

Response to the referee 2

Thank you for the positive evaluation and suggestions. Your comments are essential and significantly contribute to the quality of our manuscript. Our responses are indicated in italics.

1. I find the paper very interesting and timely especially at the time when regional integration and African Continental free-trade debates have been reinvigorated in high level policy circles across the continent. The distinction of extensive and intensive margins in the analyses adds more value. The work definitely adds value to literature and policy. I think the literature review can be improved with more recent work on the subject and on Africa particularly.

***Response:** Thank you very much for the positive comments.*

2. I think the literature review can be improved with more recent work on the subject and on Africa particularly. The most current literature the authors review is 2006, which is near obsolete in my view.

***Response:** Thank you for the feedback. This will be certainly included in the revised version.*

3. In the methodology, I would like a bit of theoretical foundation prior to equation 2.

***Response:** We are grateful for the comment. We will include theoretical support for the equation in the revised version.*

4. In terms of the estimation technique, the authors prefer Heckman for obvious reasons, after arguing the weaknesses of the PPML. They base their final choice of estimation technique on the analysis of the Mills ratio, which is sound. However, I think that the actual estimation should have included the PPML, and compare results for extra robustness. I would recommend this, perhaps only for one model.

***Response:** Thank you very much for the comment. We repeated the exercise we did for Table 1 in the manuscript using PPML estimator. We find the results reported in the Table 1 below. It will be included in the revised version.*

In PPML, the dependent variable (trade flow) is introduced in levels instead of logarithms. The findings in Table 1 show that the coefficients of two indicators of infrastructure and institutional quality variables have expected sign and are significant at substantial level. However, the sign of $\ln(\text{border_transport}_{it})$ is changed to negative and statistically significant. We also find sign reversal for some traditional gravity variables ($\ln \text{pop}_{it}$, $\ln \text{area}_i$ and Comrelig_{ij}) that may be because of convergence problems of PPML estimator in Stata. Furthermore, the coefficients of some variables are statistically insignificant.

Table 1: The impacts of border and transport efficiency, quality of economic institutions, physical and communication infrastructure on trade flow of African countries (PPML).

Variable	I (a) Outcome	II(a) Outcome	III(a) Outcome	IV(a) Outcome
$\ln \text{gdpc}_{it}$	0.462** (0.207)	0.361 (0.249)	0.376 (0.272)	0.342* (0.204)
$\ln \text{gdpc}_{jt}$	0.683***	0.654***	0.666***	0.603***

	(0.047)	(0.046)	(0.024)	(0.050)
$\ln pop_{it}$	-2.790***	-2.514**	-5.334***	-3.947**
	(0.936)	(1.202)	(1.903)	(1.830)
$\ln pop_{jt}$	0.981***	1.006***	1.025***	1.001***
	(0.064)	(0.066)	(0.042)	(0.076)
$\ln dis_{ij}$	-0.865***	-0.793***	-0.993***	-0.801***
	(0.112)	(0.123)	(0.068)	(0.143)
$Comrelig_{ij}$	-0.337	-0.426	-0.658***	-0.378
	(0.295)	(0.302)	(0.138)	(0.341)
$comcol_{ij}$	0.136	0.054	0.147	0.007
	(0.213)	(0.221)	(0.172)	(0.225)
$colrel_{ij}$	0.167	0.106	0.611***	0.440**
	(0.238)	(0.212)	(0.123)	(0.196)
WTO_i	0.953***	0.945***	0.958***	0.904***
	(0.186)	(0.195)	(0.105)	(0.236)
WTO_j	8.199***	0.266	4.228*	1.965
	(2.123)	(0.213)	(2.283)	(1.937)
$\ln area_j$	0.043	-0.018	-0.011	-0.003
	(0.049)	(0.063)	(0.024)	(0.049)
$\ln area_i$	2.659***	2.235**	4.313***	3.605***
	(0.689)	(0.873)	(1.378)	(1.313)
$\ln a_{ij}$	0.486**	0.533***	0.274**	0.354*
	(0.189)	(0.189)	(0.108)	(0.199)
Sea_i	-2.798	6.584***	15.710***	11.382**
	(1.843)	(2.397)	(6.024)	(5.387)
RTA_{ij}	0.454**	0.421**	0.223**	0.432*
	(0.182)	(0.186)	(0.099)	(0.222)
$\ln(eco_institutionst_i)$	0.304**			0.680***
	(0.148)			(0.228)
$\ln(border_transport_{it})$		0.897**		1.889***
		(0.404)		(0.628)
$\ln(physical_communication_{it})$			0.398*	0.706**
			(0.206)	(0.302)
$_cons$	-29.452***	-33.096***	-17.194	-54.959***
	(6.396)	(11.909)	(20.370)	(18.133)
<i>Obs.</i>	71416	61,341	37,307	32260
<i>Pseudo R2</i>	0.466	0.475	0.466	0.503

Notes, *** significant at 1%, ** significant at 5%, * significant at 10%, standard error in parenthesis. i and j indicate the reporter and partner country, respectively. Source: Authors calculations.

5. I like the IV models that the authors include for robustness check.
Response: We are grateful for the comment.

6. Please check the last sentence of the paragraph above Table 3.
Response: Thank you for noting it. We have revised it.