**CHAPTER 10: Documentation of Data and Analysis**

**Data file**

The data file CAP11\_DataCotonou contains all the essential variables concerning the countries that benefit from the Cotonou scheme (data pertaining to the GSP and GSP- Drugs schemes are available on request). The data sources are described in the chapter.

**Do file**

The do file brings in the data and labels the variables used in the analysis. We define the variables following Table 11.2.

label variable util\_acp "Ratio of EU actual imports under Cotonou scheme to the value of total imports, computed at the country-HS6 product line level"

label variable m\_acp "Preferential margin computed as (MFNik - PREFik)/MFNik"

label variable imp "EU import flows from Cotonou eligible countries, at the HS6-digit level”

label variable lngdp "log of GDP of exporting country"

label variable lnpop "log of population of exporting country"

label variable lndist "log of the distance, in kilometres, between each exporter i and the EU"

label variable comlang "Dummy = 1 if the exp. country language is spoken in at least one EU states"

label variable colony "Dummy = 1 if colonial links existed between the exporter and at least one EU states"

label variable free\_index "Heritage Foundation and Wall Street Journal Freedom Index"

label variable id\_exp "Identity of exporting country”

The rest of the file is divided in 3 parts.

1. **Tobit Model.**

*(a)* we retrieve a measure of compliance costs by estimating a model (Model 1 in the paper) of the preferences utilisation determinants. More precisely, the aforementioned costs are estimated residuals retrieved from a regression, explaining the utilisation ratio of the Cotonou preferential regime. As the dependent variable ranges from 0 to 100, we estimate a double-censored Tobit model.

xi: tobit util\_acp m\_acp lngdp lnpop lndis comlang colony free\_index , ll(0) ul(100) cluster(id\_exp)

*(b)* Then, we regard the difference between observed and predicted values as an estimated measure of compliance costs

predict y\_fit if e(sample), xb

g costs\_acp=(util\_acp-y\_fit)

drop y\_fit

sum costs\_acp

To facilitate the interpretation of the results, we compute a measure that increases as the costs increase, by considering the difference between the maximum sample residual and each of the residuals.

g costs\_acp2=(1082.61- costs\_acp)

sum costs\_acp2

*(c)* We generate the log of costs and preferential margin; we construct the interaction term between these variables, then we drop the observations laying in the first and last percentile of the costs distribution.

g lncosts\_acp=ln(1+costs\_acp2)

g lnmarg\_acp=ln(1+ m\_acp)

g int\_acp=lnmarg\_acp\*lncosts\_acp

centile(lncosts\_acp), centile(1,99)

drop if lncosts\_acp< 4.342206

drop if lncosts\_acp> 6.696539 & lncosts\_acp!=.

1. **Poisson Model.**

*a)* We estimate equation (11.1) adopting a Poisson model (Santos Silva and Tenreyro, 2006) to address the problem of heteroskedastic and non-normal residuals in gravity regressions.

xi: poisson imp lnmarg\_acp lncosts\_acp int\_acp i.id\_exp , cluster (nc\_6) difficult technique(bfgs)

*(b)* We test the significance of our key regressors

test lnmarg\_acp lncosts\_acp int\_acp

test lnmarg\_acp int\_acp

1. **Negative Binomial Regression.**

*(a)* As a robustness check, we replicate our gravity estimations by adopting a Negative Binomial model

xi: nbreg imp lnmarg\_acp lncosts\_acp int\_acp i.id\_exp , cluster (nc\_6) difficult

*(b)* We test the significance of our key regressors

test lnmarg\_acp lncosts\_acp int\_acp

test lnmarg\_acp int\_acp