Application of Modern Computer Software Technical in International Trade Education

--- Taking Empirical Teaching of Tariff as an Example

Li Diao [1, a], Yi Cai [2, b]

[1] Economics and Management School, Wuhan University, Wuhan, 430072, China

[2] Economics and Management School, Wuhan University, Wuhan, 430072, China

[a]email: fxyd39@126.com, [b]email:545193969@qq.com

Keywords: international trade, tariff, metering software Stata

Abstract. The main purpose of this paper is to discuss how to verify the case analysis of the typical problems in international trade by modern computer software Stata in practical teaching class of international trade in the category of economics. Study on tariff and market power is taken as examples to discuss the empirical study of tariff. Based on literature written by Christian Broda, Nuno Limao and David E. Weinstein (2008), the conclusions in this paper are tested and verified through model setup and data analysis etc. operation procedures.

There are abundant empirical studies on tariff, focusing on many topics, such as tariff change, influence of tariff (on financial revenue, industry and consumers), tariff policies and behaviors, optimal tariff etc.. The studies made by Christian Broda, Nuno Limao and David E. Weinstein (2008) on tariff and market power are taken as examples to discuss the empirical study on tariff.

Christian Broda, Nuno Limao and David E. Weinstein (hereafter referred to as Christian et al) think the tariff level of a country is related to its market power, and the greater control power the market has, the more protection the country can get from the tariff. Although the tariff distorts the consumption and production, foreign supplied products, if lack of flexibility, may improve the conditions of trade and bring about benefits. Therefore, it can be predicted that the tariff of a country has positive correlation with the control power of the market. The main purpose of Christian et al is to measure the importance of such correlation and its influence on trade policies (mainly on tariff policies). Trade agreement will also have influence on the tariff policies of a country, so it is difficult to define the influence of the market power on the tariff separately. Therefore, the data of the sample countries in this paper are obtained before these countries join WTO, and also based on some presumptions. The year in which the tariff rate has no change in some countries before and after they join WTO is also included.

According to the classical optimal tariff theory, the optimal tariff rate is equal to the reciprocal of the elasticity of export supply, i.e. $t^* = 1/es$. Meanwhile, it is also related to the import volume and the price of imported goods in the country. In a more specific way, we can use the formula $e_s=(dm/dp)(p/m)$, in which, m is the import volume, and p is the price. In order to estimate the influence of market power, it is necessary to get the data regarding tariff, domestic output and elasticity of export supply. In order to guarantee the representativeness and commonness of the sample countries, 15 countries (and regions) are selected as the subjects investigated, and those countries differ from each other in the basic conditions. In consideration of possible difference between the elasticity of import demand and the elasticity of export supply, HS4 category is standard trading data which can differentiate the goods in different industries. All the original data about tariff come from TRAINS, and the trading data comes from COMTRADE (the trading data in Taiwai area come from TRAINS). Due to the large quantity of the data, the specific observed values will not be listed herein, and only the statistical characteristics of the data will be demonstrated in the following table(Table1).

Table1 tariff and elasticity of export supply of 15 counties

Country	Quantity of observed values	Trading period	Tariff			Reciprocal of the elasticity of export supply						
						Mid			Mean		Standard Deviation	
			Mid	Mean	Standard Deviation	Low	Middle	High	All	Deduct the highest part	All	Deduct the highest part
Algeria	739	Mar-93	15.6	23.8	17.4	0.4	2.8	90.9	118	23	333	47
Belarus	703	Mar-98	10	12.4	7.8	0.3	1.5	61.4	85	15	257	36
Boliva	647	Mar-93	10	9.8	0.8	0.3	2	90.9	102	23	283	49
China	1,125	Mar-93	30.3	37.9	26	0.4	2.1	80.1	92	17	267	35
Czech	1,075	Mar-93	5.1	9.5	17.6	0.3	1.4	25.5	63	7	233	18
Ecuador	753	Mar-94	10.6	9.8	5.5	0.3	1.5	55.7	76	13	243	30
Latvia	872	Mar-94	1	7.3	10.5	0.2	1.1	9	52	3	239	8
Lebanon	782	Feb-97	15	17.1	14.8	0.1	0.9	31.3	56	7	215	18
Lithuania	811	Mar-94	0	3.6	7.4	0.3	1.2	23.9	65	6	235	16
Oman	629	Mar-94	5	5.7	8.7	0.3	1.2	25.2	209	7	3536	21
Paraguay	511	Mar-94	14	16.1	11.3	0.4	3	153	132	67	315	169
Russia	1,029	Mar-96	5.7	10.7	11	0.5	1.8	32.8	48	8	198	18
SaudiArabia	1,036	Mar-93	12	12.1	2.6	0.4	1.7	50.2	71	11	232	25
Taiwan	891	92-96	7.5	9.7	8.5	0.1	1.4	131	90	20	241	43
Ukraine	730	Feb-96	5	7.4	7.6	0.4	2.1	77.5	86	16	254	34
Average	782	—	10	9.8	8.7	0.3	1.6	62.6	85	13	243	30

In the above list, the quantity of the observed values in the countries is based on the trade from which the tariff data can be obtained. The basic principle of the estimation for the reciprocal of the elasticity of export supply references to Feenstra(1994) and Broda & Weinstwin (2006), and the estimation is based on the stepwise regression method. Considering the estimation of elasticity may have certain deviation, I verify the above results of the elasticity estimation in three aspects, including: for the same product, whether the elasticity estimated according to the data from different countries is highly correlated; for different products, whether the estimated elasticity is different in general; whether the market power depends on the size of a country. Through the verification, [1]it can be found that the result of the estimation of elasticity is reasonable. In addition, the data concerning the elasticity of supply and demand can be obtained in the process of estimating the reciprocal of elasticity of export supply. Based on this, the basic regression equation for estimation is as follows:

$$T_{ig}=\beta_i f(w_{ig})+\eta_{iG}+x_{ig}y+u_{ig}$$

Where, $w=1/e_s$, and i represents different countries; g represents different products and G represents different industries. Although the basic theory reveals linear relation, we will still take into account the different forms of the function f from the angle of empirical estimation. In addition, due to the individual characteristics of different countries, such as the location, the economic development level, institutional environment and industrial structures etc. factors which may affect the tariff, the influence of country and industry will be taken into account in the process of regression. In order to analyze the influence of country and industry, three methods are used: 1. Only the influence of country is counted, and the influence of industry is regarded as a part of the deviation, i.e. $\eta_{iG}=\eta_i+v_{iG}$. 2. Both the influence of country and the influence of industry are taken into account, i.e. $\eta_{iG}=\eta_i+\eta_G+v_{iG}$. 3. In the two factors (country and industry), the factor which is relatively important is taken into account. According to the above presumptions, OLS and Toibit are used to make estimation on the models ($x_{ig}y$ is not taken into account temporarily). The folder named “replication” in the root directory of disc C shall be presumed to be the data and program folder, and the data file shall be named “baseline_table.dta”. “baseline.do” is presumed to be the “do” file of estimation model. The core program language is as below:

```
cd "C:\replication\"
use "baseline_table", clear
qui tab section, gen(secdum)
set more off
qui reg avgtariff inv_exp_el avgtar_* , r nocons outreg inv_exp_el avgtar_* using table.out, replace nolabel
qui reg avgtariff mid_hi avgtar_* , r nocons
outreg mid_hi avgtar_* using table.out, append nolabel
qui reg avgtariff lninv avgtar_* , r nocons
outreg lninv avgtar_* using table.out, append nolabel
sum avgtariff
```

[1] For specific method, please refer to the part related to elasticity estimation in Christian et al (2008).

```
sum avgtariff if section==9
rename secdum9 exc_secdum9
qui       reg avgtariff inv_exp_el    avgtar_* secdum* , rnocons
outreg inv_exp_el    avgtar_* using table.out, append nolabel
qui       reg avgtariff mid_hi avgtar_* secdum* , r nocons
outreg mid_hi    avgtar_* using table.out, append nolabel
qui       reg avgtariff lninv avgtar_* secdum* , r nocons
outreg lninv    avgtar_* using table.out, append nolabel
qui tobit avgtariff lninv avgtar_* secdum* , nocons ll(0)
qui    do threshold_tab_7.do
qui    reg avgtariff mid hi avgtar_* secdum* , r nocons
outreg mid hi    avgtar_* using table_7.out, append nolabel
```

The folder named “replication” in the root directory of disc C shall be presumed to be the folder which includes storage program file and data file, and the data file .dta shall be named “baseline_table”, and the .ado file, used to output the estimation result shall be named “outreg”. The variable “section” represents the industries categorized according to HS2 standard (21 categories in total, i.e. 21 industries), while “avgtariff” indicates the average tariff of the products categorized according to HS4 standard, i.e. t value. “inv_exp_el” represents the reciprocal of elasticity of export supply, i.e. w value. “avgtar_*” denotes the dummy variable of a country (15 countries in total). For instance, in the data sheet, the name of the dummy variable of Argentina is avgtar_ALG. mid_hi is the target variables which are relatively high and accounts for 66.67% of w value, and mid is the target variables which are in the middle and accounts for 33.33% of w value, while hi is the target variables which are the highest and accounts for 33.34% of w value. “lninv” is the variable obtained by taking the logarithm of “inv_exp_el”. In the actural operation, the file name and the variables can be adjusted according to the actual needs.

In the above program language, the first two lines are to set up the route and open the data file, and the data file can also be opened by inputting the detailed route of the data file. The third line is to generate dummy variables for the 21 categories. The line started with qui is the estimation for models. In precedence order, the first three models only take the influence of country into account, and the last six models take the influence of both country and industry into account.

Notice: “rename secdum9 exc_secdum9” indicates that the dummy variables of the industries in the ninth category are renamed “exc_secdum9”, while in following lines below this line, in the program language of estimation models, “secdum*” is used to represent the dummy variables of the industries. Since the dummy variables of the industries in the ninth category have been renamed, they will not be included in “secdum*”. It denotes that in the following models, the ninth category is not excluded, i.e. wood and wood products, charcoal, softwood and softwood products, straw, straw stalk and stipa etc. products in these industries are not included. The line started with “outreg” indicates that the results are input into the file named “table.out”. “.out” file can be opened by the program of notepad, and can be easily copied to excel files.

Additionally, the program language “qui do threshold.do” means that the “do” program file named “threshold” is run to estimate the eighth model, which is a threshold model. The core program language of “threshold.do” is:

```
local i = 1
local m = 0
scalar rss`m' = 5000000
while `i' < 100 {capture drop inv_exp_hi_int inv_exp_not_hi_slop inv_exp_hi_slop egen perc`i'= pctile(inv_exp_elas) , p(`i')
gen inv_exp_not_hi_slop= 0
gen inv_exp_hi_slop = 0
gen inv_exp_hi_int=0
replace inv_exp_not_hi = inv_exp_elas if inv_exp_elas < perc`i'
replace inv_exp_hi = inv_exp_elas if inv_exp_elas > perc`i'
replace inv_exp_hi_int=1 if inv_exp_elas > perc`i'
```

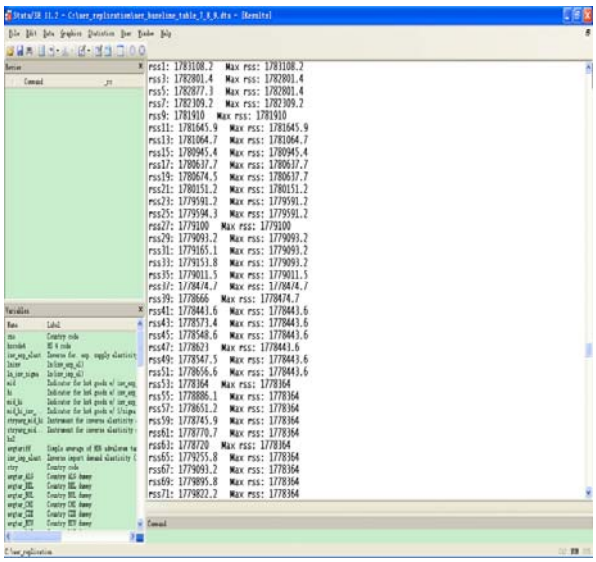
```
reg avgtariff inv_exp_not_hi inv_exp_hi inv_exp_hi_int    avgtar_* secdum*, r nocons
scalar rss`i' = e(rss)
if rss`i' < rss`m'    {local m = `i'    scalar rss`m' = rss`i'}
else    display "rss`i': " rss`i' "      Max rss: " rss`m'
local i = `i' + 2 }
replace       inv_exp_not_hi_slop = 0
replace inv_exp_not_hi_slop = inv_exp_elas if inv_exp_elas < perc`m'
replace inv_exp_hi_slop    = 0
qui replace    inv_exp_hi_slop    = inv_exp_elas if inv_exp_elas > perc`m'
replace inv_exp_hi_int= 0
replace    inv_exp_hi_int=1 if inv_exp_elas > perc`m'
display "threshold percentile: `m'"
reg avgtariff inv_exp_hi_int inv_exp_not_hi inv_exp_hi    avgtar_* secdum*, r nocons
```

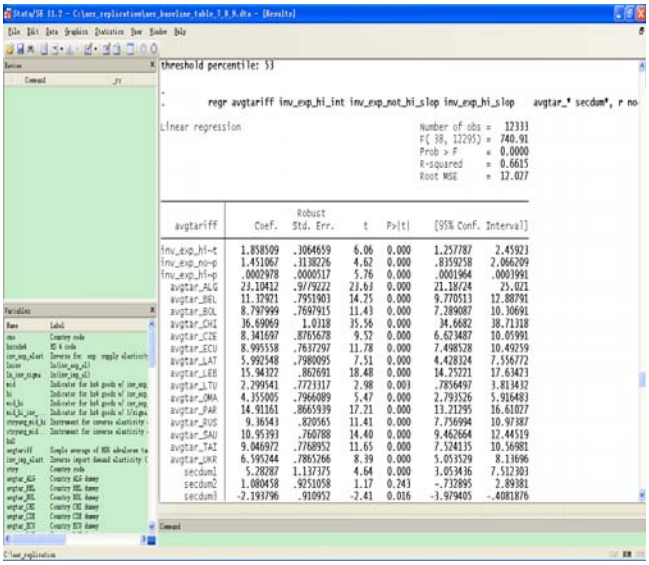
The sixth line in the program language of threshold.do indicates that the first 1% value of the variable “inv_exp_elas” shall be the starting point to search threshold. The tenth line means that when the observed value of “inv_exp_elas” is smaller than the current threshold, the new variable “inv_exp_hi” shall be given a new observed value of “ inv_exp_elas”. The twelfth line means that the dummy variable “inv_exp_hi_int” shall be set up, and the value shall be 1 when the observed value of “inv_exp_elas” is bigger than the current threshold. The thirteenth line indicates that regression shall be done under current threshold. The fourteenth line means that the regression sum of squares is assigned to the scalar rss1. Paragraph of program language：

```
scalar rss`i' = e(rss)
if rss`i' < rss`m'    {local m = `i'    scalar rss`m' = rss`i'}
else
display "rss`i': " rss`i' "      Max rss: " rss`m'
local i = `i' + 2
```

Compare rss1and rss0: if rss1is smaller than rss0, set “rss0=rss1”; if rss1is bigger than rss0, there shall be no change, and the value of rss1 and rss0 shall be displayed in the program window. After the first circulation is finished, the second one shall be started, and in each circulation, add 2 to the temporary variable i.

The most optimal threshold point shall be found during the circulation between the eighth line (counted backwards) and the second line (counted backwards), and the value of the inv_exp_not_hi, inv_exp_hi and inv_exp_elas shall be re-assigned according to the optimal threshold, and the m value that corresponds to the optimal threshold shall be displayed in the program window. The last line is the regression under the optimal regression. Run this program separately in stata11, and the window displays as below:

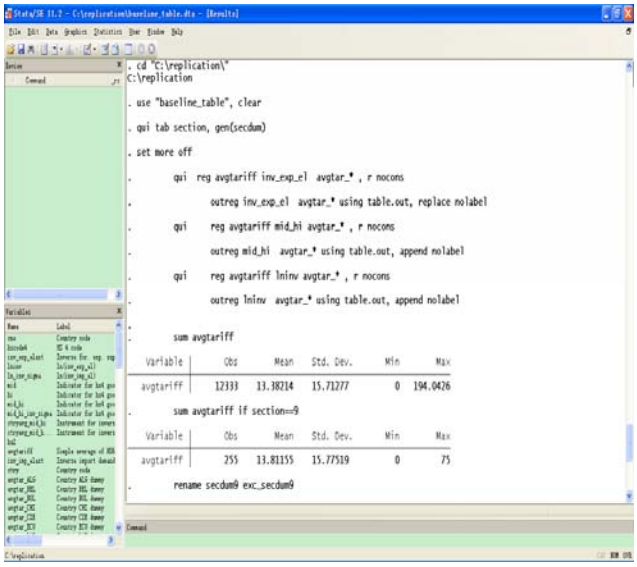


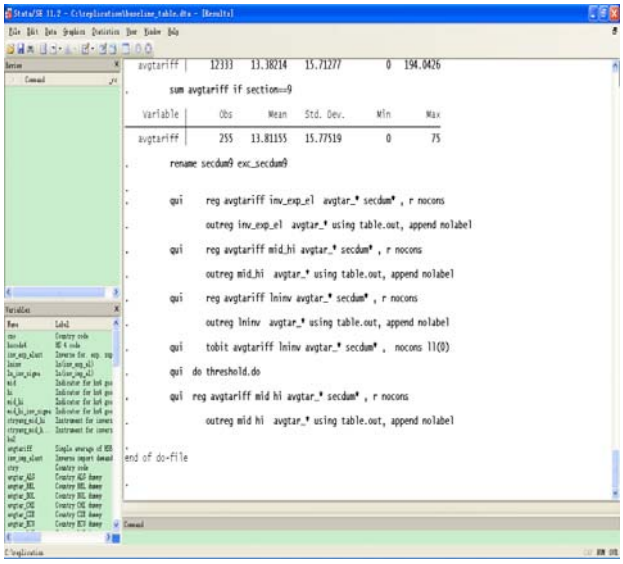


Through running “threshold.do” program separately, the threshold can be assigned at 53% of the flexible data, and the value is 1.82432.

To finish the model estimation, the two “do” files, baseline.do and threshold.do shall be compiled in advance. This can be realized by opening the “do” file compiler in stata and input the above program language in the corresponding “do” file. After the program languages for the two files are done, click the shortcut key used to run the program directly to run “baseline.do” in the window of “do” file compiler, or input “do baseline.do” in the program naming window and then click enter key. If the directory of program file is not consistent with the current directory, the current directory can be changed or the complete path name of the file “baseline.do” can be input.

Run “baseline.do”, and the below result is displayed:





After finish running the “baseline.do” program file, we can get the result output file “table.out”. The following result (Table2) can be obtained after the data in this file is put in order:

Table2 The result for the “baseline.do” program

Fixed Effect	Country			Country and industry					
Model No.	1	2	3	4	5	6	7	8	9
Estimation Method	ols	ols	ols	ols	ols	ols	tobit	threshold	ols
inv_exp_elast	0.0003			0.0004					
	[0.0001]			[0.0000]					
mid_hi		1.239			1.464				
		[0.248]			[0.242]				
mid									1.556
									[0.276]
hi									1.367
									[0.277]
inv_exp_not_hi								1.4511	
								[0.3138]	
inv_exp_hi								0.0003	
								[0.0001]	
inv_exp_hi_int								1.8585	
								[0.3065]	
lninv			0.121			0.169	0.167		
			[0.042]			[0.041]	[0.045]		
avgtar_ALG	23.8027	23.01	23.659	24.613	23.651	24.385	24.373	23.1041	23.655
	[0.6407]	[0.657]	[0.642]	[0.9513]	[0.962]	[0.952]	[0.931]	[0.9779]	[0.962]
avgtar_BEL	12.3766	11.573	12.296	12.69	11.699	12.543	12.433	11.3292	11.704
	[0.2956]	[0.337]	[0.297]	[0.7641]	[0.783]	[0.765]	[0.940]	[0.7952]	[0.783]
avgtar_BOL	9.7645	8.967	9.65	10.17	9.221	9.994	9.959	8.798	9.225
	[0.0309]	[0.171]	[0.061]	[0.7362]	[0.756]	[0.739]	[0.958]	[0.7698]	[0.756]
avgtar_CHI	37.8788	37.078	37.738	38.253	37.241	38.005	37.964	36.6907	37.246
	[0.7743]	[0.798]	[0.779]	[0.9874]	[1.010]	[0.993]	[0.890]	[1.0318]	[1.010]
avgtar_CZE	9.461	8.651	9.399	9.7441	8.725	9.612	8.773	8.3417	8.73
	[0.5353]	[0.538]	[0.532]	[0.8546]	[0.867]	[0.855]	[0.894]	[0.8766]	[0.867]
avgtar_ECU	9.8103	9.005	9.732	10.367	9.394	10.235	10.195	8.9956	9.397
	[0.1996]	[0.260]	[0.203]	[0.7300]	[0.749]	[0.732]	[0.935]	[0.7637]	[0.749]
avgtar_LAT	7.2623	6.449	7.243	7.2811	6.277	7.22	6.907	5.9925	6.281
	[0.3556]	[0.404]	[0.357]	[0.7632]	[0.785]	[0.764]	[0.910]	[0.7980]	[0.785]
avgtar_LEB	17.1013	16.29	17.089	17.144	16.151	17.099	17.079	15.9432	16.155
	[0.5310]	[0.564]	[0.531]	[0.8444]	[0.860]	[0.844]	[0.923]	[0.8627]	[0.860]
avgtar_LTU	3.6174	2.808	3.571	3.6146	2.616	3.514	-6.061	2.2995	2.621
	[0.2611]	[0.316]	[0.264]	[0.7404]	[0.760]	[0.742]	[0.982]	[0.7723]	[0.760]
avgtar_OMA	5.6223	4.858	5.608	5.6929	4.755	5.633	4.891	4.355	4.759
	[0.3448]	[0.377]	[0.349]	[0.7775]	[0.790]	[0.779]	[0.949]	[0.7966]	[0.790]
avgtar_PAR	16.0925	15.304	15.931	16.339	15.421	16.108	15.996	14.9116	15.424
	[0.4983]	[0.525]	[0.502]	[0.8402]	[0.853]	[0.842]	[0.996]	[0.8666]	[0.853]
avgtar_RUS	10.6562	9.841	10.555	10.891	9.873	10.709	9.987	9.3654	9.878
	[0.3435]	[0.385]	[0.347]	[0.7754]	[0.799]	[0.779]	[0.898]	[0.8206]	[0.799]
avgtar_SAU	12.1274	11.319	12.021	12.455	11.446	12.264	12.162	10.9539	11.451
	[0.0821]	[0.185]	[0.093]	[0.7194]	[0.740]	[0.722]	[0.895]	[0.7608]	[0.740]
avgtar_TAI	9.6905	8.89	9.612	10.318	9.338	10.18	9.712	9.047	9.341
	[0.2862]	[0.335]	[0.289]	[0.7482]	[0.768]	[0.750]	[0.911]	[0.7769]	[0.768]
avgtar_UKR	7.357	6.554	7.233	8.0824	7.114	7.887	6.823	6.5952	7.118
	[0.2832]	[0.335]	[0.290]	[0.7460]	[0.767]	[0.749]	[0.933]	[0.7865]	[0.767]
Observations	12333	12333	12333	12333	12333	12333	12333	12333	12333
R-squared	0.61	0.61	0.61	0.66	0.66	0.66		0.66	0.66
parameters	16	16	16	36	35	36	35	38	36
Adj R2	0.61	0.61	0.61	0.66	0.66	0.66		0.66	0.66

Note: The data in the bracket is the standard error.

From the final estimation result of program running, it can be seen that the effect of the first three models are almost consistent with the effect of the forth, the fifth and the sixth models, so the analysis is made according to the latter. It can be found from model 4 the market power has obvious positive influence on the tariff. However, due to the abnormality of the elasticity value of export supply, the value of the coefficient is relatively low. When the market power is relatively high, and the tariff will be very high, what’s more, the marginal effect to increase market power shall be small. This can be seen from the model 8. Model 5 indicates that the products with relatively big reciprocal of elasticity of export supply need to pay relatively high tariff. In addition, from the

model 9, we can see that such kind of effect has similar influence on both the relatively high market power and relatively low market power. In model 6, according to the semilogarithmic model estimation, market power still has obvious positive effect on tariff. The result is quite close to the result obtained from tobit model estimation of model 7.

However, the basic OLS estimation may be the biased estimation due to the measurement error and negligence of variables. Since the purpose of this case is to define and measure the influence of market power on tariff, it is considered that the estimation based on instrument variable shall be introduced.

Method of introducing instrument variable: the market power on a certain product of a country is described by the market power of the other 14 countries on the same product. In the elasticity analysis, the market power on different products in different countries is divided into three types: low, medium and high. Therefore, the instrument variable of $f(w_{ig})$ shall be the average of the classified variables of the other 14 countries, and the name of its variable is “ctryavg_mid_hi” in the datasheet. Similar to the basic models, there are following conditions: only the influence of countries is taken into account; both the influence of countries and the influence of industries are taken into account. Additionally, considerations are also given to the influence of different countries at the level of industry, and three functional forms of f are taken into consideration in each condition. The program file used to estimate those models shall be named “instru.do” , and the core program language is:

```
use "baseline_table", clear
qui tab section, gen(secdum)
sum avgtariff
sum avgtariff if section==9
rename secdum9 exc_secdum9
qui ivregress gmm avgtariff avgtar_*    (inv_exp_el= ctryavg_mid_hi) , first r nocons
outreg inv_exp_el using table_2.out, replace nolabel
qui ivregress gmm avgtariff avgtar_*    (mid_hi= ctryavg_mid_hi) , first robust noconstant
outreg mid_hi    using table_2.out, append nolabel
qui ivregress gmm avgtariff avgtar_* (lninv =ctryavg_mid_hi), first robust noconstant
outreg lninv using table_2.out, append nolabel
qui ivregress gmm avgtariff avgtar_* secdum* (inv_exp_el= ctryavg_mid_hi) , first r nocons
outreg inv_exp_el using table_2.out, append nolabel
qui ivregress gmm avgtariff avgtar_* secdum* (mid_hi= ctryavg_mid_hi) , first robust
noconstant
outreg mid_hi using table_2.out, append nolabel
qui ivregress gmm avgtariff avgtar_* secdum* (lninv =ctryavg_mid_hi), first robust noconstant
outreg lninv using table_2.out, append nolabel
char section[omit] 9
char cno[omit] 39
qui xi: ivregress gmm avgtariff i.section*i.cno (inv_exp_el= ctryavg_mid_hi) , first r
outreg inv_exp_el using table_2.out, append nolabel
qui xi: ivregress gmm avgtariff i.section*i.cno (mid_hi= ctryavg_mid_hi) , first robust
outreg mid_hi using table_2.out, append nolabel
qui xi: ivregress gmm avgtariff i.section*i.cno (lninv =ctryavg_mid_hi), first robust
outreg lninv using table_2.out, append nolabel
```

Set up “instru.do” file in stata and input the above program language, and then running the file.

There is estimation for nine models in the file “instru.do”, and the result is output to the file “table_2.out”. The following result (Table3) can be obtained after the data in this file is put in order.

In the program language of “instru.do:

```
char section[omit] 9
char cno[omit] 39
```

Table3 The estimate results for nine models

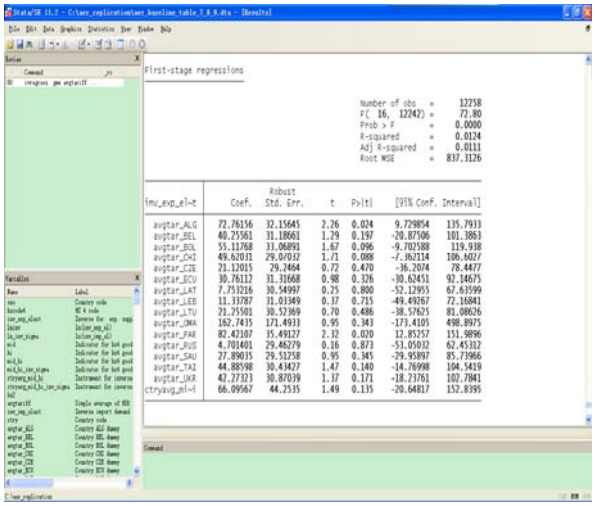
Fixed Effect	Country			Country and industry			Difference between nations at the level of industry		
Estimation Method	GMM	GMM	GMM	GMM	GMM	GMM	GMM	GMM	GMM
Model No.	1	2	3	4	5	6	7	8	9
inv_exp_elast	0.04			0.089			0.075		
	-1.45			-1.62			(2.69)**		
mid_hi		3.959			8.884			9.016	
		(5.18)**			(7.51)**			(8.36)**	
lninv			0.751			1.705			1.764
			(5.16)**			(7.38)**			(8.34)**
Observations	12258	12258	12258	12258	12258	12258	12258	12258	12258
parameters	16	16	16	35	35	35	284	282	283
1st stage F	5	1649	1335	2	653	517	3	691	544

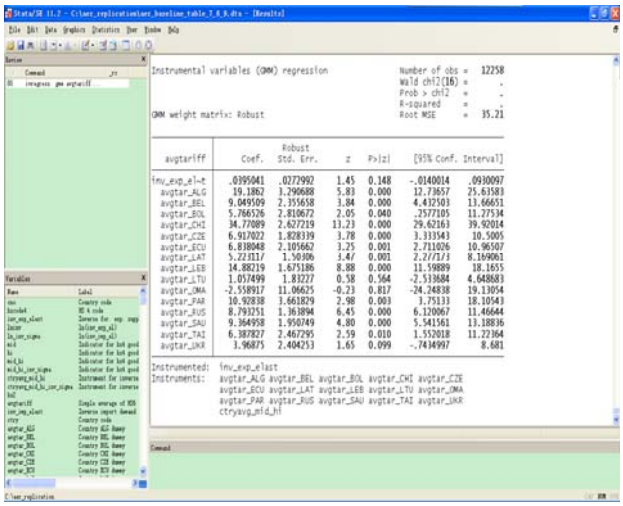
Note: the data inside the bracket is the value of t. * means it is obvious at the level of 5%, and ** means it is obvious at the level of 1%.

The industries in the ninth category and the No. 39 country (i.e. White Russia) are defined as the basic category respectively.

Due to the limit of space, only the conditions of the estimation for the most important variables are listed in the above table. In order to know the specific results of the regression of each model, the program language that corresponds to the estimation in “instru.do” can be input separately in input window of the stata program language and executed.

Take model 1 as sample. Open baseline_table.dta, and run qui tab section, gen(secdum). Input “ivregress gmm avgtariff avgtar_* (inv_exp_el= ctryavg_mid_hi) , first r nocons” in the program language input window. The following result is obtained after the enter key is clicked:





The operation of the estimation for the other models is similar to this one.

After the instrument variable is counted in, the positive influence of market power is greater than the result obtained in OLS estimation. For example, the coefficient is 1.7 (model 6 and model 9) under the condition that both the influence of country and the influence of industry are taken into account. It is 10 times larger than the result obtained from OLS estimation, and it is obvious at the level of 1%. Similar conclusion can be drawn in the estimation for the dummy variables of elasticity (model 5 and model 8). Moreover, it can be found that if a country has middle level or high level market power on a product, the tariff level will be increased by 9%. This verifies and supports the correlation theory for the relation between market power and tariff.

Although the factor of the difference between countries is taken into account in the foregoing model, the decisive factors that can affect the tariff of a country may be discussed separately. The above-mentioned basic theory will be tested here to see whether it is still valid in the separate analysis of a country. For the countries discussed above, the instrumental variable GMM is used separately to estimate the decisive factors that affect the tariff with a view to the factor of industries. In the 15 countries discussed in the above text, during the observation period of samples, there is almost no difference between Bolivia, Oman and Saudi Arabia in terms of tariff rate of all walks of life, so these three countries will not be discussed for the moment. In the above description of the

statistical features of the data in the case of this paper, it has been mentioned that the abnormal observed value (outliers) may affect the validity of the model estimation. In the general analysis, small quantity of outliers will not have obvious effect on the estimation results. However, in the analysis of each country, the influence of the outliers can not be neglected. In order that the estimation results will not be affected by those outliers, in the language of program file, these outliers will be deleted.

In order to make it easy to compare with the general conditions, in the estimation for a single country, estimation will also be made for the results corresponding to all the observed values (exact the same as model 9 in table 2, because model 9 is the analysis of the difference between countries at the level of industry) and the results obtained after Bolivia, Oman and Saudi Arabia are deleted. Presume the name of program used to estimate each country with consideration to the factor of industry is inside.do, and the core program statement is as below:

```
use "C:\replication\baseline_table.dta", clear
qui tab section, gen(secdum)
set more off
qui sum avgtariff if section==9
rename secdum9 exc_secdum9
char cno[omit] 39
qui xi: ivregress gmm avgtariff i.section*i.cno (lninv =ctryavg_mid_hi), first robust
outreg lninv using table_3.out, replace nolabel
qui xi: ivregress gmm avgtariff i.section*i.cno (lninv =ctryavg_mid_hi) if cno!=44 & cno!=169 & cno!=186, first robust
outreg lninv using table_3.out, append nolabel
bysort cno: egen q=iqr(avgtariff)
bysort cno: egen q25_avgtariff=pctile(avgtariff), p(25)
bysort cno: egen q75_avgtariff=pctile(avgtariff), p(75)
gen low=q25-3*q
gen hi=q75+3*q
gen inside=1 if (avgtariff>low & avgtariff<hi)
local countryname CHI RUS TAI UKR CZE ALG BEL ECU PAR LTU LEB LAT
foreach j of local countryname {
qui ivregress gmm avgtariff  (lninv =ctryavg_mid_hi) secdum* if ctry=="`j'" & inside==1, first robust  outreg lninv using table_3.out, append nolabel }
```

In the above program language, the first paragraph is to set up the running environment for the program, and do some preparatory treatment for the data. If such treatment has been done in the estimation of basic models, the first paragraph can be deleted. If estimation is made directly on the conditions of each country without running “baseline.do” and “instru.do”, it is necessary to use the first paragraph.

The second paragraph is used to estimate the overall models (including the overall model in which the three countries are deleted). The program statement is consistent with the program statement in the corresponding model estimation in “instru.do”. When estimation is made on the general model in which the three countries are deleted, “if cno!=44 & cno!=169 & cno!=186” shall be used to delete Bolivia, Oman and Saudi Arabia, in which, cno is the name of the variable of country number.

The third paragraph is used to delete the outliers in the observed tariff values of each country. The basic method is: put the data in ascending order, and take the value between the result of the first quartile deducting three times quartile deviation and the result of the third quartile adding three times quartile deviation, i.e. take the value at the boundary point of the first 1/4 observed values and the value at the boundary point of the first 3/4 observed values, and then calculate the difference (positive number), and then deduct three times of the difference from the value at the boundary point of the first 1/4 observed values, and the result will be taken as the lower limit for data selection; then adding three times of the difference to the value at the boundary point of the first 3/4 observed values, and take the result as the upper limit for data selection. The function of the

program instruction iqr is to take the value of quartile deviation. “pctile(avgtariff), p(25)” means the value of the first quartile is found, and “pctile(avgtariff), p(75)” means the value of the third quartile is found. The variable “inside” is used to differentiate the normal observed value and the abnormal observed value (outliers). If the value of “inside” is 1, it means the observed value is normal.

The last paragraph is used to estimate the model of each country (the order is based on the size of GDP), and the result is output to “table_3.out”. The statement of loop computation in the last paragraph can be obtained by writing the regression statement of each country. That is to say, the last paragraph can achieve the same purpose with the following paragraph:

qui ivregress gmm avgtariff (lninv =ctryavg_mid_hi) secdum* if avgtar_CHI==1& inside==1, first robust

outreg lninv using table_3.out, append nolabel

……

qui ivregress gmm avgtariff (lninv =ctryavg_mid_hi) secdum* if avgtar_ LAT==1& inside==1, first robust

outreg lninv using table_3.out, append nolabel

In the process of programming, only the statement used to estimate the model of the first country need to be written, and then the function of loop computation can be realized easily and quickly by means of copying and pasting and then changing the name of the variable “country”. This method will have no impact on the programming efficiency when the loop is small, but if the loop is big or there are many countries, this method will affect the programming efficiency. It depends on the actual conditions in practical operation.

After the above program language is written into the “inside.do”, set up the running environment of the program and run “inside.do”.

The below table (Table4) can be obtained after the data in “table_3.out” is put in order:

Table4 The estimate results for every country

Model No.	1	2	3	4	5	6	7
Sample	all	exceptthe3	CHI	RUS	TAI	UKR	CZE
Constant	14.965	15.197	22.677	13.728	-0.157	0.647	5.25
	(8.15)**	(7.50)**	(3.60)**	(6.73)**	-0.07	-0.86	(5.51)**
lninv	1.764	2.113	7.597	2.419	1.98	0.708	0.16
	(8.34)**	(8.49)**	(4.26)**	(3.99)**	(2.54)*	(2.68)**	-0.67
Observations	12258	9952	1089	1021	841	685	1000
Elasticity of tariff	0.13	0.15	0.2	0.23	0.22	0.12	0.03
Model No.	8	9	10	11	12	13	14
Sample	ALG	BEL	ECU	PAR	LTU	LEB	LAT
Constant	38.38	17.396	11.536	17.265	5.557	12.935	1.846
	(8.36)**	(10.49)**	(7.46)**	(3.08)**	(4.80)**	(5.52)**	-1.14
lninv	5.402	2.28	1.55	2.443	0.826	2.406	1.408
	(5.60)**	(4.41)**	(4.59)**	(3.54)**	(3.06)**	(4.47)**	(2.35)*
Observations	739	703	753	510	768	754	868
Elasticity of tariff	0.23	0.18	0.16	0.15	0.36	0.16	0.2

Note: The data in the bracket is t value. * means it is obvious at the level of 5%, and ** means it is obvious at the level of 1%. Tariff elasticity is obtained by the coefficient of each sample model lninv divided by average of tariff (%).

From the estimation result, it can be seen that the coefficients of the corresponding variables of the other 11 countries, except Czech are obviously positive. Czech joined GATT in 1992, and its tariff is the same with that in Slovak (the allies of Czech), while Slovak is the member of GATT, so the tariff of Czech can not reflect the motivation of the trading terms, and therefore, it is not obvious. If the model estimation is made once more after Czech is deleted, there will be no obvious influence on the results. In the above model estimation, the coefficient of each country is distributed at both sides of 1.76 and 2.11 (coefficient of general model). The mean and midpoint are quite close to this area. If we see the result by country, we can find that the lninv coefficient of China is the biggest. Generally, the countries and regions with bigger economic scale will have bigger lninv coefficient. This conclusion is basically consistent with the above conclusion that the stronger the market power is, the higher the tariff is. In addition, in view of elasticity of tariff, the difference between countries is small also comply with the theory. For the economic body with expected higher market power, the tariff has already been very high, and the marginal effect to increase the tariff more is small, so the elasticity of tariff has no big difference with that of smaller economic body. Therefore, in view of the

conditions of different countries, there may be certain difference in estimation results, but the core conclusion is consistent.

Additionally, the data in this paper can also be used to discuss the influence of the government's motive for tariff and the lobbying ability of the industries on the taxation. The readers who are interested in it can refer to the original text of Christian et al.

Acknowledgement

This article is sponsored by the central university basic scientific research business expenses special fund and Wuhan university common courses 2011(201119). Thank you so much for all your supports.

References

[1] Christian Broda, Nuno Limao, David Weinstein. *Optimal Tariffs and Market Power: The Evidence.* The American Economic Review , 2008, 98(5)

[2] Robert C. Feenstra & Clinton R. Shiells，*Bias in U.S. Import Prices and Demand.* National Bureau of Economic Research, 1994 - 30

[3] T. Hitiris; E. Petoussis。*Price and tariff effects in a dynamic specification of the demand for imports* .Applied Economics，1984-02

[4] Christopher F. Baum. *An Introduction to Modern Econometrics Using Stata.* Stata Press, 2006

[5] Alan C. Acock. *A Gentle Introduction to Stata.* Stata Press, 3rd Revised edition,2012