

## Reply to the referee 1

Jiang Yushi and Dinkneh Gebre Borojo- The Impacts of Institutional Quality and Infrastructure on Overall and Intra-Africa Trade- No. 2018-75.

Thank you for the constructive comments and positive evaluation. Our answers are indicated in italics.

### Major comments

-The discussion of the problem could be broadened to mention (i) the regional integration processes that aim to reinforce trade ties among African countries, as well as the proposed continent-wide FTA, (ii) go beyond the continent intra-trade average of 11% and show some differences across the regional blocks (see the 2017 WTO's "World Trade Statistical Review", page 12; or the 2018 report), and (iii) the WTO's trade facilitation agreements of 2013 that aim to improve the trade environment and to which a large number of African countries have adhered.

*Response: We agree. We will broaden problem statement and include differences across regional trade blocks in revised version of the paper.*

-The literature review section is too short (only three references have been cited), and tend to be repetitive of the more detailed discussion of the literature in the introduction (more than a dozen papers cited). The authors may consider dropping the literature section. Alternatively, the discussion in the introduction can be reduced and properly summarized, and the literature section will then offer in-depth discussions. The invoked literature could include the paper by Seck (2017). Most of what the authors aim to do have been addressed in the paper which tends to be more comprehensive: impact of various soft and hard infrastructure from both the exporter and importers perspectives (measured in terms of tariff and distance equivalence), both total and intra-regional trade, a comparison between Africa and the rest of the developing world, a simulation exercise, in addition to trade composition (agricultural vs. manufacturing commodities), in the context of a gravity model.

*Response: We agree. Literature review section will be discussed in-depth in the revised version of the paper. We will also drop some literatures cited in introduction section and we will provide detailed discussion in literature review part. We will include work of Seck (2017) in the revised version of the paper.*

-The discussion of the gravity model could start with the standard non-linear form, and then show why the variables are, in the final reduced form, specified in log terms.

*Response: We will extend equation 1 in page 8 and equation 2 in page 9 to their natural logarithm form and we will discuss about variables in the revised version of the paper.*

-A more careful interpretation of the interaction terms is warranted. See the paper by Brambor et al. (2005).

*Response: We thank the referee for suggesting to care for interpretation of the interaction terms. We did the interaction exercise carefully in the submitted paper. We will see the paper recommended by the referee and will review if there is anything to consider.*

-Additional variables to be considered: common currency, tariffs (to the extent that they are not highly correlated with RTA – they allow for instance in the simulation to obtain tariff-equivalences).

**Response:** *We will control for common currency and tariff in the revised version.*

-In addition to PPML and a selection model, negative binomial model can also be used to address the issue of zero trade. It compares with PPML with respect to the dispersion of the data. The author need to mention this, and possibility explain why the negative binomial approach may be ruled out.

**Response:** *We agree that it's important to discuss and compare Heckman selection model to Negative Binomial Pseudo Maximum Likelihood (NBPML) and Zero Inflated Models. We will add this in the revised version.*

-The Heckman type selection model also need to be compared with *hurdle*-type model, the difference being whether the DGPs modeled in the first stage (probability of trading) and second stage (trade flows) are independent. The Heckman regression output show also include the LR test for independence, which the authors need to report and discuss.

**Response:** *We agree that we should compare the model used in the study with hurdle-type model. Regarding LR, it is not listed because we used two-step procedure of Heckman selection estimator.*

-The estimation results tables need to include additional statistics: total number of obs., Wald statistic for joint significance, etc. (in addition to LR test of independence). The footnotes under the tables should rather read “Notes” instead of “Sources”, and the it would be interesting to remind the reader what the explained variable is and how it is specified.

**Response:** *Total number of observations is already included in the tables (No. ob). We have also mentioned censored observation (Censored). They are listed in the last two rows of Table 1 and 2. We will add Wald chi2 (2) for joint significance in the revised version. Certainly, we will use “Note” instead of “Sources” and we will explain variables in the revised version.*

-In discussing the results (“variables of main interest”), it is important to compare with the findings in the previous literature: significance, sign, magnitude, etc.

**Response:** *We will compare our findings with the findings of other literatures.*

-Why the three infrastructure variables have been included separately in the regression analyses, why not account for them simultaneously? Is there any issue of multicollinearity? The authors need to show that.

**Response:** *As there is a relatively high level of collinearity issue among border and transport facilitation, physical and communication infrastructure and quality of institutional indicators, and among individual indicators of each aggregate indicator, we prefer to treat them individually and we all repeated the exercise using them simultaneously (Table 1, Table 2 and Table 9). We will discuss this issue in the revised version.*

-The IV regression results should also include various test results: endogeneity, weak instrument, over-identification, etc.

**Response:** *We agree that including test results is mandatory and we will include them in the revised version. Its good observation and adds to the quality of the paper.*

-In the end, the authors need to tell which one the three dimensions of the trade environment matters the most. This is particularly important from a policy standpoint, as government may need to prioritize among the reforms aiming at boosting trade. Better than the distance equivalence measures, the authors may consider rescaling the variables so that they become comparable, and then the estimated coefficients will provide such information.

**Response:** *We agree with this suggestion and we are grateful to the referee for giving us this feedback. We will standardize variables of our interest and make them comparable for policy framework.*

### **Minor comments**

-Paragraph 1 (page 2): provide data supporting the claim “(...) African countries are among those having the least trade share compared to other regions in the world market”; also add intra-regional trade data from elsewhere to be compared with the intra-trade figure of 11%.

**Response:** *Surely, we will provide data in the revised version.*

-Paragraph 2, page 2: provide some references in support of the claim “Some literatures indicate that tariff and non-tariff barriers are responsible for high cost of trade and the lower performance of developing countries in the world trade.”

**Response:** *We agree. We will provide literatures that support our claim.*

-Paragraphs are relatively too long. (e.g.: paragraph 2, page 2). Think of breaking them up and organize around key ideas.

**Response:** *We will reorganize them.*

-A careful reading is required to correct many typos and improve the writing

**Response:** *We admit that there are many typos and writing errors. We will correct them in the revised version.*

-List of countries in the appendix: instead of the abbreviations, use the entire spellings of countries' name.

**Response:** *We will write full name of countries instead of abbreviations.*

-It is customary to denote the significance level by \*\*\* (1%), \*\* (5%) and \* (10%). The authors may want to follow suit.

**Response:** *We will correct it.*