Supplementary Materials to

Prospects and Challenges for Supply chain trade under the Africa Continental Free Trade Area

Melo and Twum

1. Detailed methodology: measures of supply chain participation

Borin and Mancini (2019) develop decompositions at the aggregate, bilateral and sectoral levels using measures to avoid double-counting that give identical estimates of double counting for the two approaches in GVC analysis: (i) the country where value-added originates (source based), and; (ii) the country that ultimately absorbs it in final demand (sink-based). In the source-based approach, value-added is accounted for the first time it leaves the country or origin while in the sink-based approach, it is considered the last time it crosses national borders. Borin and Mancini develop indices of GVC trade for both approaches at the aggregate, bilateral, and sectoral levels (see figures in sections 3 and 4). They show that double-counting estimates are the same under both approaches. This paper uses the source-based approach.

The source-based approach gives indicators that extend the 'vertical specialization' index first introduced by Hummels et al. (2001) on the basis of national input-output tables is suitable to study the processes in which export flows are involved. With the source perspective, an item is counted as value-added at the first stage of production while in the sink-based approach, it is accounted for at the last shipment Below are the decompositions used for the figures and tables in sections 3 and 4 in the main text.

For a sample of N countries where U_N is the identity matrix of rank N, the GVC related trade for source country (s) exporting to destination country (r):

$$\mathrm{GVCX_{sr}}\ (\text{GVC related trade}) = \mathrm{U_N E_{sr}} - \mathrm{DAVAX_{sr}} \qquad \text{A1(a)}$$

$$\mathrm{GVC_s}(\ \text{GVC participation rate}) = \frac{\sum_{r\neq s}^{G} \mathrm{GVCX_{sr}}}{\mathrm{U_N E_{s*}}} \qquad \text{A1(b)}$$

$$\mathrm{u_N E_{sr}}(\text{bilateral exports}) = DVAsource_{sr} + DDCsource_{sr} + FVAsource_{sr} + FDCsource_{sr} \qquad \text{A2}$$

where;

$\mathrm{E_{sr}}$: bilateral exports ;
$\mathrm{DAVAX_{sr}}$: the domestic value added in exports from exporter (s) absorbed by importer (r);
$DVAsource_{sr}$ and $FVAsource_{sr}$: domestic value added and foreign value added exports from country s to country r;
$DDCsource_{sr}$ and $FDCsource_{sr}$: domestic double counted and foreign double counted exports from country *s* to country *r.*

$\mathrm{GVCX_{sr}}$ in equation A1(a) is the GVC share in bilateral exports. It is the sum of domestic and foreign value-added exports ($\mathrm{U_N E_{sr}}$) excluding the domestic value added in exports absorbed directly by its importer (DAVAX). $\mathrm{DAVAX_{sr}}$ is the sum of i) the portion of production that is produced entirely in country s and exported to country r with no intermediates from outside its borders and ii): the

intermediates inputs that are entirely produced in country s and exported to country r for the production of final goods that are entirely consumed by the domestic market of the importing country, r[1]. $\mathrm{GVC_s}$ expressed total GVC related trade from country s as a share of its total gross exports.

$\mathrm{GVC_S}$ in A1 (b) is the GVC participation rate. It avoids double counting of trade flows of intermediates. The measure can be decomposed into backward and forward GVC participation components:

$$\mathrm{GVC_s} = \mathrm{GVCbackward_s}(GVC_{bs}) + \mathrm{GVCforwad_s}(GVC_{fs}) \qquad \text{A3}$$

Backward GVC related trade (GVC_{bs}) measures country's exports that include value added previously imported from abroad. For example, if Rwanda imports maize from Uganda for the production of fortified foods for export, then Rwanda is said to be participating in backward GVC participation when it exports fortified foods.

Forward GVC related trade (GVC_{fs}) measures the country's exports that are used by the importing country to produce for export. In the example of fortified foods exports, Uganda is engaging in forward GVC participation because its exports are used as intermediates by Rwanda for the production of its fortified food exports. We use the Borin and Mancini indicators of GVC related trade presented in (1) and (3) in the sections 3 and 4 to explore the participation of African RECs, and the EAC in particular, in supply chain trade.[2]

Length of the production chain. GVCs also evolve along the dimension of tasks (which can be carried out domestically or abroad). The length of a supply network, say in textiles and apparel, is the result of the forces of agglomeration (better robots reduce the benefits of specialisation, i.e. reduce the number of separate tasks or the propagation length) and the forces of dispersion (better communication costs lower the marginal cost of coordination and hence favour outsourcing). [3] Average propagation length measures the number of production stages in a production process in a country, a region (or the world).[4]

As an illustration, take Figure A1 that shows the total number of production stages (i.e. the length of the production network for the textile, leather and footwear sector). This is a prominent sector of GVC entry for lesser-developed countries. The data are from the OECD-WTO TiVA database for 2008. The figure shows that the total number of production stages is less than 2 when all industries are covered with 12% taking place in a foreign country (Diakantoni and Escaith (p.6)). In figure A1, Brazilian firms are relatively concentrated with only 11% of the stages sourced internationally. China and Vietnam

[1] GVC measures are also available at a sectoral level. In this case equations 1(a-b) will be adjusted to account for the sectoral –bilateral perspective. We do not show the underlying equations in this paper, however the reader can find them in Borin et al. 2019 (section 4.2 p. 31)

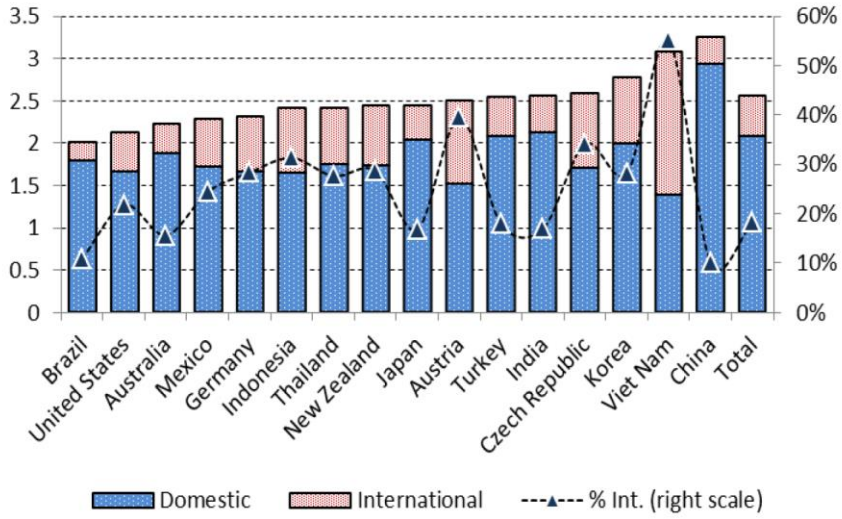
[2] Here vertical is used in the sense that all the stages of the supply chain are carried out domestically as shown for the T&A example in figure 1.

[3] Baldwin and Forslid (2013) discuss how a change in automation (improved information technology) combined with a change in coordination costs (improved communication costs) determines a firm's optimal number of stages and hence the propagation length.

[4] Antras and Chor (2019) and Wang et al. (2016) define a GVC position index that measures the distance from any production stage between the final demand and the initial factor inputs in a production line. If a country's representative production chains towards final products are longer than those towards primary products, the country is considered to be operating in a relatively upstream position. As an example, Inomata (2017, figure 1.10) shows that for an aggregate of the production networks in East Asia over the period 1995-2009, China stayed in the most downstream segment of the lengthening network in East Asia. This is an indication that China held a dominant role as a final assembler of regional products. See discussion in annex A1.

have much longer supply chains. For Vietnam, 55% of sourcing is foreign while for China, 90% of sourcing is domestic.

Figure A1: Supply chains in Textile, Leather and Footwear (2008)



Source: Diakantoni and Escaith (2014, figure 1) computed from the OECD-WTO TiFVA data base for 2008.
Explanatory notes: Vietnam (China) has 3 (3.2) production stages in the sector with 57% (10%) originating abroad. The left axis is the number of stages and the right axis is the percentage of intermediates sourced internationally

In practice, cross-border activities along a production chain are limited. This reflects three factors that influence the length and geography of production networks. First, for transaction costs, a task perspective operates like compound interest with transaction costs increasing more than proportionally with the number of stages.[5] Second, the effect of a marginal variation in trade costs along the chain is much larger when there is more than one international transaction. [6] Third, a small decrease in tariffs (or more generally border-related transaction costs) can induce a tipping point at which vertical specialization (i.e. cross-border trade in tasks) kicks in (Yi (2003)). These obstacle explain the low percentage of stages involving more than one country in figure 1.[7]

2. Trends in backward and forward integration rates: African RECs and comparators

[5] An example, taken from Diankantoni and Escaith illustrates the importance of lowering trade costs along a supply chain. Let trade costs apply in proportion to the value of the good. Ferrantino (2012) shows that the total costs of delivering the final good to the consumer increases exponentially with the number of production stages. With 5 [10] stages and an ad-valorem transaction cost of 10%, the ad-valorem tariff equivalent is 34% [75%].

[6] Rouzet an Miroudot (2013) estimate that EU pay an average tariff of 3.7% on imported products from India with only 51.5% being paid at the EU border. Because the cumulative effect of tariffs (and other border costs) is bound so long as intermediate goods of domestic origin along the supply chain are substitutable, to some extent at least, complex GVCs cannot develop when tariffs are above a certain threshold. The same applies to RVCs where tariffs are zero, but other border trade costs can be important.

[7] This measure depends on the level of aggregation in the data.

Table A1: Trends in backward and forward GVC integration

	Backward integration				Forward integration			
	1990	2005	2015	Trend	1990	2005	2015	Trend
World	0.24	0.29	0.28		0.18	0.20	0.20	
China	0.05	0.18	0.14		0.17	0.18	0.21	
India	0.08	0.15	0.16		0.17	0.20	0.19	
ASEAN	0.37	0.40	0.36		0.13	0.17	0.18	
EAC	0.09	0.16	0.19		0.18	0.20	0.19	
ECOWAS	0.09	0.09	0.07		0.22	0.25	0.28	
MERCOSUR	0.07	0.12	0.12		0.17	0.18	0.18	
SADC	0.15	0.16	0.16		0.21	0.24	0.24	
COMESA	0.24	0.20	0.17		0.17	0.22	0.24	

Source: Authors calculations using GVC database from Borin and Mancini (2015, 2019).
Note: GVC participation measures the share of country's exports that either makes use of value-add imported from another country or is exported to another country for further processing. It is the share of GVC related trade for a country over its total gross exports. See Box 1 for details on GVC measures used in this paper: backward, forward, regional and non-regional value chain participation

3. Specification tests

To decide between random and fixed effects we run the Hausman Test to test whether the errors are correlated with the regressors where the null hypothesis is that they are not and thus Random effects would be preferred to fixed effects. The results below indicate that fixed effects is the preferred specification.

Table A2: Results from Hausman Test

	Coefficients (b) fe	(B) re	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
lnGDP_pc	-.2658604	-.1057652	-.1600951	.0299415
lnMANUSHR	.0684709	.0994516	-.0309807	.0223681
lnFDI_pc	.0656876	.095022	-.0293344	.0061154
lntarZM	-.0086488	-.0247683	.0161195	.0042132
lnTRCOST	-.0576557	-.4301622	.3725065	.0547643
lnMOBS	.0392144	.0424477	-.0032333	.0022826
year				
2003	.0050246	-.0034771	.0085017	.0030661
2006	.1011041	.0186471	.082457	.0117668
2009	.1520051	-.0037082	.1557133	.0218873
2012	.1949566	-.0146592	.2096157	.0277919
2015	.1459398	-.0935582	.239498	.0308871

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(11) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 147.77
Prob>chi2 = 0.0000

Next, we run a test to ascertain whether we need to include time-fixed effects when running a FE model. We use the command testparm, a joint test to see if the dummies for all years are equal to 0,

in which case no time fixed effects are needed. From our results below, we confirm that we must include time fixed effects in our regression specifications.

Table A3: Results from testparm specification in Stata

```
. testparm i.year

 ( 1)  2003.year = 0
 ( 2)  2006.year = 0
 ( 3)  2009.year = 0
 ( 4)  2012.year = 0
 ( 5)  2015.year = 0

       F(  5,   377) =    5.81
            Prob > F =    0.0000
```

Lastly, we run a test to measure the presence heteroscedasticity. Under the test we use, the null is homoscedasticity (or constant variance). In the table below, we reject the null and conclude heteroscedasticity is present. To account for this, we estimate a FE model using Driscoll-Kraay standard errors, which also account for auto-correlation and cross-sectional dependence.

Table A4: Test to ascertain the presence of heteroscedasticity

```
Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (91)  =    4.4e+05
Prob>chi2 =     0.0000
```

4. Robustness checks

1995-2005 vs 2006-2015

We first run our main estimations on two sub-periods between 1995 and 2015 (1995-2005 and 2006-2015). Table A6 and A7 report the results. These are largely the same with stronger significance and magnitudes in the later periods (2006-2015) for some of our policy and non-policy variables of interest.

Regression including services share in GDP

We also add as an independent variable, the services share in GDP, which is strongly correlated with the manufacturing share in GDP. Both variables show no significant association with overall GDP participation similar to our main estimation in Table 1: trade costs and tariffs on intermediates remain negatively and significantly associated with overall GVC participation. Our findings are similar for the coefficients on forward and backwards GVC participation with services share in GDP showing a negative and significant association with forward GVC participation and no positive association between manufacturing share in GDP and backward GVC participation (under our main specification manufacturing share in GDP was positively associated with backwards GVC participation for our all-country sample). As in the results in table 1, there are no significant results for our specification on the group of African countries. See Table A6 and A7.

Balanced vs. Unbalanced dataset

Next, we run our main regression on a balanced panel and compare results to our main regression estimation using an unbalanced panel. Table A5 below shows the share of missing values for the independent variables in out regression specification; missing variables range from between 18 to 29 percent.

Table A5: Share of missing variables

Variable	Number of missing observations	Share of total observations
GDP per capita	560	18.9
FDI per capita	798	27
Manufacturing share in GDP	727	24.6
Trade costs	794	26.8
Tariffs on imports of intermediate goods	746	25.2
Number of mobile phone subscribers per 100 inhabitants	834	28.2

To create a balanced dataset, we interpolate missing variables using the nearest observation. Specifically, we use the stata code *mipolate* which estimates missing values using known values either before or after missing values, depending on which is nearer. When values before and after are equally distant from a known value we use the mean of the two values. Results using this balanced dataset are similar to those based on the unbalanced dataset. For most, the direction and significance remain the same, but estimated coefficients have lower magnitudes because of smoothing. Notably, the phone subscriber variable displays a negative significant association with forward GVC participation when using the balanced dataset. See Table A6 and A7.

Yearly variable estimates vs. 3-year averages

In our paper, we use 3-year averages to account for yearly fluctuations and to give time for adjustment to changes in values of the regressors. Under a yearly specification, the results are largely similar in direction and significance except for the loss of significance for the GDP and trade cost coefficients for forward GVC participation and gaining significance on the trade cost coefficient for backward GVC participation. See Table A6 and A7.

Balanced vs. unbalanced dataset for estimations on group of African countries

We look at results for both unbalanced and balanced datasets for our regression on the group of African countries. The results differ in three ways: using a balanced dataset, we find a positive and significant association between FDI and overall GVC participation but a negative and significant association for the number of mobile phone subscribers; we also see a negative association between tariffs on intermediate imports and backward GVC participation. See Table A6 and A7.