SMALL TRADE FLOWS, COMPLIANCE COSTS AND TRADE PREFERENCES: THE CASE OF EU IMPORTS FROM AFRICAN LDCs

The motivation of the paper is well founded: seeking an explanation for the low rate of utilisation of tariff preferences in low value imports from Africa. However, the authors' development of the topic is limited and less interesting. They provide an extensive description of the data (which I do not believe adds much to the paper due to the high level of aggregation considered), while the development of the model and the discussion of the results is minimal and not very enlightening.

Let us review with some detail the two main sections of the paper.

SECTION 2

In relation to the analysis of Tables A1 and A2 in this section, the authors could interpret the results more directly if the indicators in columns 2 and 3 were presented in the same way as the PUR (column 1). That is, bearing in mind that the PUR indicates the % of imports (in value) that have made use of tariff preferences, it would be logical for column 3 to also include the % of import transactions that use tariff preferences, rather than the % of transactions that do not use such preferences, as is the case in the current version of the manuscript. This would help to clarify the existing divergence between the utilisation rates expressed in terms of value and number of operations (for example, in the case of Angola the utilisation rate would drop from 76% in terms of value to 2% in terms of the number of operations).

In addition, and bearing in mind that it is impossible to establish the number of import operations that make up each of the observations of the model, is the figure referring to the % tariff preference utilisation in terms of the number of operations really valid? That is, in the sample, observations of one sole transaction might be getting mixed up with the observations that group 2, 3, 4... N import flows. In order for this indicator to really be meaningful, we would have to be sure that either one observation refers to one sole transaction or one sole importer, or that the number of operations per observation is constant for all the items in the sample. Finally, the means presented in column 7 refer to highly aggregate data, so it would be interesting to have information on their standard deviation.

It would be worth if the authors' could provide more detailed information on the nature and source of the data used beyond that in the second paragraph on page 7 (for example, the institution responsible for compiling the COMEXT database, the link, the origin of the data, etc.).

SECTION 3

In the first place, the formal specification of the model has not been explained sufficiently and it is not clear that is correct from an econometric point of view. In light of the objective of the paper and the topic of the special issue that applies, it is true that it makes no sense to explain econometric issues to any great extent. However, the authors should at least establish a theoretical framework that encompasses the tool employed in the analysis. All they would have to do is mention that the tool used in the analysis is a discrete choice model and refer interested readers to papers such as Train (2003).

What the equation on page 6 is reflecting is the utility function linked to using tariff preferences, not the probability of choosing to make use of those preferences. An equation like the following is perhaps more appropriate:

$$U_{PA,n} = \alpha + \beta.PVOP_n + \sum\delta_{PAm}.COUNTRY_{mn} + \sum\Phi_{PA,n}.PRODUCT_{jn} + \epsilon_{PA,n}$$
$$U_{NPA,n} = \beta.PVOP_n + \sum\delta_{NPAm}.COUNTRY_{mn} + \sum\phi_{NPA,n}PRODUCT_{jn} + \epsilon_{NPA,n}$$

whereby
$U_{PA,n} \rightarrow$ utility of the person in charge of import operation "n" making use of tariff preferences.
$U_{NPA,n} \rightarrow$ utility of the person in charge of import operation "n" not making use of tariff preferences.
$\alpha \rightarrow$ alternative specific constant linked to the option of using tariff preferences.
$COUNTRY_{mn} \rightarrow$ dummy variable that takes a value of 1 if import flow "n" comes from country "m" and 0 otherwise.
$PRODUCT_{jn} \rightarrow$ dummy variable that takes a value of 1 if import flow "n" refers to product "j" and 0 otherwise
$\epsilon_{PA,n} \rightarrow$ random error normal identically and independently distributed
$\epsilon_{NPA,n} \rightarrow$ random error normal identically and independently distributed

instead of $P(util\_rate=1)= \alpha + \beta.PVOP_n + \sum\delta_m.COUNTRY_{mn} + \sum\phi.PRODUCT_{jn} + \epsilon_{PA,n,}$ as considered by the authors, because the probability of the exporter making use of tariff preferences is equal to $Prob(U_{PA,n} > U_{NPA,n})$, which, bearing in mind that the authors have opted for a probit model (that is, assuming that $\epsilon_{PA,n}$ and $\epsilon_{NPA,n}$ are distributed iid as a normal) is the same as the integral of the density function of ϵ_n.

Why do the authors use a PROBIT model instead of a LOGIT model, when the solution to the former is not closed and therefore requires a simulation?

Furthermore, the authors should establish at some stage in the paper that the dummy variable UTIL_RATE takes a value of 1 if the importer decides to make use of tariff preferences and 0 otherwise, together with the meaning of sub indexes j, k, m and n.

On the other hand, Table 5 provides the results of the different specifications, but only one of them is detailed in the manuscript. The authors should provide the functions of all three specifications in the table of results.

Results in column (1) of Table 5 refer to the model detailed on page 6. Is that correct? In any case, why is no information provided regarding the estimation and significance of the specific constant (α) and the coefficients linked to the country and product dummies? What software have the authors used to perform the estimations? Please give information about the meaning of * and ** in Table 6. Does the lack of * mean that the coefficient obtained is not significant at 1%?

The way the authors interpret the results, on the basis of the coefficients obtained, is also misleading (perhaps incorrect). They interpret the coefficients as if we were dealing with a linear model in which the value of the coefficient directly captures the effect that a marginal variation in the attribute has on the probability of choice. However, isn't the absolute value of the coefficient meaningless in economic terms in discrete choice models? Shouldn't the variation in the probability of choice be calculated in response to variations in the explanatory variables using micro-simulation and sample enumeration and shouldn't the direct interpretation of the coefficient be confined to the sign obtained?

One of the aims of the paper is to establish a relationship between the probability of using and not using tariff preferences and the potential value of said preferences (that is, what applying the preferences would save the importer, which is equal to the value of the import multiplied by the corresponding tariff), as well as to ratify the evidence available that suggests a lower use of trade preferences in the case of low value imports by means of a model.

In order to do so, the authors specify the explanatory variable as PVOP = value of the import x MFN tariff. Have the authors tried specifying these two variables separately? Therefore, as it is, their current approach makes it impossible to isolate one effect from another. If the authors specified the variables separately, they could test both hypotheses:

- <u>Variable that captures the value of the import</u>: according to the initial hypothesis the paper considers, the sign is expected to be positive, indicating that, ceteris paribus, increases in the value of the import transaction increase the probability of the importer using tariff preferences. The authors could also try specifications in which the value of the import flow is included as a dummy variable instead of a continuous variable. This would allow them to test the importance of rules of origin in the decision under analysis by specifying a dummy variable that takes a value of 1 if the import transaction is worth less than $6,000 (imports that are exempt from submitting a certificate of origin). A positive sign would indicate that imports exempt from presenting a certificate of origin are more likely to make use of tariff preferences than those which are required to submit this document.

- <u>Variable that captures the MFN tariff of the shipment</u>: the sign is also expected to be positive in this case, as increases in the benefits stemming from applying the preference system are expected to increase the probability of making use of such preferences. In relation to this variable, it would be interesting to see what happens when it is specified as a dummy variable instead of a continuous variable. One of the hypotheses of the authors is precisely that what the exporter does is to compare the benefits he would reap from applying preferences (and which are equal to the tariff that would no longer have to be paid) to the cost incurred through applying it, such that only when the benefit surpasses the cost would the exporter request the application of the preference. However, the specification the authors use cannot ratify their hypothesis, as no proxy of the cost of the utility function has been included. On a different note, the literature review in Section 1 refers to the 4% threshold estimated by Manchin (2006): the procedure that requesting tariff preferences entails is only worthwhile when the tariff that no longer has to be paid is greater than 4%. They could test the threshold obtained by Manchin by introducing a dummy variable in the model that takes a value of 1 if the tariff is greater than 4% and 0 when it is less than or equal to 4%, or even several dummy variables linked to different levels of tariffs: for example, less than 2%, between 2% and 4% and greater than 4%. In that case, we would expect the 2% and between 2% and 4% dummy variables to display negative signs and the greater than 4% dummy a positive sign. The authors could test the most suitable levels of these thresholds by applying econometric tests.

In any case, in order for the results obtained to be considered valid, we should be certain that the level of disaggregation of the observations refers to one sole import flow, which the authors themselves recognise is not the case (first paragraph in Section 2): a preference utilisation rate of between 0% and 100% can still imply that more than one transaction is recorded but that all the transactions either utilise preferences or not. This aspect could still be upheld if all the transactions that their observation includes referred to the same importer, but if this is not the case, if the decisions of various importers are included in the same observation (who, in agreement, have decided the same), the authors would be making a conceptual error,

as in some cases we would be assessing individual decision, while in others we would be evaluating a joint decision taken by a group of importers.

Furthermore, if one same observation aggregates several import transactions, we would be cancelling out the effect of the "paperwork" associated to each transaction in order to be able to apply tariff preferences on that probability of opting to use tariff preferences. That is, an observation may include an import transaction value of €100,000, which has not made use of a tariff preference. We would be considering this as one sole import transaction and that if tariff preferences of, say, 2% were applied, and we would be saving €2,000. However, in reality the observation includes 100 transactions of €1,000 each, implying a tariff saving of €20. Therefore, bearing this in mind, while it does not seem very rational in the case of a €100,000 transaction eligible for a saving of €2,000 that an importer would not do the extra "paperwork" required, but it is not illogical to forego the use of preferences when they entail a saving of only €20, but involve much more paperwork.

Finally, the authors seek to capture the effect that the type of commodity has on the probability of deciding to use tariff preferences. Initially, the idea is correct, because as they say in the paper, the severity of rules of origin varies depending on the product. However, it looks like that the "good intentions" behind this specification make no sense whatsoever given the level of aggregation the authors consider. Without knowing in detail how rules of origin work, intuitively we can assume that they vary substantially within one same section. As a result, what the coefficient of the dummy variable linked to a section or to another TARIC may indicate is entirely meaningless. In this sense what would be really interesting is to perform an analysis with sufficiently disaggregated data to really reflect the differences in rules of origin.