bilateral trade between small economies. We apply two different strategies to account for this. First, we replace cases of zero exports by small values within the panel framework and conduct a baseline two-way error component fixed effects estimation as well as a Hausman-Taylor (HT) estimation. The HT estimation does not wipe out time-invariant regressors but allows for correlation between the country-pair specific effects with the time-varying and time-invariant explanatory variables. We obtain the striking result that the Sub-Saharan average tariff rate in combination with non-membership in an FTA has a significantly negative effect on trade, and the COMESA FTA accelerates bilateral exports. Second, we apply the Poisson pseudo-maximum likelihood (PPML) estimator proposed in Santos Silva & Tenreyro (2006) as well as the corresponding panel estimator. The PPML estimates do not support these findings from the loglinear models. This is in line with the empirical results of Afesorgbor (2013) and the theoretical result of Santos Silva & Tenreyro (2006) on biased least squares estimates on the log-log notation in case of heteroskedasticity. In contrast, the significance and direction of the impact of limited market access are robust to different estimation techniques. We detect a negative effect of remaining market barriers on intra-regional bilateral trade.

The drawback of the broad time span is the lack of non-tariff barriers and transportation cost data, which are important determinants of intra-African trade. We control for these measures and for multilateral resistance with exporter- and importer-specific effects. Time effects control for trends and exogenous disturbances in intra-regional trade pattern, and events such as epidemics, crises and natural disaster.

The paper is organized as follows. Section 2 introduces the model setup and estimation techniques. Section 3 presents the data. The main regression results are discussed in section 4. In section 5 we interpret our estimation results along with supplementary findings to prove their robustness, and we compare our results to those in the literature. Finally, section 6 concludes.

2 The modeling framework and estimation procedure

The baseline gravity model was introduced to international economics by Tinbergen (1962) and motivated theoretically by Anderson & Van Wincoop (2003). Essentially, the model describes trade flows between two entities as the product of their economic sizes divided by the distance between them. Intuitively, trade between two regions i and j increases with their trade potentials reflected in national incomes, and decreases with transportation costs approximated by the distance between economic centers of the two entities. Transportation and transaction costs are additionally proxied by tariffs, common language and common border indicator variables, etc. In our setup we also account for country-pair and exporter- and importer-specific unobserved effects such as tradition or preferences that may enhance or impede trade integration. The country-specific effects also account for multilateral resistance described in Anderson & Van Wincoop (2003). A common time trend in increasing South-South trade or specific events such as global shocks are reflected in time effects. We set up our regression models on the baseline gravity equation. Along with the traditional variables we include additional regressors which may influence the countries' competitiveness and their value of exports. We use different regression models in which exports are modeled as a function of the following variables.

$$\begin{aligned}\mathrm{EX}_{ijt} &= f(\ln \mathrm{GDP}_{it},\ \ln \mathrm{GDP}_{jt},\ \ln \mathrm{DIST}_{ij},\ \mathrm{CB}_{ij},\ \mathrm{CL}_{ij},\ \mathrm{IncomeDiff}_{ijt},\ \mathrm{Tariffterm}_{ijt}, \\ &\quad \ln \mathrm{Urban}_{it},\ \ln \mathrm{School}_{it},\ \ln \mathrm{School}_{jt},\ \ln \mathrm{ExCo}_{it},\ \ln \mathrm{ImCo}_{jt},\ u_{ijt},\ \theta) \qquad (2.1)\end{aligned}$$

EX_{ijt} denotes real exports from country i to country j, u_{ijt} denotes an error term, and θ a vector of parameters, $\ln$ denotes the natural logarithm. We include real GDP data for the importing and exporting country as a measure for income, $\ln \mathrm{GDP}_{it}$ and $\ln \mathrm{GDP}_{jt}$. Bilateral transport costs and formal barriers to trade are proxied by the distance between capitals, $\ln \mathrm{Dist}_{ij}$. CB_{ij} is an indicator variable for a common border between two economies, and we add another dummy being one whenever the two countries share a common official language, CL_{ij}. Both indicate reduced trade costs. As argued in Hanson (2012), middle-income economies often specialize in manufacturing exports, while low-income countries' exports concentrate in the sectors agriculture, raw-materials and apparel. Even though most countries of the Tripartite are low-income countries, the members are highly heterogeneous in terms of GDP and the level of development. Thus, we include the absolute value of the difference in exporter's and importer's logarithm of real GDP per capita to quantify disparities in the level of development, denoted by $\mathrm{IncomeDiff}_{ijt}$. The share of urban population in the exporting country, $\ln \mathrm{Urban}_{it}$, contains information on the production structure. An economy with high urban population shares may have a higher share of manufacturing products whereas a low degree of urbanization is commonly observed in agricultural economies. Thus, urbanization tendencies may indicate structural transformation and a more diversified product space. The schooling attainment terms $\ln \mathrm{School}_{it}$ and $\ln \mathrm{School}_{jt}$ are used to measure factor endowments and technological development. We argue that intra-African trade depends on a country's competitiveness and its factor endowment as the growth in intra-Tripartite trade is mainly based on manufactured goods. This is similar to the approach of Egger (2004a), who examines the export patterns for developed countries and includes the difference in high-skilled to low-skilled labor ratios between trade partners. We add corruption indices, $\ln \mathrm{ExCo}_{it}$ and $\ln \mathrm{ImCo}_{jt}$, to complete the list of regressors. They control for distinct effects not observed in the common aggregates. We hope to proxy transaction costs and political factors that may influence trade flows. The effect of corruption on trade is not clear-cut, it can either enhance trade due to the possibility of blackmailing customs officials, or lower exports due to higher uncertainty and costs (Dutt & Traca (2010)). The latter can be interpreted as a non-tariff barrier to trade. The underlying data behind the explanatory variables introduced here, including data sources and transformations, are described to a greater extent in the data section 3.

Most importantly, we want to quantify the effect of tariff barriers on real exports. Since bilateral tariff rates are not available for a sufficient number of countries and years, we set up the following two specifications for the tariff term. In the first approach we introduce indicator variables for the three FTAs being equal to one whenever two trading partners are members of the same FTA in a certain year and zero otherwise. In what follows we denote this model by I in which the tariff term in Eq. (2.1) is explicitly defined as a vector of the following variables

$$\mathrm{Tariffterm\ I}_{ijt} \equiv (\mathrm{COMESAFTA}_{ijt},\ \mathrm{EACFTA}_{ijt},\ \mathrm{SADCFTA}_{ijt},\ \ln \mathrm{Marketbarrier}_{it}) \qquad (2.2)$$

The variable $\ln \mathrm{Marketbarrier}_{it}$ measures the average overall bilateral tariff rates each exporting country has to pay on its exports to the world. Hence, this variable captures an exporter-specific effect and thereby represents a barrier to the world market, independent of the bilateral trading partner.

In the second approach we control for the country-pair specific elimination of tariffs whenever the trading countries are in the same FTA to detect the impact of remaining tariffs.[3] For this purpose we generate an auxiliary dummy variable, AUX_{ijt}, which takes on the value one if the two trading partners are at least in one common FTA. The interaction term between this indicator and a tariff rate allows us to measure the impact of the effectively applied tariffs,
$\text{appliedtariff}_{ijt} = (1 - \text{AUX}_{ijt}) \ln \text{Tariff}_t$, country i has to pay on its exports whenever the two countries are not a member of the same trade agreement. Here $\ln \text{Tariff}_t$ denotes the logarithm of the Sub-Saharan simple average of the effectively applied tariff rate on total imports from the world. Due to a lack of data for bilateral applied tariff rates (46 percent missing in our sample), we use the average rate to approximate the individual tariff rate country j is applying on its imports from country i. World market barriers are again measured by $\ln \text{Marketbarrier}_{it}$. We denote this setup by II with the tariff term

$$\text{Tariffterm II}_{ijt} \equiv (\text{appliedtariff}_{ijt},\ \ln \text{Marketbarrier}_{it}) \tag{2.3}$$

In a first step models I and II are specified as linear panel regression equations. We regress the natural logarithm of real exports from country i to country j, $\ln \text{EX}_{ijt}$, on the set of explanatory variables in Eq. (2.1). We set up a two-way error component model that allows for country-pair specific effects, which are different for each direction of trade, i.e. $\mu_{ij} \neq \mu_{ji}$. Additional time effects are denoted by $\lambda_t,\ t = 1, \ldots, T$. Thus, u_{ijt} is written as $u_{ijt} = \mu_{ij} + \lambda_t + \varepsilon_{ijt}$. The remainder denotes the i.i.d. error term, $\varepsilon_{ijt} \sim (0, \sigma^2_{\varepsilon})$. Since our research question demands for a predetermined set of countries, the country-pair effects are most likely correlated with the regressors. That is why we use the fixed effects (FE) model and apply the within estimator to obtain consistent coefficient estimates for the time-varying regressors X_{ijt}. This choice is statistically tested by the Hausman specification test that compares the consistent estimator (FE model, $\tilde{\theta}_{Within}$) with the efficient estimator, that is only consistent under the null hypothesis that $\text{E}(\mu_{ij}|X_{ijt}) = 0$ (random effects model, $\hat{\theta}_{GLS}$). The test statistic is given by $\hat{q}_1 = \hat{\theta}_{GLS} - \tilde{\theta}_{Within}$ with H_0 : $\text{plim}\ \hat{q}_1 = 0$.

Hausman-Taylor estimation

The within estimator eliminates all time-invariant variables from our regressors. Moreover, there may be unobserved heterogeneity such that the explanatory variables are correlated with the country-pair specific effects. The distance variable in Eq. (2.1) is presumably highly correlated with the country-pair specific effect. In addition, some determinants of the tariff terms and the difference in real GDP per capita may also have country-pair specific features. Taken together, this motivates the use of the procedure proposed in Hausman & Taylor (1981). We include instruments to account for potential endogeneity and re-estimate models I and II. To do so we group the regressors as follows. Let $X1_{ijt}$ and $X2_{ijt}$ refer to the exogenous and endogenous time-varying

[3]The complete elimination of tariffs when being in an FTA poses a simplifying assumption that holds true for the EAC FTA, but for the COMESA FTA and the SADC FTA still many exceptions exist.

variables, respectively. $Z1_{ij}$ denote exogenous time-invariant variables. The HT model is specified as

$$\ln \mathrm{EX}_{ijt} = \alpha_1 X1_{ijt} + \alpha_2 X2_{ijt} + \gamma_1 Z1_{ij} + \gamma_2 Z2_{ij} + \mu_{ij} + \mu_i + \mu_j + \lambda_t + \varepsilon_{ijt} \tag{2.4}$$

We add country-specific effects μ_i and μ_j to address multilateral resistance. The model is estimated using a two-stage least squares (2SLS) procedure in which $X2_{ijt}$ is instrumented by the deviation from the individual mean, $X2_{ijt} - \overline{X2}_{ij\cdot}$, and $Z2_{ij}$ is instrumented by the individual mean $\overline{X1}_{ij\cdot}$ (see Breusch, Mizon & Schmidt (1989)). The HT estimation of Eq. (2.4) is consistent and efficient. In our case of heteroskedasticity-robust standard errors the Sargan test is used to test for exogeneity of the instruments. The null hypothesis is $\text{plim}(1/M) \sum_{m=1}^{M} \overline{X1}'_{m\cdot} \mu_m = 0$, where μ_m denotes the country-pair effects with $m = ij$. The test statistic follows a χ^2 distribution with $(p - K)$ degrees of freedom in which p and K denote the number of instruments and regressors, respectively.

Poisson estimation

Since we use the log-log specification in the FE and HT setups, we have to get rid of zero bilateral exports as regularly observed in large datasets. Before taking the natural logarithm we replace zero export data ad hoc by a small number, i.e. 0.001. Since the data are given in 1000 USD, this reflects real exports of 1 USD a year, which is a negligible value. More importantly, zero trade data may be indeed missing values or rounded down small values. These two measurement errors are frequently observed for distant entities and small countries. Thus, the errors in the multiplicative form of the gravity equation are presumably correlated with the regressors. This leads to inconsistent estimates. Related to this problem, Santos Silva & Tenreyro (2006) demonstrate that in general, due to Jensen's inequality (in our case $\ln \mathrm{E}[\mathrm{EX}_{ijt}] \neq \mathrm{E}[\ln \mathrm{EX}_{ijt}]$), in any case of heteroskedasticity the loglinearization of the gravity equation yields a built in bias of the OLS estimates of the elasticities in a loglinear regression compared to the true parameters in the multiplicative form. Santos Silva & Tenreyro (2006) show that the inclusion of country-specific fixed effects reduces the severity of the bias, but does not completely remove it. To address this shortcoming, we estimate the gravity equation in a multiplicative form with their proposed Poisson pseudo-maximum likelihood (PPML) estimator. The conditional expectation of real bilateral exports has the following form

$$\mathrm{E}[\mathrm{EX}_{ijt} | W_{ijt}, \theta, \mu_i, \mu_j, \gamma_t] = \exp(\theta W_{ijt} + \mu_i + \mu_j + \gamma_t) \tag{2.5}$$

Note, W_{ij} denotes the matrix of logarithmized and indicator explanatory variables, and the dependent variable is real exports instead of its logarithm. The coefficients can still be interpreted in terms of elasticities.[4]

The estimator offers an elegant way to cope with zero trade data for several reasons. It is consistent under heteroskedasticity, and simulation studies in Santos Silva & Tenreyro (2006) show that in contrast to other estimators the bias due to rounding-down errors in the dependent variable is almost negligible. Trade flows do not need to follow a Poisson distribution (shown in Gourieroux,

[4]The estimation is implemented in the Stata module `ppml`. The authors offer supplementary material and recent findings on a webpage: $http : //privatewww.essex.ac.uk/\ jmcss/LGW.html.$

Monfort & Trognon (1984)). The PPML estimator requires that the conditional expectation of the dependent variable is proportional to its conditional variance. This ensures that, if a realization for $\exp(\theta W_{ijt})$ is high, the variance is considered to be relatively high, too, and thereby the corresponding observation weighs as much as any other observation with small realizations. Since this proportionality is not always reflected in the data, nonlinear estimators that assume constant conditional variances through all entities have been proposed in case of overdispersion. Among others, Burger, Linders & van Oort (2009) suggest the use of the negative-binomial zero-inflated model in case of many zeros. However, their estimated elasticities depend on the scale of the dependent variable, as does the overdispersion parameter. Moreover, Santos Silva & Tenreyro (2011) show in simulation studies that zero-inflated data are not a problem for the PPML estimator at all. Aside from this, we also do not face an incidental parameter problem when including importer- and exporter-specific effects as each country acts as an importing and an exporting nation. For example, if we consider 10 countries, we observe 90 country pairs, but only 20 country-specific effects. Thus, we do not estimate specific effects for each unit of observation (see also Egger et al. (2011)). Moreover, Fernández-Val & Weidner (2013) show that there is no incidental parameter problem in Poisson regressions with two fixed effects as long as the regressors are strictly exogenous.

We also consider the Poisson panel regression model first suggested by Hausman, Hall & Griliches (1984) and estimate a simple fixed effects model allowing for country-pair and time effects. We use robust standard errors proposed by Wooldridge (1999).[5]

In the estimation we proceed as follows. The different estimators are applied to models I and II in the baseline specification with GDP data, distance, and common border and language dummies, and the tariff terms. We successively increase the number of regressors. The results are robust to this inclusion and to the order of including them. The dataset and the results of our preferred models are outlined below.

3 The dataset

The panel consists of annual bilateral export data of 24 member countries of the COMESA, the EAC and the SADC from 1995 to 2010. We exclude Seychelles and Libya from the country list due to missing data points for almost every year. We arrive at a dataset with about 28 percent missing export data remaining. Nominal bilateral exports in 1000 USD are obtained from the UN COMTRADE Statistics database. There, zero trade flows and missing values are both treated as 'not reported'. We scale the data by the export value index provided by the World Bank World Development Indicators (WDI) in order to have real export figures. The data quality on bilateral exports in Sub-Saharan Africa is generally quite low. For instance, sometimes shipments to the African continent are misleadingly counted as net exports of the port of arrival nation. We consider any kind of measurement errors by controlling for outliers. We detect implausibly high exports from Djibouti to South Africa in the years 1995-1997. We report them as 'not reported'. Since we cannot differentiate between zero trade and missing values, we compare the data to those reported

[5]The FE Poisson quasi-ML estimator is implemented in the Stata module `xtpqml` by Simcoe (2007). Cameron & Trivedi (2005) outline why there is no incidental parameter problem in Poisson panel regressions including country-pair fixed effects.

in the IMF Directions of Trade Statistics. Most of the 'not reported' entities are aligned to zeros in this database. Thus, we set each 'not reported' value to zero in our sample.

The GDP data in constant prices and real exchange rates (base year: 2005), and population data are taken from the UNCTAD Statistics database. Distance between capitals in kilometers, and data on whether or not countries share a common border and a common language are obtained from the CEPII database.

Human capital is represented by data from the Barro & Lee (2013) schooling dataset. We use the sum of the shares of population over the age of 14 who attend primary, secondary or tertiary schools and account for double-counting. Since the data are only observed at a 5-year basis we linearly interpolate them. We have to cope with missing data for six countries: Angola, Djibouti, Eritrea, Ethiopia, Comoros and Madagascar. These data points are replaced by the averages from similar countries with respect to size and trends in terms of real GDP per capita and then compared with the relative size and trends in the Cline Center schooling data (Nardulli, Peyton & Bajjalieh (2010)). The share of urban population is taken from the World Bank's World Development Indicators, the countries' ranks in the control of corruption is obtained from the World Bank's Worldwide Governance Indicators. We use the percentile rank among all countries (ranges from 0 (lowest) to 100 (highest) rank). A higher rank in index means less corruption.

We include the simple average bilateral tariff rate each exporting country has to pay on its exports to all countries in the world. This measures the limited market access due to formal trade barriers. We additionally make use of the Sub-Saharan simple average of the effectively applied tariff rates on total imports from the world whenever two trading partners are not in a common FTA. Data on tariffs are obtained from the World Integrated Trade Solution (WITS) UNCTAD Trade Analysis Information System (TRAINS) database. An overview of variables, data transformations and sources is offered in Table 6 of the appendix.

4 Results

The fixed effects results are presented in Table 1. The coefficient estimates for the baseline gravity models I and II are given in columns (1) and (2). The baseline model is extended by the degree of urbanization, the schooling attainment rates, and exporter and importer corruption. The results are presented in columns (3) and (4). The baseline results are robust to these changes. As expected the coefficients for exporter's and importer's real GDP are significantly positive, and the latter is close to unity in both models. Concerning the tariffs we find that a restricted market access is reflected in the significantly negative impact of the average export tariff on bilateral real exports. A one percent decrease in the tariff results in an increase in exports by about 0.18 to 0.19 percent. Using the dummy variable approach in model I we find a significantly positive effect on bilateral trade if both the exporting and importing country join the COMESA FTA. Considering the on average applied tariff rate on imports in model II we detect a significantly negative relationship between these costs of trading and real exports. A one percent increase of the average tariff rate the exporter has to pay, results in a decrease in exports of about 0.22 to 0.24 percent (columns (2) and (4)). Remarkably, the effect of both trading partners being in the EAC FTA becomes significantly negative in the extended model specification. As expected the within R^2 is quite low in all FE models. The Hausman test results unambiguously support our choice of modeling. The the country-pair fixed effects are highly correlated with the explanatory variables.

Table 1: Fixed effects panel regression results

	(1)	(2)	(3)	(4)
	Baseline model		Extended model	
	I	II	I	II
ln Exporter real GDP	1.6409***	1.5673***	1.1546**	1.0900**
	(0.4690)	(0.4621)	(0.4925)	(0.4863)
ln Importer real GDP	0.9549*	0.8977*	0.9713*	0.9165*
	(0.5190)	(0.5140)	(0.5367)	(0.5328)
ln Diff in per capita GDP	0.1810	0.2097	0.2123	0.2446
	(0.3852)	(0.3854)	(0.3784)	(0.3786)
COMESA FTA	0.8497***		0.7722**	
	(0.2572)		(0.2534)	
EAC FTA	−0.2818		−0.3924*	
	(0.2088)		(0.2067)	
SADC FTA	0.0825		0.0867	
	(0.2473)		(0.2517)	
ln Market barrier	−0.1906***	−0.1922***	−0.1807***	−0.1814***
	(0.0557)	(0.0558)	(0.0511)	(0.0513)
appliedtariff		−0.2435***		−0.2169***
		(0.0806)		(0.0790)
ln Exporter urban population			1.3930	1.3311
			(1.2022)	(1.1999)
ln Exporter schooling rate			3.6568**	3.6193**
			(1.6608)	(1.6552)
ln Importer schooling rate			−0.0900	−0.7934
			(1.7325)	(1.7277)
ln Exporter corruption			−0.0143	−0.0189
			(0.0793)	(0.0791)
ln Importer corruption			−0.0833	−0.0970
			(0.0743)	(0.0742)
within R^2	0.023	0.022	0.026	0.025
overall R^2	0.307	0.314	0.300	0.302

Note: ***significant at the 1%, **5%, *10% level. Robust standard errors. Constant included. Country-pair and time effects. Number of observations: 8464.

As discussed in section 2, the FE model wipes out the effects of the time-invariant variables and a random effects analysis is arguably subject to an endogeneity bias such as described in Egger (2004b). In line with his analysis, we control for potential correlation of unobserved country-pair effects with the distance between capitals. We also estimate models including the difference in real GDP per capita in the list of endogenous variables, but results do not change much and the Sargan test is in favor of our specification presented in Table 2. The Hausman test supports both models. Table 2 presents the results of the Hausman-Taylor model with time effects, country-pair fixed effects and exporter- and importer-specific effects. The last two account for features such as specialization in a certain export good, transportation costs and multilateral resistance.

The first two columns of Table 2 contain the baseline specification. The distance coefficient is significantly negative in model specification I, and a common border between two countries as well as a common language have the expected significantly positive influence on real exports. The significant coefficients for the COMESA FTA and the tariff terms are similar to the FE results in direction and size. The Sargan test of overidentification provides evidence that model I and II are well specified and the choice of instruments is appropriate (p-value of the Sargan test about 0.289 and 0.633, respectively).

Next we extend our regression models by the share of urban population, schooling attainment and exporter and importer control of corruption indices as additional exogenous regressors. The results for the extended HT setup are presented in Table 2, columns (3) and (4). The highest level of schooling attained in the exporting country has a positive and surprisingly large effect on exports. The coefficients of the other additional regressors are not significant. Compared to the baseline models, the estimates on exporter's real GDP are very close to unity. This fits well to the theoretical results in Anderson & Van Wincoop (2003) and is in line with estimates for OECD countries. Moreover, there is a slight decrease of the limited market access coefficient, implying a smaller impact of tariff rates on exports. Similarly, the absolute value of the COMESA FTA coefficient decreases slightly, but remains significant at a 1% level. In each model, the Sargan test supports our choice of instruments (p-values of 0.522 and 0.835). The exporter- and importer-specific effects are individually and jointly significant. The centered R^2 from the 2nd stage IV regression is 0.20 in model I and 0.18 in II.

Comparing the results from the HT specifications to those obtained from the simple fixed effects regression we find that in general coefficient estimates do not vary substantially. Most importantly, the coefficients' direction and size on the market access and the applied tariffs are robust, regardless of the underlying estimation method.

As discussed above, our results may yet be driven by the fact that we have 28 percent 'not reported' observations on exports. Replacing these by small values might be subject to criticism. In a next step, we set them to zero (as zeros are often used to code small trade values as well as missings) and estimate regressions on both models I and II in multiplicative form considering the real export data and not their logarithms. In a preliminary step, not reported here, we apply the PPML estimator disregarding the panel structure. In general the outcomes of the traditional gravity variables are in line with our previous findings, however, we do not find evidence on any significant role of tariffs and FTA membership. We account for the panel structure of our dataset within two approaches, for which the results of the extended models are presented in Table 3. First, in columns (1) and (2) we use the Poisson panel estimator with robust standard errors proposed by Wooldridge (1999) and implemented in Stata by Simcoe (2007). We include country-pair fixed effects and time dummies. Second, we use the PPML estimator, include time dummies and allow for exporter- and importer-specific effects. Clustered standard errors for the country-pairs are used to control for the panel structure. Columns (3) and (4) in Table 3 refer to this. Additional results for the baseline model and the PPML estimates without country-(pair) specific effects are available on request. A third PPML specification would additionally include country-pair specific effects but raises the issue of multicollinearity and is therefore discarded from the analysis.

The results in Table 3 are robust to the different Poisson models. Real GDP data are no longer

Table 2: Hausman-Taylor panel regression results

	(1)	(2)	(3)	(4)
	Baseline model		Extended model	
	I	II	I	II
ln Exporter real GDP	1.6186***	1.5489***	1.1255**	1.0674**
	(0.4793)	(0.4459)	(0.4903)	(0.4960)
ln Importer real GDP	0.9379*	0.8828*	0.9514*	0.9002*
	(0.5105)	(0.5061)	(0.5387)	(0.5212)
ln Distance	−1.3970*	−0.2318	−1.4117**	−0.3294
	(0.7153)	(0.9469)	(0.6916)	(0.8858)
Common border	3.0387***	4.0988***	3.0355***	4.0157***
	(0.8922)	(1.1203)	(0.8882)	(1.0748)
Common language	1.0380**	1.0747**	1.0404**	1.0759**
	(0.4123)	(0.4279)	(0.4133)	(0.4480)
ln Diff in per capita GDP	0.0658	0.0855	0.0766	0.0960
	(0.1791)	(0.1827)	(0.1793)	(0.1859)
COMESA FTA	0.8015***		0.7268***	
	(0.2424)		(0.2478)	
EAC FTA	−0.3380		−0.4484*	
	(0.2239)		(0.2115)	
SADC FTA	0.1042		0.1095	
	(0.2504)		(0.2546)	
ln Market barrier	−0.1911***	−0.1915***	−0.1810***	−0.1806***
	(0.0561)	(0.0534)	(0.0513)	(0.0517)
appliedtariff		−0.2424***		−0.2162***
		(0.0805)		(0.0793)
ln Exporter urban population			1.4096	1.3335
			(1.2112)	(1.1746)
ln Exporter schooling rate			3.6880**	3.6271**
			(1.7278)	(1.7337)
ln Importer schooling rate			−0.6697	−0.7987
			(1.7637)	(1.7675)
ln Exporter corruption			−0.0123	−0.0160
			(0.0792)	(0.0809)
ln Importer corruption			−0.0885	−0.0946
			(0.0711)	(0.0778)
Sargan test (p-value)	0.2886	0.6328	0.5216	0.8353

Note: ***significant at the 1%, **5%, *10% level. Bootstrapped standard errors with 1000 replications are used. Constant, time dummy variables, country-pair and exporter- and importer-specific effects included. Endogenous variable: ln Distance. Number of observations: 8464.

significant determinants of real exports. The within estimator deletes the time-invariant variables in columns (1) and (2), but in the PPML setup the distance and common border coefficients are significant at a 1% level and are about -1.15 and 0.95, respectively. The common language dummy is insignificant in all model specifications that include exporter- and importer-specific fixed effects and thus does not determine the size of real exports. Regarding our main research question we find that the effect of the average tariff rate a country has to pay for exports to the world (ln Market barrier) is much less than suggested by the log-log gravity equations in the HT model in Table 2 (range: -0.192 to -0.181). Now, this limited market access has a small significantly negative effect on exports, the coefficients range from -0.047 to -0.042. Aside from this, being in one or more FTAs does not matter at all. The same holds true for the coefficients of the COMESA and EAC FTAs. Comparing the impact of both tariff measures in the PPML framework to the one in the HT model we find an upward bias in the panel setup. More generally, the estimates in the HT and FE models in the log-log notation seem to be upward biased. This is in line with the theory discussed in Santos Silva & Tenreyro (2006) and their findings from an empirical study on trade flows.

In summary, our preferred specifications of the Poisson models I and II in columns (3) and (4) suggest that real exports increase if the trading countries share a common border, and with higher exporter schooling attainment and lower corruption in the importing country. Exports significantly decrease with distance, limited market access and the level of education in the importing economy. The latter may reflect the negative relation between the level of education and import demand for more sophisticated products (for details see section 5). The R^2 for the two extended PPML models are computed as the square of the correlation between the export data and their fitted values and take on the values 0.869 and 0.868, respectively.

5 Discussion and robustness checks

In what follows we discuss our findings and draw a comparison to those in the literature, especially with respect to the role of trade agreements and tariff rates. Furthermore, we conduct supplementary robustness checks and give detailed interpretations to our main estimation results.

FTA membership status

We focus on the effectiveness of FTAs in increasing real exports in the COMESA-EAC-SADC Tripartite countries. Since we look at ratified FTAs and not RTAs, our results are comparable to the literature only to a certain extent.

While we estimate a positive effect only for the COMESA free trade agreement membership in the panel framework, the effect is not significant using Poisson regressions. This is in line with the comparative analysis from Korinek & Melatos (2009) on a panel from 1981 to 2005. They show that the existence of an RTA promotes intra-regional trade in agricultural products. It is noteworthy that the effect for the COMESA is the smallest among the considered agreements, and that it is only relevant in the FE estimation. They cannot replicate this finding using PPML. This is supported by Afesorgbor (2013) for aggregated exports within a similar sample period. The COMESA RTA

Table 3: Poisson pseudo-maximum likelihood regression results

	(1)	(2)	(3)	(4)
	Poisson panel model		PPML model	
	I	II	I	II
ln Exporter real GDP	−0.2464	−0.2198	−0.4158	−0.3480
	(0.5140)	(0.5136)	(0.5708)	(0.5496)
ln Importer real GDP	0.4021	0.4248	0.2519	0.2892
	(0.2836)	(0.2809)	(0.1822)	(0.1762)
ln Distance			−1.1596***	−1.1494***
			(0.1988)	(0.1978)
Common border			0.9499***	0.9576***
			(0.2686)	(0.2641)
Common language			0.2610	0.2525
			(0.3644)	(0.3620)
ln Diff in per capita GDP	0.1268	0.1267	0.0448	0.0558
	(0.3235)	(0.3187)	(0.0993)	(0.0988)
COMESA FTA	0.2064		−0.1520	
	(0.1354)		(0.2692)	
EAC FTA	−0.1265		0.0346	
	(0.1303)		(0.2463)	
SADC FTA	−0.0169		−0.0281	
	(0.1379)		(0.1751)	
ln Market barrier	−0.0466**	−0.0420*	−0.0469**	−0.0472**
	(0.0226)	(0.0218)	(0.0215)	(0.0209)
appliedtariff		−0.0365		−0.0020
		(0.0441)		(0.0561)
ln Exporter urban population	1.4757	1.2409	1.1507	1.1995
	(1.0468)	(0.9709)	(1.0951)	(1.0371)
ln Exporter schooling rate	1.5962***	1.5095**	1.8933***	1.7014**
	(0.6145)	(0.6503)	(0.7279)	(0.7193)
ln Importer schooling rate	−1.8519***	−2.0009***	−1.6716***	−1.7510***
	(0.5771)	(0.6195)	(0.5878)	(0.6379)
ln Exporter corruption	0.0223	0.0171	0.0189	0.0219
	(0.0822)	(0.0828)	(0.0852)	(0.0850)
ln Importer corruption	0.0626***	0.0627***	0.0642***	0.0654***
	(0.0199)	(0.0199)	(0.0218)	(0.0225)
No. observations	7728	7728	8464	8464
Country-pair effects	yes	yes	no	no
Country-specific effects	no	no	yes	yes

Note: ***significant at the 1%, **5%, *10% level. Constant and time dummies included. Robust standard errors in columns (1) and (2), country-pair clustered standard errors in columns (3) and (4).

dummy is positive and significant in most estimations, except in those using a PPML setup. Geda & Kebret (2008) analyze exports from 1980 to 2004 within a Tobit estimation and emphasize that macroeconomic variables as well as infrastructure are positively related to trade pattern within the COMESA RTA.

The membership in the EAC customs union does not indicate any export promoting effect in our study, in most specifications we obtain insignificant coefficient estimates.[6] A possible explanation for this is may be given in Buigut (2012) who shows that intra-EAC imports have largely increased for all countries, while intra-EAC exports were mainly driven by Kenya and Uganda. Busse & Shams (2005) also reveal that Kenya benefits most from intra-bloc trade. Moreover, external events that are common for all EAC nations, may cause this insignificance. EAC economies suffered more than other Tripartite nations from the coffee crisis between 2000 and 2005. In addition, two economically weakened countries joined the EAC in 2007: Burundi, that suffered from a civil war (1993-2003), and Rwanda, that was involved in the Second Congo War (1998-2003). Apart from this, the EAC economies were affected by various region-specific hunger crises.

Afesorgbor (2013) finds that if both trading countries are members of the SADC RTA, exports can be up to three times higher. The impact diminishes in the PPML estimation, but remains significant. Carrere (2004) estimates an HT model for the years 1962 to 1996 and finds a positive impact on real imports for four African regional agreements, among them the SADC. Compared to this, in our analysis the SADC FTA membership does not show any significant impact on bilateral exports. Two possible reasons for the differing outcomes are worth mentioning. First, the SADC FTA exists since 2008, whereas the RTA was founded in 1992. Given that our analysis only covers three years of the FTA, any positive effect on bilateral exports might materialize not immediately, but within the passage of time. This is in line with Coulibaly (2009). He estimates the impact of the duration of a membership in an RTA on bilateral exports within a panel of 56 exporting and 90 importing countries and 40 years (1960-1999). The analysis reveals that the effectiveness of the SADC RTA in promoting exports increases over time. Commenting on a more general issue, second, most studies use IMF Directions of Trade Statistics, which are given in million USD and which are therefore more likely to be subject to rounding errors than the UN COMTRADE data denoted in 1000 USD we consider.

In summary, the dummy variable approach literature offers mixed evidence on trade enhancing effects of RTAs. In a meta-analysis Afesorgbor (2013) integrates 14 individual empirical studies with 139 results. 40 percent of the estimated coefficients are larger than one, which they interpret as an upward bias. While 35 percent of the results are between zero and one, 25 percent predicted a negative effect on exports. Actually, the approach is subject to criticism as FTA indicator variables suggest an immediate effect of the tariff liberalization and cannot fully capture the effect of a stepwise tariff reduction as observed in east and south Africa (see again Coulibaly (2009)).

Tariff barriers

[6]We even observe a significantly negative effect of the EAC FTA dummy on bilateral exports in the HT estimation controlling for the influence of education, urbanization and corruption (Table 2). This may be attributed to a bias in the estimation of loglinear gravity equations.

Due to poor data records only few panel studies for the African continent include tariff rates. Iwanow & Kirkpatrick (2008) use an on average applied tariff rate on all incoming products from WITS for 124 developed and developing countries in 2003 and 2004. In line with our results, they find a significantly negative effect on bilateral trade. Hayakawa (2013) uses bilateral tariff data from WITS and interpolates them. He supports the view that the omission of bilateral tariffs does not raise an estimation bias. This is in accordance with our finding that the coefficient estimates do not vary when the proxy for limited market access is included. Noteworthy, in our analysis the market barrier tariff coefficient is significantly negative throughout the different estimation techniques. We further identify a negative effect of Sub-Saharan import tariffs on bilateral exports in the HT model II. This and the positive effect of the COMESA FTA membership, both become insignificant in most PPML specifications.

Our findings are in part in line with those from a general equilibrium analysis of the COMESA-EAC-SADC Tripartite FTA conducted by Willenbockel (2013). Within a trade policy simulation for eight different scenarios of the Tripartite FTA on aggregated and country-specific levels they show that all countries benefit in terms of welfare gains and total exports under different scenarios. The author considers the elimination of remaining intra-COMESA and intra-SADC tariffs, and a complete elimination of all intra-Tripartite tariffs in combination with a reduction of non-tariff barriers.

Traditional variables

Signs and values of the traditional gravity variables estimated in our study are in line with the literature on intra-African trade, with one exception. The effect of GDP becomes insignificant in the PPML estimation once we include either country-specific or country-pair specific effects. Previous studies (e.g. Afesorgbor (2013)) do not control for exporter- and importer-specific effects and find a positive effect of the countries' real GDP. The inclusion of unobservable effects has received marginal attention so far. Herrera (2012) points out this specification problem and suggests the exclusion of the GDP variables when including time-varying country-specific effects. Since we only control for time-invariant country-specific characteristics, we still include GDP measures.

Aside from this, we conduct two robustness checks addressing the set of traditional variables. First, our results are robust to using real GDP per capita data instead of total real GDP. Second, we resort to the literature on trade gravity equations and include a landlocked dummy into our models. Such an inclusion has proven valuable whenever a huge proportion of products are shipped, however, we question its relevance looking only at intra-regional trade in Sub-Saharan Africa. This is reflected in insignificant coefficient estimates.

Throughout most model specifications and estimation techniques the distance between capitals has a large and highly significant negative effect on bilateral exports. This highlights the importance of transportation costs and infrastructure on intra-Tripartite trade.

Education

As reported in section 4, schooling attainment in the importing country and real exports seem to have an inverse relationship in the PPML setup when we allow for country-specific or country-pair specific effects. Several distinct mechanisms may be at work here. A higher level of education in the importing country may result in an increase in import demand for sophisticated goods that are imported from beyond Sub-Saharan Africa. Additionally, differences in factor endowment may determine specialization and drive up foreign (Sub-Saharan) supply of skill-intensive goods to a low-skill importing country.

We are aware of the fact that we completely approximate missing schooling attainment data for six countries as otherwise we would face a great reduction in the number of observations. This may be controversial as we use the average attainment rates over countries with similar GDP per capita profiles. To make sure that this finding above is not just a statistical artifact we re-estimate the models shown in Table 3 removing the corresponding countries from the set of observations. The results are presented in Table 4. Not surprisingly we find some changes in the significance of the real GDP coefficient estimates. Most remaining coefficient estimates are in line with those presented above. The distance and common border coefficients are slightly smaller in absolute values. The influence of schooling attainment and corruption is almost unaltered. Most importantly, market barrier coefficients are significant and range between -0.051 and -0.047, which is well in line with the results above. The FTAs still do not have any significant impact on exports within the Poisson models. This remarkable finding is also confirmed for the HT model, for which the same robustness check is conducted (tabulated results are not shown here to conserve space). In this model, the coefficients for the market barrier term are significant at a 1% level and range between -0.246 and -0.216.

Corruption

Furthermore, we find a significantly positive, yet small role of low corruption in the importing country (0.09 in the HT model and 0.06 in the PPML specifications). The positive sign indicates that the better corruption is curtailed, the higher the bilateral exports. Recently Dutt & Traca (2010) also report such a trade taxing or extortion effect of corruption. In contrast, the authors also identify an export enhancing effect of corruption. One may think of exporting firms who bribe the customs officials of the importing country to bypass trade barriers, such as high tariffs. Interestingly, we do not confirm this for our sample of Sub-Saharan Africa. Aside from this, for a discussion on the usefulness of this index see e.g. Knoll & Zloczysti (2012).

Non-tariff barriers

It goes without saying that our analysis is limited to observable data. The drawback of the long time span used in our analysis is the lack of non-tariff barriers (NTBs) and transportation cost data. Non-tariff barriers that lower exports are also difficult to measure, however, they are recognized to be important impediments to trade. In order to give a first indication of how severe such barriers are, we conduct a subsample analysis and use the World Bank's Doing Business Indicators for the years 2006 to 2010. We seek to capture the effect of NTBs on bilateral trade by including the number of documents to be filled out, the deflated costs to export or import per container, and the

Table 4: Robustness check schooling attainment: Poisson regression results

	(1)	(2)	(3)	(4)
	Poisson panel model		PPML model	
	I	II	I	II
ln Exporter real GDP	-0.6927^{*}	-0.6842^{*}	-0.9720^{**}	-0.8360^{**}
	(0.3993)	(0.3896)	(0.3970)	(0.3994)
ln Importer real GDP	0.5398^{*}	0.5489^{*}	0.2533	0.3100
	(0.2881)	(0.2845)	(0.2000)	(0.1951)
ln Distance			-0.9818^{***}	-0.9627^{***}
			(0.1633)	(0.1648)
Common border			0.7916^{***}	0.8261^{***}
			(0.1798)	(0.1841)
Common language			$0.5594^{(*)}$	0.5520^{*}
			(0.3400)	(0.3240)
ln Diff in per capita GDP	0.3044	0.2890	-0.0337	-0.0219
	(0.2975)	(0.2926)	(0.0931)	(0.0960)
COMESA FTA	0.1874		-0.3600	
	(0.1575)		(0.2565)	
EAC FTA	-0.0643		0.2062	
	(0.1253)		(0.2802)	
SADC FTA	-0.0295		0.1767	
	(0.1456)		(0.1437)	
ln Market barrier	-0.0511^{**}	-0.0487^{**}	-0.0471^{**}	-0.0480^{**}
	(0.0204)	(0.0201)	(0.0209)	(0.0203)
appliedtariff		-0.0291		-0.04063
		(0.0464)		(0.0517)
ln Exporter urban population	1.0604	0.9398	0.8662	0.9237
	(0.9997)	(0.9431)	(1.0994)	(0.9720)
ln Exporter schooling rate	1.6346^{**}	1.5998^{**}	1.948^{**}	1.4550^{*}
	(0.6480)	(0.6961)	(0.7850)	(0.7720)
ln Importer schooling rate	-1.7011^{***}	-1.8190^{***}	-1.7621^{***}	-1.8727^{***}
	(0.5583)	(0.5909)	(0.5845)	(0.6397)
ln Exporter corruption	0.0789	0.0755	0.0799	0.0840
	(0.0832)	(0.0813)	(0.0854)	(0.0847)
ln Importer corruption	0.0661^{***}	0.0663^{***}	0.0711^{***}	0.0728^{***}
	(0.0199)	(0.0199)	(0.0200)	(0.0220)
No. observations	4528	4528	4624	4624
Country-pair effects	yes	yes	no	no
Country-specific effects	no	no	yes	yes

Note: ***significant at the 1%, **5%, *10% level. Constant and time dummies included. Robust standard errors in columns (1) and (2), country-pair clustered standard errors in columns (3) and (4). Data for Angola, Djibouti, Eritrea, Ethiopia, Comoros, and Madagascar are excluded.

lead time to export or import in days as additional variables. The relation between these indicators and real export data may be at least twofold. At first sight, we may expect that NTBs lower trade. However, in many countries with low tariff rates, non-tariff measures such as administrative costs are said to be the last resort to protect home markets to a small extent. Thus, a positive relation to real bilateral exports would not come as a surprise. We do find limited evidence for effects of NTBs on bilateral exports within the Tripartite nations among different models. We briefly comment only on the results for the PPML estimation to conserve space. For this short sample, the common border and the distance between capitals are detected as main determinants for trade pattern between economies, and the EAC membership coefficient is significantly positive. But we also find evidence that an increase in last year's import costs (excluding tariffs) by 1 percent reduces real exports by 0.53 - 0.61 percent. Aside from this, the used indicators do not account for the majority of NTBs such as infrastructural quality or predictability of transparency of legal decisions.

The role of NTBs and trade facilitation activities have been examined for different countries and years in several studies. Karugia, Wanjiku, Nzuma, Gbegbelegbe, Macharia, Massawe, Freeman, Waithaka & Kaitibie (2009) conduct such an analysis for the EAC. They analyze trade patterns in the maize and beef sector within a spatial equilibrium model based on data from a regional survey in 2007. The authors show that the low level of intra-EAC trade of these goods can be increased by improving administrative procedures at border points, reducing the extent of road blocs and implementing an efficient monitoring system. Iwanow & Kirkpatrick (2008) examine the determinants of manufacturing exports for 124 developing countries in 2003 and 2004. They consider a trade facilitation index that consists of the number of all documents required, the time necessary to comply with all procedures to export/import goods, and the costs associated with this, as well as an infrastructure index which contains a measure of paved roads, rail density, number of telephone and mobile phone subscribers. Controlling for these variables the "African dummy" becomes insignificant. In absolute terms the infrastructure index has the largest effect, and is of special importance within the African subsample. Considering data from 2004 to 2007, Portugal-Perez & Wilson (2012) analyze 101 developing countries. They aggregate a pool of 18 variables from the WEF's Global Competitiveness Report, the Doing Business Report and Transparency International to end up with four indicators: Physical infrastructure, information and communication technology, border and transport efficiency, and business and regulatory environment. They confirm the findings of Iwanow & Kirkpatrick (2008), that physical infrastructure has a large effect on bilateral exports, a 1 percent increase of physical infrastructure results in an 0.2-0.5 percent increase of exports. Unfortunately, we cannot contribute to this new avenue of empirical research on non-tariff barriers as much of the data are not available for the Tripartite countries over our sample period.

6 Conclusion

This study analyzes the impact of regional integration efforts among the COMESA-EAC-SADC Tripartite countries on bilateral exports. Within a panel framework, we apply the Hausman-Taylor instrumental variable estimation procedure on bilateral export data between 1995 and 2010 and estimate an extended gravity model for the member states. We consider traditional explanatory variables as well as indicators for the level of education, corruption, and urbanization. We also account for multi-membership in regional trade agreements and include two tariff measures. In addition to the loglinear model, we conduct a Poisson pseudo-maximum likelihood estimation on the gravity equation in multiplicative form. This allows us to work with zero export data and to avoid the bias made in least squares estimation of the loglinear model in case of heteroskedasticity. Such a bias may induce misleading conclusion when evaluating tariff reducing policies and the achievements of free trade agreements.

Indeed, our findings provide evidence for a bias in the log-log notation. The COMESA FTA has a positive impact on real exports with coefficient estimates about 0.8 in the Hausman-Taylor regression only. Furthermore, the EAC and SADC FTAs do not show any positive effects on real exports. Moreover, the Sub-Saharan average import tariff does not significantly determine intra-regional bilateral trade in our analysis. To this extent our study confirms the pessimistic view of the effectiveness of African free trade agreements.

However, we detect a negative effect of the average tariff barrier of an exporter (which represents limited global market access) on intra-Tripartite trade, independent of the underlying estimation technique and model setup. This gives a rationale for accelerating integration efforts as formal tariff barriers still matter.

References

Afesorgbor, S. K. (2013). Revisiting the effectiveness of African economic integration. A meta-analytic review and comparative estimation methods., *Technical report*, University of Aarhus.

Afesorgbor, S. K. & van Bergeijk, P. A. (2014). Measuring multi-membership in economic integration and its trade-impact: A comparative study of ECOWAS and SADC., *South African Journal of Economics* **82**(4): 518–530.

Anderson, J. & Van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle, *The American Economic Review* **93**(1): 170–192.

Baier, S. & Bergstrand, J. (2007). Do free trade agreements actually increase members' international trade?, *Journal of International Economics* **71**: 72–95.

Baier, S. L., Bergstrand, J. H., Egger, P. & McLaughlin, P. A. (2008). Do economic integration agreements actually work? Issues in understanding the causes and consequences of the growth of regionalism, *The World Economy* **31**(4): 461–497.

Barro, R. J. & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950-2010, *Journal of Development Economics* **104**: 184–198.

Breusch, T., Mizon, G. & Schmidt, P. (1989). Efficient estimation using panel data, *Econometrica* **57**: 695–700.

Buigut, S. (2012). An assessment of the trade effects of the East African Community customs union on member countries, *International Journal of Economics and Finance* **4**(10): 41–53.

Burger, M., Linders, G.-J. & van Oort, F. (2009). On the specification of the gravity model of trade: Zeros, excess zeros and zero-inflated estimation, *Spatial Economic Analysis* **4**(2): 167–190.

Busse, M. & Shams, R. (2005). Trade effects of the East African Community, *The Estey Centre Journal of International Law and Trade Policy* **6**(1): 62–83.

Cameron, A. C. & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*, Cambridge University Press.

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

Cernat, L. (2001). Assessing regional trade arrangements: Are South–South RTAs more trade diverting?, *Policy Issues in International Trade and Commodities Study Series 16*, UNCTAD, United Nations, New York and Geneva.

COMESA (2015). Tripartite Summit Bulletin 2, 15th e-COMESA newsletter.

Coulibaly, S. (2009). Evaluating the trade effect of developing regional trade agreements: A semi-parametric approach, *Journal of Economic Integration* **24**(4): 709–743.

Dutt, P. & Traca, D. (2010). Corruption and bilateral trade flows: Extortion or evasion?, *The Review of Economics and Statistics* **92**: 843–860.

Egger, P. (2004a). Estimating regional trading bloc effects with panel data, *Review of World Economics* **140**(1): 151–166.

Egger, P. (2004b). On the problem of endogenous unobserved effects in the estimation of gravity models, *Journal of Economic Integration* **19**(1): 182–191.

Egger, P., Larch, M., Staub, K. E. & Winkelmann, R. (2011). The Trade Effects of Endogenous Preferential Trade Agreements, *American Economic Journal: Economic Policy* **3**: 113–143.

Fernández-Val, I. & Weidner, M. (2013). Individual and time effects in nonlinear panel models with large N,T, *CeMMAP working papers CWP60/13*, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Geda, A. & Kebret, H. (2008). Regional economic integration in Africa: A review of problems and prospects with a case study of COMESA, *Journal of African Economics* **17**(3): 357–394.

Gourieroux, C., Monfort, A. & Trognon, A. (1984). Pseudo maximum likelihood methods: Applications to Poisson models, *Econometrica* **52**: 701–720.

Hanson, G. (2012). The rise of middle kingdoms: Emerging economies and global trade, *Journal of Economic Perspectives* **26**: 41–64.

Hausman, J., Hall, B. H. & Griliches, Z. (1984). Econometric models for count data with an application to the patents-R&D relationship, *Econometrica* **52**(4): 909–938.

Hausman, J. & Taylor, W. (1981). Panel data and unobservable individual effects, *Econometrica* **49**(6): 1377–1398.

Hayakawa, K. (2013). How serious is the omission of bilateral tariff rates in gravity?, *Journal of the Japanese and International Economies* **27**: 81–94.

Herrera, E. (2012). Comparing alternative methods to estimate gravity models of bilateral trade, *ThE Papers 5*, University of Granada.

Iwanow, T. & Kirkpatrick, C. (2008). Trade facilitation and manufacturing exports: Is Africa different?, *World Development* **37**(6): 1039–1050.

Karugia, J., Wanjiku, J., Nzuma, J., Gbegbelegbe, S., Macharia, E., Massawe, S., Freeman, A., Waithaka, M. & Kaitibie, S. (2009). The impact of non-tariff barriers on maize and beef trade in East Africa, *Working Paper 29*, ReSAKSS.

Kirkpatrick, C. & Watanabe, M. (2005). Regional trade in Sub-Saharan Africa: An analysis of East African trade cooperation, 1970-2001, *The Manchester School* **73**(2): 141–164.

Knoll, M. & Zloczysti, P. (2012). The Good Governance Indicators of the Millennium Challenge Account: How Many Dimensions are Really Being Measured?, *World Development* **40**(5): 900–915.

Korinek, J. & Melatos, M. (2009). Trade impacts of selected regional trade agreements in agriculture, *Trade Policy Working Paper 87*, OECD.

Longo, R. & Sekkat, K. (2004). Economic obstacles to expanding intra-African trade, *World Development* **32**(8): 1309–1312.

Musila, J. W. (2005). The intensity of trade creation and trade diversion in COMESA, ECCAS and ECOWAS: A comparative analysis, *Journal of the African Economies* **14**(1): 117–141.

Nardulli, P., Peyton, B. & Bajjalieh, J. (2010). Gauging cross-national differences in education attainment, *Technical report*, Cline Center for Democracy, University of Illinois.

Othieno, L. & Shinyekwa, I. (2011). Prospects and challenges in the formation of the COMESA-EAC and SADC Tripartite Free Trade Area, *Research Series 87*, Economic Policy Research Centre.

Portugal-Perez, A. & Wilson, J. (2012). Export performance and trade facilitation reform: Hard and soft infrastructure, *World Development* **40**(7): 1295–1307.

Sandrey, R. (2013). An analysis of the SADC Free Trade Area, *Stellenbosch: tralac* .

Santos Silva, J. & Tenreyro, S. (2006). The log of gravity, *The Review of Economics and Statistics* **88**(4): 641–658.

Santos Silva, J. & Tenreyro, S. (2011). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator, *Economic Letters* **112**(2): 220–222.

Simcoe, T. (2007). XTPQML: Stata module to estimate Fixed-effects Poisson (Quasi-ML) regression with robust standard errors, Statistical Software Components, Boston College Department of Economics.

Tinbergen, J. (1962). Shaping the world economy: Suggestions for an international economic policy, *New York: Twentieth Century Fund* .

UNCTAD (2014). *Handbook of Statistics 2014*, United Nations, New York and Geneva.

Willenbockel, D. (2013). General equilibrium assessment of the COMESA-EAC-SADC Tripartite FTA, *Munich Personal RePEc Archive Paper* .

Wooldridge, J. (1999). Distribution-free estimation of some nonlinear panel data models, *Journal of Econometrics* **90**: 77–97.

Yang, Y. & Gupta, S. (2005). Regional trade agreements in Africa: Past performance and the way forward, *Working Paper Series 5*, International Monetary Fund.

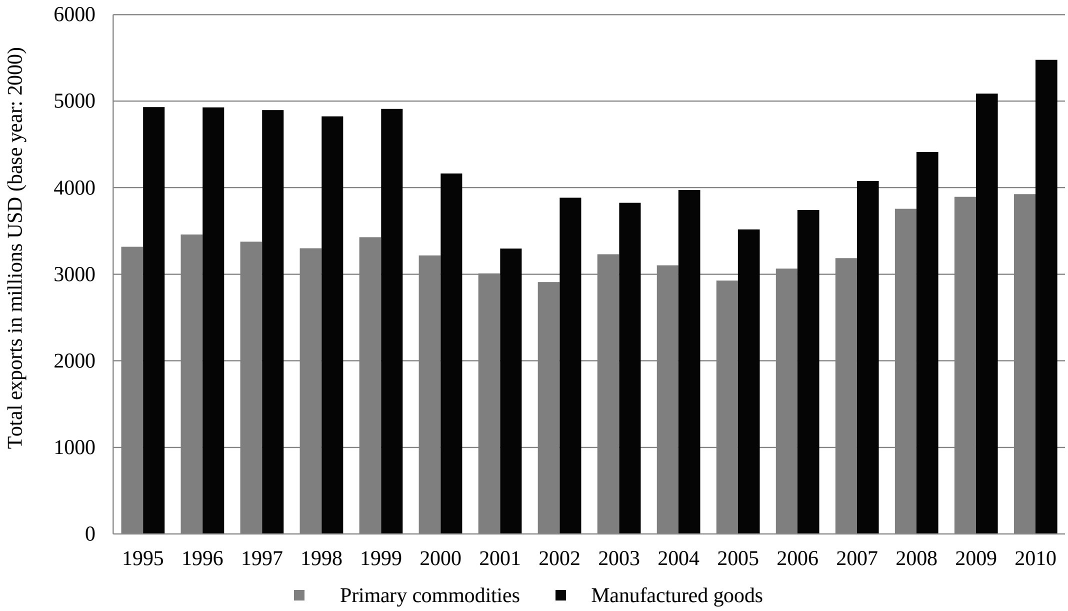

Figure 1: Total real intra-Tripartite exports by main sectors

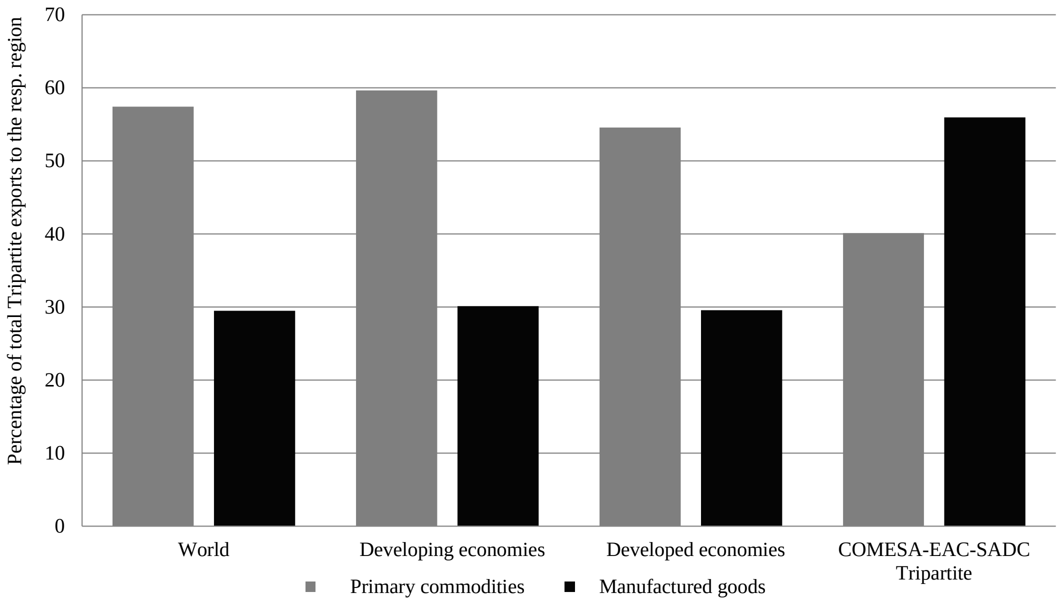

Figure 2: Real sectoral exports from the Tripartite countries to different regions in the year 2010

Table 5: List of countries and membership status

	COMESA		EAC	SADC		Tripartite
	PTA	FTA		PTA	FTA	FTA
Angola	1981-2006	-	-	1992	-	2011
Burundi	1981	2004	2007	-	-	2011
Botswana	-	-	-	1992	2008	2011
DR Congo	1981	-	-	1992	-	2011
Djibouti	1981	2000	-	-	-	2011
Egypt	1999	2000	-	-	-	2011
Eritrea	1994	-	-	-	-	-
Ethiopia	1981	-	-	-	-	2014
Kenya	1981	2000	2000	-	-	2011
Comoros	1981	2006	-	-	-	2011
Lesotho	1981-1997	-	-	1992	2008	2011
Libya	2005	2006	-	-	-	2011
Madagascar	1981	2000	-	1992-2009 reinstated 2014	-	-
Mauritius	1981	2000	-	1992	2008	2011
Malawi	1981	2000	-	1992	2008	2011
Mozambique	1981-1997	-	-	1992	2008	2011
Namibia	-	-	-	1992	2008	2011
Rwanda	1981	2004	2007	-	-	2011
South Africa	-	-	-	1992	2008	2011
Sudan	1981	2000	-	-	-	2011
Swaziland	1981	-	-	1992	2008	2011
Tanzania	1981-2000	-	2000	1992	2008	2011
Uganda	1981	-	2000	-	-	2011
Zambia	1981	2000	-	1992	2008	2011
Zimbabwe	1981	2000	-	1992	2008	2011

Note: The Table contains the years of admission to the preferential trade areas (PTAs), which correspond to the RTAs for COMESA and SADC, and the years of admission to the FTAs, or the time span of membership.
Sources: www.comesa.int; www.sadc.int; www.eac.int.

Table 6: Variable description, data transformation and source

Variable	Description and Transformation	Data Source and Availability
Real exports	Nominal bilateral export data in 1000 USD scaled by the export value index	UN COMTRADE Statistics Database (2013) World Bank World Development Indicators (WDI)
Real GDP	GDP in Millions $US, constant 2005 prices and constant exchange rate	UNCTADStat (2013) 1995 - 2010
Difference in real GDP per capita	$DIF_{ijt} = \mid ln(\frac{GDP_{it}}{POP_it} - ln(\frac{GDP_{jt}}{POP_{jt}}) \mid$ Real GDP data; Total population in thousands	UNCTADStat (2013) 1995 - 2010
Distance	Distance in kilometers between capitals	CEPII database
Common border	Indicator variable: set to 1 if there is a common border two countries	CEPII database
Common language	Indicator variable: set to 1 if two countries share an official language	CEPII database
FTAs	Dummy variable set to unity if both trading partners are member countries of the same FTA	www.comesa.int; www.eac.int; www.sadc.int
Market barrier	Average over all countries of bilateral effectively applied tariff rate on imports, simple average, all products, by exporter	World Bank WITS Trains (2014) 1995 - 2010
Applied tariff	$\text{appliedtariff}_{ijt} = (1 - \text{AUX}_{ijt}) \ln \text{Tariff}_t$ Tariff_t: Average over all Sub-Saharan countries of effectively applied tariff rate on imports from all countries, simple average, all products AUX_{ijt}: Variable set to unity if both trading partners are member countries of the same FTA	World Bank WITS Trains (2014) 1995 - 2010 (Missing data for Sudan)
Urban population	People living in urban areas as percent of total population	World Development Indicators (2014) 1995 - 2010
Schooling attainment	Sum of primary, secondary and tertiary attainment, highest level attained (percent of population aged 15 and over) Missing data for 6 countries: averages from similar countries in terms of real GDP and in terms of Cline Center schooling data Missing years: linearly interpolated justified by WDI (2014) data	Barro and Lee Database (2014) 1990, 1995, 2000, 2005, 2010 (Missing data for Angola, Djibouti, Eritrea, Ethiopia, Comoros, Madagascar)
Control of corruption	Percentile Rank 0-100 Missing years: Average between two years	World Bank Worldwide Governance Indicators (2014) 1996 - 2010 (Missing data for 1997, 1999 and 2001)
Trading across borders	Time to exports and imports (days); Documents to export and import (number); Cost to export and import (deflated US$ per container)	World Bank Doing Business Report (2014) 2006 - 2010