This application relates to chapter 3 and partly replicates Disdier, Fontagné and Mimoui (2008) but using more recent data.

It focuses on agri-food products (HS chapters 01 to 24) and investigates the trade effects of SPS and TBTs on the exports of 189 countries to 23 OECD partners (20 European countries, Chile, Japan and Mexico) in 2012. The empirical application consists in the estimation of a gravity equation as described in module 3.

The database "dataset_final.dta" is available on the UNCTAD course website and already includes the variables needed for the estimations. As in previous application the NTM data come from the CEPII website and are based on the TNT database. The trade data are extracted from the BACI database also provided by the CEPII. Other gravity variables, such as distance, contiguity, common language and past colonial links are also from the CEPII. Exporting and importing countries' size is proxied by the gross domestic product taken from the World Development Indicators database. Finally, we also control for the bilateral applied protection using tariff data retrieved from the using the MAcMap version of the UNCTAD/TRAINS database.

The database is built at the Harmonized System Classification 4-digit level and two variables account for NTMs:

- A simple dummy variable set to 1 if the importing country notifies at least one SPS/TBT measure on one HS 6-digit product within the HS 4-digit sector (0 otherwise);
- A frequency index computed for each importing country and HS 4-digit sector and defined as the number of HS 6-digit lines affected by at least one SPS/TBT measure divided by the total number of HS6 lines within the HS 4-digit sector.

Note that the bilateral tariffs at the HS 4-digit level are computed as the simple mean of applied tariffs for all HS 6-digit lines within the HS4 sector and for each country-pair. Tariffs are for the year 2010. Finally, exporting countries are divided into two groups (OECD versus developing exporters), and intra-European trade flows are dropped from the sample to avoid biases (within the European Union, tariffs are set to zero and the mutual recognition is applied for NTMs).

The first step consists in opening the data into STATA and finalizing the dataset.

Before opening the dataset you may want to change the default directory that is the directory STATA will use unless another directory is specified. We created a folder called STATA_NTM (c:/ STATA_NTM) and copied all files that are needed to run the application (data files and do-files) into that folder. To change the default directory just run

```
cd c:/ STATA_NTM/Application2
```

```
/*If you want to save the output of your session you must open a so-called log file */
log using application_2.log, replace /*it is in a text format*/
```

To open the dataset in STATA, the command "use" should be applied. Before running the estimations, we add some labels to the variables with the label var command. Labels are self-explanatory.

```
use dataset_final.dta, clear
* If you had not changed the default directory the right command would be
* use c:/ STATA_NTM/dataset_final.dta, clear

label var lgdp_o "log GDP exporter"
label var lgdp_d "log GDP importer"
label var ldist "log distance"
label var mtariffs "mean bilateral tariffs (HS4, simple average)"
label var dum_spstbt_jk "= 1 if at least 1 SPS/TBT at the importer-HS6 level"
label var freq_spstbt_jk "frequency index for SPS/TBT by importer"
label var contig "common border"
label var comlang_ethno "common language (spoken by at least 9% of the pop in both
countries"
label var colony "colonial links"
label var hs2 "HS 2-digit code"
label var hs4 "HS 4-digit code"
label var imp_hs2 "Importer-HS2 fixed effects"
label var exp_hs2 "Exporter-HS2 fixed effects"
label var imp_hs4 "bilateral HS4 imports"
label var limphs4 "log of bilateral HS4 imports"
label var oecd_d "=1 if importer is an OECD country"
label var oecd_o "=1 if exporter is an OECD country"
label var dc_o "=1 if exporter is a developing country"
label var eur_o "=1 if exporter is an EU country"
label var eur_d "=1 if importer is an EU country"
describe
```

We can now run a first set of basic estimations to get a rough idea of the possible impact of NTMs on bilateral exports of agri-food products. Our dependent variable is the log of imports from country B into country A of food products as represented by the corresponding 4-digit categories of the Harmonized System (HS) classification. This is slightly more aggregated by what is usually assumed to represent a product that is at least the 6 digit categories of the HS classification.

All specifications include the gross domestic products of both partners (proxies of their size) and fixed effects defined at the HS 2-digit level to control for unobservable sector characteristics. Standard geography gravity variables are also included. Error terms are clustered at the country-pair level. We also control for the heteroscedasticity with the "robust" option. Finally as it is usually done in the gravity literature, continuous variables are in logs and estimated coefficients can therefore be interpreted as elasticities. Our sample is restricted to strictly positive trade flows.

Note that as we are in a cross section set up the areg command is easier to implement than the xtreg command which works essentially with panel data. Both commands involve variables transformation to account for the one dimension fixed effect.

You should run all specifications at once and then compare estimated coefficients across specifications.

Estimation #1 only tariffs are controlled for but not NTMs

```
xi: areg limphs4 lgdp_o lgdp_d ldist contig comlang_ethno colony mtariffs, absorb(hs2) cluster(groupbil) robust
```

Estimation #2 controls for tariffs and includes a dummy for SPS/TBTs

```
xi: areg limphs4 lgdp_o lgdp_d ldist contig comlang_ethno colony mtariffs dum_spstbt_jk, absorb(hs2) cluster(groupbil) robust
```

Estimation #3 controls for tariffs and include a frequency index for SPS/TBTs

```
xi: areg limphs4 lgdp_o lgdp_d ldist contig comlang_ethno colony mtariffs freq_spstbt_jk, absorb(hs2) cluster(groupbil) robust
```

Estimation #4 focuses on European imports and includes a dummy for SPS/TBTs

```
xi: areg limphs4 lgdp_o lgdp_d ldist contig comlang_ethno colony mtariffs dum_spstbt_jk if eur_d == 1, absorb(hs2) cluster(groupbil) robust
```

Estimation #5 focuses on European imports and includes a frequency index for SPS/TBTs

```
xi: areg limphs4 lgdp_o lgdp_d ldist contig comlang_ethno colony mtariffs freq_spstbt_jk if eur_d == 1, absorb(hs2) cluster(groupbil) robust
```

When significantly different from zero all estimated coefficients have the expected sign and SPS/TBT impact is consistent independently of the indicator chosen. Note however that the quantitative interpretation of coefficients estimated for the dummy indicator and the frequency index variable is not the same. In the former case the elasticity equivalent can be approximated by taking the exponential of the estimated coefficient and then subtracting one. For instance estimation 2 suggests that the presence of a TBT/SPS reduces bilateral imports by exp(-0.30)-1 that is approximately 26 percent. With a frequency index the estimated coefficient is a semi-elasticity and then the interpretation is somewhat less straightforward. The coefficient indicates the impact on the dependent variable in relative terms of a variation of one unit of the index. In other words results are not immediately comparable quantitatively. What is however certain is that they share the same sign and both are significantly different form zero.

Regressions limited to EU destinations indicate that these destinations are most likely to dive overall results as they concern about four fifth of the trade relationships of the sample.

In the next step, we replace countries' gross domestic products by two sets of fixed effects: one for the exporter and one for the importer. These fixed effects are interacted with HS 2-digit fixed effects. In other words we control for any effect which is country and sector specific on both the importer and the exporter side.

The inclusion of a larger set of fixed effects is essentially motivating by the necessity to reduce as much as possible any source of endogeneity whether we consider omitted variables or reverse causality. GDP variables disappear as they are perfectly collinear with a linear combination of the fixed effects included.

We replicate the previous set of regressions with our extended group of control variables. As before you can run all estimations at once and then compare the results across specifications.

Note that we use the reg2hdfe command. This command transforms variables to deal with two-dimension fixed effects estimation. The command should be installed first as it is not part of the standard STATA command package. To do so run

```
ssc install reg2hdfe
```

You can now launch your regressions

```
* Estimation #6: control only for tariffs but not for NTMs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtariffs, id1(imp_hs2) id2(exp_hs2)
cluster(groupbil)

* Estimation #7: control for tariffs and include a dummy for SPS/TBTs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtariffs dum_spstbt_jk, id1(imp_hs2)
id2(exp_hs2) cluster(groupbil)

* Estimation #8: control for tariffs and include a frequency index for SPS/TBTs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtariffs freq_spstbt_jk, id1(imp_hs2)
id2(exp_hs2) cluster(groupbil)

* Estimation #9: focus on European imports and include a dummy for SPS/TBTs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtariffs dum_spstbt_jk if eur_d == 1,
id1(imp_hs2) id2(exp_hs2) cluster(groupbil)

* Estimation #10: focus on European imports and include a frequency index for SPS/TBTs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtariffs freq_spstbt_jk if eur_d == 1,
id1(imp_hs2) id2(exp_hs2) cluster(groupbil)
```

What you should notice is that although results are similar from a qualitative point of view (coefficient signs do not change) if compared with the previous set of estimations we do observe some significant differences in terms of coefficient values. The most striking is the estimated coefficient of distance. It is now about -0.75 while it was about -0.1 in estimations with one dimension fixed effects. This suggests that estimations 1 to 5 are likely to suffer from some omitted variable bias.

If we look at estimation 8 the impact of SPS/TBTs is now equivalent to a trade reduction of 18 percent. It was equal to minus 27 percent in estimation 2. If we move to results obtained for estimation #9, then the impact of SPSP/TBTs measured by a frequency index appears to be non-significantly different form zero.

Once again this is clear evidence that accounting for the largest possible set of fixed effects is crucial.

As in Disdier, Fontagné and Mimouni (2008), we can also examine the potential difference in the trade impact of tariffs and NTMs across exporters. To do so, interaction terms between the tariff and the NTM variables and the two dummies identifying OECD and developing exporters should be included in estimations.

We first construct our set of interaction terms simply by multiplying the variables we want to interact with each other.
For instance the mtar_oecd variable is the interaction between the tariff variable and the dummy variable signaling the exporting countries which are OECD countries.

```
gen mtar_oecd = mtariffs * oecd_o
gen mtar_dc = mtariffs * dc_o
gen dumntm_oecd = dum_spstbt_jk * oecd_o
gen dumntm_dc = dum_spstbt_jk * dc_o
gen freqntm_oecd = freq_spstbt_jk * oecd_o
gen freqntm_dc = freq_spstbt_jk * dc_o
```

Let us first run estimation 11. Estimation #11 is comparable to estimation #7 but with interactions.

```
* Estimation #11: control for tariffs and include a dummy for SPS/TBTs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtar_oecd mtar_dc dumntm_oecd
dumntm_dc, id1(imp_hs2) id2(exp_hs2) cluster(groupbil)
```

Estimated coefficients suggest that tariffs are more of an issue for oecd exporters than for non-oecd ones and that the negative impact of SPS measures and TBT's is similar whether exports are form OECD countries or from non-OECD ones.

Note that Estimation #11 is equivalent to the following estimation called #11'

reg2hdfe limphs4 ldist contig comlang_ethno colony mtariffs dum_spstbt_jk mtar_oecd dumntm_oecd, id1(imp_hs2) id2(exp_hs2) cluster(groupbil)

Let us run estimation # 12.

* Estimation #12: control for tariffs and include a frequency index for SPS/TBTs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtar_oecd mtar_dc freqntm_oecd freqntm_dc, id1(imp_hs2) id2(exp_hs2) cluster(groupbil)

Coefficients obtained for the interaction variables suggest that tariffs are again a concern for exporters from OECD countries only. As to SPS measures/TBTs no coefficient appears to be significantly different form zero. These results are consistent with those obtained for estimation 8 where no effect of SPS/TBT was found. We could have expected some different impact across exporting country groups as in estimation 11 but this not the case.

In the next two estimations we focus on exports directed to EU countries only.

* Estimation #13: focus on European imports and include a dummy for SPS/TBTs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtar_oecd mtar_dc dumntm_oecd dumntm_dc if eur_d == 1, id1(imp_hs2) id2(exp_hs2) cluster(groupbil)

Results on tariffs obtained previously are confirmed. We now observe however that ntms matters for all exporters and that their impact in magnitude is comparable from a statistical point of view.

* Estimation #14: focus on European imports and include a frequency index for SPS/TBTs
reg2hdfe limphs4 ldist contig comlang_ethno colony mtar_oecd mtar_dc freqntm_oecd freqntm_dc if eur_d == 1, id1(imp_hs2) id2(exp_hs2) cluster(groupbil)

Results on tariffs obtained previously are confirmed. We now observe however that ntms again matters for all exporters and with frequency indices their impact in magnitude is comparable across exporters groups from a statistical point of view (estimated coefficients are not different from a statistical point of view based on a standard Wald test).

This exercise is important as it shows that results vary significantly across specifications. The impact of NTM could be either negative or not significantly different from zero depending on the NTM presence indicator we opt for. The choice of the sample also plays an important role. It is also important to reduce as much as possible the incidence of endogeneity by including the largest set of fixed effects that is that still allows the identification of the impact we are interested in, in our case the impact of NTM on bilateral trade flows. It would have been difficult to identify this effect if exporter and HS4 product fixed effects had been included as we work in a cross section and observations of the dependent variable are at the country pair HS4 level.

/* you must close your log file at the end of the session*/

log close