

# Final Paper: Democracy and Trade Policy in the Developing Countries

*Lucie Lu*

*May 15, 2018*

Since the 1980s, economic liberalization has advanced remarkably. There has been a marked shift in the orientation of the trade and industrial policies for developing countries. Many developing countries moved away from a heavy reliance on direct intervention and inward-looking trade regimes toward less controlled and more export-oriented trade regimes. Simply put, they chose to liberalize their trade regimes. Many developing countries have decided to integrate their economies into a global economy by dismantling their protectionist trade policies. Countries such as Argentina, Brazil, Indonesia, Mexico have conducted substantial trade reforms respectively, including reducing tariff rates, reforming import licensing procedures in the 1980s. What drove such a substantial change in their trade policy?

What has driven the trend in trade liberalization? For the developing countries, it is costly for them to abandon the protectionist, import-substituting industrialization strategy given that they have been the orthodoxy among developing country policymakers in the post-war period. A temporary protection for the infant industry is sometimes necessary for infant industries to gain competitiveness in trade in the developing countries. However, over time, many infant industries refuse to mature but become inefficient, and what is worse, they hold vested interests of the oligarchies. In this case, interest groups and political leaders usually benefit from the status quo of the protectionist policies. Change in the protectionist status quo is unexpected. This makes the question puzzling. If the political leaders benefit from the status quo, why do they conduct reforms? Why would they choose to lower the trade barriers?

## Argument

Rodrick (1994) argues that it is very difficult for politicians to conduct liberalization reforms in normal times, but times of crises can provide an opportunity moment for undertaking structural reforms. What are the ‘times of crises’ to drive the structural change? Some scholars claim the economic crises may force countries to give up the protectionist old policies and move toward a market-friendly and open one. Crises may underscore the problems of the old protectionist view of a country, so countries conduct reform and liberalize to recover from the crises (Krugers, 1997). Others claim that external pressure such as a country’s accession to the international institutions helps facilitate the transition. External pressures such as the World Trade Organization (WTO), IMF or trade treaties, matter (Vreeland, 2007). When developing countries join these international institutions or sign the binding international trade agreements with other developed countries, the newly joined members will show their commitment to the international regimes (Baccini & Urpelainen, 2014).

A voluntary tariff reduction in a developing country will send a promising signal to the international community.

Now, what about political reforms? Another closely related structural change, the trend of democratization, was happening a decade before the trend to free trade. In 1975, there were around thirty democracies in the world, while by 1992, there were about eight-nine. Are these two trends causally related or happened to coincide by chance? When countries are more democratized, are they more likely to initiate trade liberalizing economic reforms? Are sharp changes in the economic reform the product of political transformation? As Rodrik argued, "Not all political transformations result in trade reform, but sharp changes in trade policy are typically the result of such transformation (1994, p. 69)." This article largely focuses on the regime type. To show that the regime type places a role in the move to free trade, I will also control for the factors of economic crisis and external pressure illustrated above.

What remains unclear is whether political reforms affect trade policy reforms, or whether trade reforms induce political transition. The research design of observational data plus classic regression model cannot fully address this endogeneity issue. The fundamental problem of drawing causal inference in an observational data is to make the right comparison. To compare apples to apples, we need to hold other things equal. Because in observational studies, treatments are observed rather than assigned, so it is not possible to consider the observed data under different treatment are randomly assigned. It is highly possible that the two groups of countries (one with democratization, the other without) are fundamentally different in many ways. All of the potential confounders (whether observed or unobserved) will obscure the causal relationship we try to infer. Milner and Kubota's article (2005), "Why the Move to Free Trade? Democracy and Trade Policy in Developing Countries" is a classic piece. They provide a systematic investigation on the relationship between regime change and trade liberalization. They have suggested that democracy has led to trade liberalization. I will cast doubt on their analysis strategy that has relied on controlling for confounding covariates through linear regression. I suggest that a more sophisticated research design, matching can to some extent solve the selection bias problem in the observational data.

I am going to replicate and extend the paper on Milner and Kubota's argument: When countries become more democratic, they are more willing to open their markets to the international economy. I have a couple of concerns on their original work. They present the average treatment effect of democracy on average tariff level in the Ordinary Least Squares (OLS) model, which at best only presents the two variables are positively correlated. Given the limitation of observational data, it is not convincing to draw a causal relationship between the two. To extend their original piece, inspired by Angrist and Pischke (2015), I will conduct a regression with matching approach to overcome the possible selection problem in the observational data. The goal of this replication paper is to present a regression-based causal story: countries that have democratized will tend to lower their levels of trade barriers.

The paper will be divided into three main sections. The first section will briefly describe the variables, measures, the reasons for adjustment. The second and third section present discussion of two methods of statistical inference and assessment.

[1] "C:/Users/Lucie Lu/Documents/531-explorations/Finalpaper/Data summary"

## Section I: Data on Trade Liberalization and Democratization

The data that I use in the following analysis is a data set compiled by Milner and Kubota (2005). It is a time-series cross-sectional (TSCS) data, covering 81 developing countries from 1980 to 1999. It include various measures not only on trade policy but also of a series of political regime measures, economic development measures and so on. The central hypothesis is that an increase in democracy will lead to a reduction in trade barriers, *ceteris paribus*.

The data Milner and Kubota have collected on the developing countries demonstrate a significant change in the measure of trade policy. *Tariff*, a country's unweighted average statutory tariff rate collected, is a key measure of a country's trade policy in year  $t$ . The historical trend shows that a dramatic decline in the average tariff level across countries has happened from 1982 to 1999. Statutory tariff rate has decreased by over sixty percent, from an average about 30 percent in 1980 to around 12 percent in 1999 in the developing countries.

More evidence of trade liberalization comes from a second measure of trade policy: the dichotomous categorization of countries into open and closed trade regimes. The data covers ninety developing countries from 1980 to 1999. In 1980, only 14 countries were scored as openness, while in 1999, 62 countries were open economies. The number of a country classified as open economy has increased by three times.

Evidence of democratization among the developing countries is also plentiful. To measure regime type, the paper uses the 21-Point Polity Score, ranging from -10 for a highly autocratic state to 10 for a highly democratic state. The average Polity IV score for about 110 developing countries that fell from around -3.4 to -4.7 and increased sharply to a score of 1.8 in 1999. It suggests that although many regimes on averages with mixes of democratic and autocratic features (-5 to 5), the wave of democratization has been evident since the mid-1970s. Similarly, the dichotomous regime shows how the number of democracies has increased over time. The series show that the process of democratization has started in the 1970s. In 1970, less than 20 developing countries were quantified as democracies. Over three decades, there has been a four-fold increase in the number of democracies among the developing countries.

## Section II: Empirical analysis: The OLS Models

### Variables and Measures

#### Dependent Variables

*Tariff* is a country's average statutory tariff rate between 1980 and 1999. However, this is poorly measured with a large number of missing data in the dataset. To be fair, it is notoriously difficult to measure and hard to find time-series and cross-sectional data on this measure.

#### Independent Variable

*Regime (Polity Score)* is a 21-point scale that measures variations both within democracies and among autocracies. This measurement comes from Polity III and Polity IV. It ranges

from -10 (highly autocratic) to 10 (highly democratic). I converted it as a dichotomous variable *Democracy* with countries scoring below 6 in the Polity Score as autocracies (=0) and those above 6 as democracies (=1). The hypothetical treatment here is “*democracy*.”

### Control Variables and Justification

*GDP per capita* is a lagged value of per capita real GDP based in 2005. The economic development is likely to affect its trade policy and its political regime change. More developed countries tend to be more economically and politically liberal. More developed countries tend to have smaller trade barriers.

*Economic Crisis* is an extreme event that occurs in that specific year rather than an economic variable with yearly change. To code that year with the economic crisis, one of the two conditions hold: either the country’s inflation rate is 40 percent or more, and it increases by 25 percent or more from the year before, or per capita GDP falls by 15 percent or more from the previous year (=1).

When countries experience crisis, the political leaders tend to conduct political and economic reforms to address the crisis. They are more likely to liberalize their trade regime because of a domestic demand or conditionality of foreign aid. They will tend to lower the tariff rates to facilitate trade. Under the pressure of economic crisis, politicians are more likely to undergo political reforms as well.

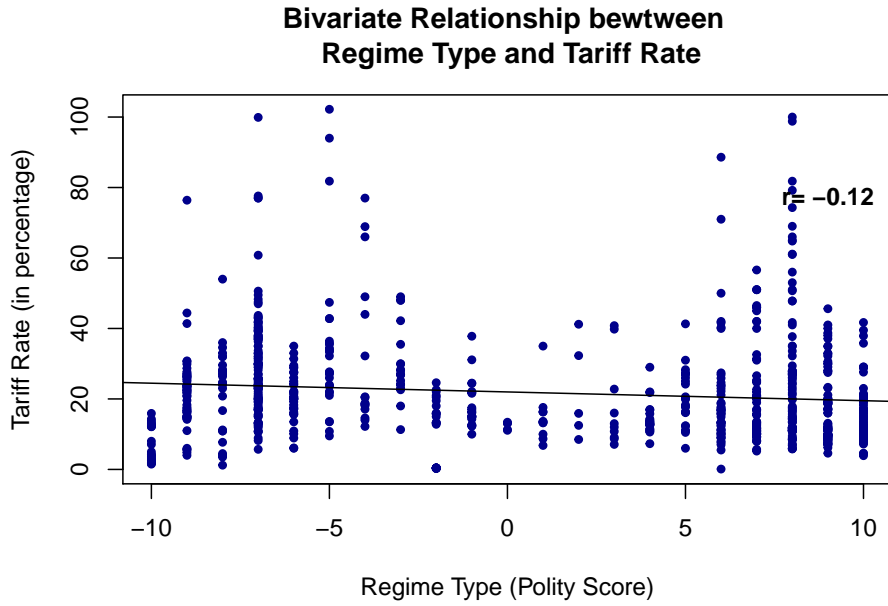
*GATT/WTO Member* is A lagged variable indicating whether a country is in the GATT/WTO(=1) or not. Joining a WTO induce member countries to lower their trade barriers. Also, WTO members tend to welcome new members intending to adopt liberal policies in both economic and political domains. Therefore, this controlled variable may have effects on both independent and dependent variables.

*Years of Office* counts the number of years a government has been in office. It ranges from 0 to 44, with the mean of 8.4. It is included in the control variable because the number of years a government in power will have an impact on both political and economic reforms. A new government might indicate a change in leadership and new idea.

The summary of statistics is included in the appendix.

### Data Structure

The correlation coefficient itself is simply a way to describe how the two variables vary together. Here, I first look at the bivariate relationship between the average tariff rate and polity score. The correlation coefficient  $r$  is -0.12. From Table 2, we can see there are negative relationships between the dependent variable, the independent variable and the control variables (except WTO members). It is noted that the covariate coefficients between the dependent variable, independent variable and cofounders are low. This correlation matrix provides evidence of a low possibility of the collinearity issues in the later regression analysis.

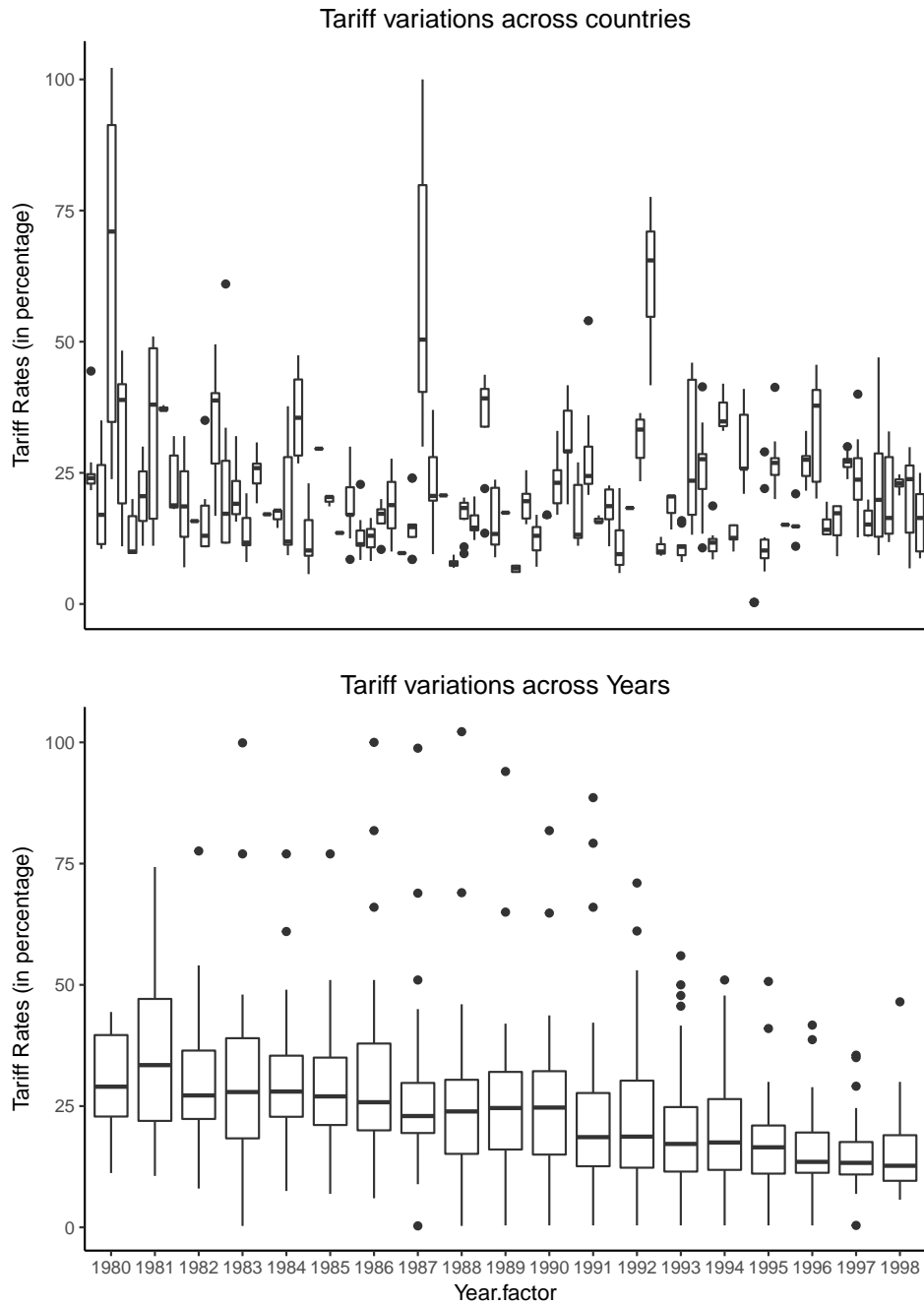


% latex table generated in R 3.4.3 by xtable 1.8-2 package % Wed May 16 00:32:51 2018

	Tariff	Regime	lnGDP	EC	WTO	YO
Tariff rate	1.00	-0.12	-0.45	-0.10	0.11	-0.07
Regime (Polity Score)		1.00	0.24	0.09	0.22	-0.39
GDP per capita (log)				-0.06		0.02
Economic Crisis				1.00	0.02	-0.08
GATT/WTO Member					1.00	-0.00
Years of Office						1.00

Table 1: Correlation Matirx: Correlations among Variables and Confounders

This figure tells me two important problems of the data structure in this dataset: 1) There are outliers in the data structure that we need to take them into account. Later on, I will use “lmrob” regression modelling strategy to handle unusual observations that do not follow the general trend of the dataset. 2) Variations of tariffs and polity scores across countries and across years are merged. For a single point on the figure, I can get the values of x (polity score) and y (tariff rate) of that point. However, I do not know that particular point stands for which year or/and which country. Before I proceed, I will examine whether variation across countries for a given year (*country differences*) or variation across years within a given country (*year differences*) is higher.

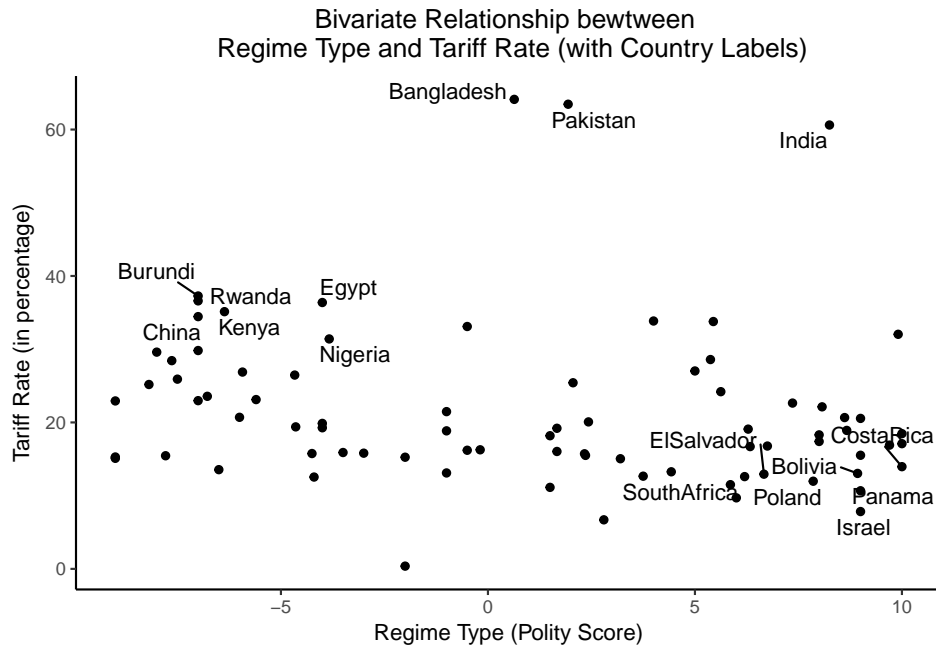


I calculated across-country variance (123.7135) and within-country-years variance (70.711). Across-country variance is around 1.7 times of the within-country-years variance. It means that differences across countries are higher than differences across years within each country. I also calculated an ICC score: It is a method of Simple Estimation of the Intraclass Correlation Coefficient. Interclass correlation is the relative values of the country- and

year-level variances (Gilman & Hill, p. 258). The variance within countries is quite high with ICC score 0.6306. In this case, the interclass correlation 0.6306 suggests that differences between across countries convey more information than differences across years within a country. This suggests that within a country, a country's tariff levels have not changed much across years. In contrast, tariffs variations across countries are high. The boxplots further demonstrate this point.

In the current dataset, the panel dataset includes the yearly tariff data for 81 developing countries over a 20-year period. Tariffs variations across countries are larger than variations across years within a country. Therefore, I collapse the dataset to 81 data points so I can know what the mean tariff rate (and other information) is for each country. Here, the country is the grouping variable.

Hence, I reduce the data to the cross-sectional dimension. I run my analysis based on time-series averages. My dependent variable is the time-series mean of tariff rate of the country  $i$  and my independent variable is the time-series mean of polity score. I also converted the time-series mean of polity score to a dichotomous variable, meaning that a developing country is democratic or non-democratic based on the polity scores on average over the 20-year period.



## Hypothesis

I have laid out the reasons to control for the three covariates: GDP per capita, economic crisis and WTO membership. The fourth one, "year of office" is added for additional check. Only covariates that meet the condition of affecting both the treatment and outcome variables confound the observed relationship between the two (Rubin, 1997; Achen, 1982). I hope that through the control of these relevant covariates, the treatment effect of democratization is

less cofounded. I lag all of the independent variables by one year. The basic OLS equation estimating the relationship between democracy and trade policy is:

$$trade_{policy}_{i,t} = \beta_0 + \beta_1 * Democracy_{i,t-1} + \beta_2 * GDP_{PC}_{i,t-1} + \beta_3 * ECCRISIS_{i,t-1} + \beta_4 * WTO_{i,t-1} + \beta_5 * OFFICE_{i,t-1} + \varepsilon_i$$

My main hypothesis is *democratic developing countries are oriented toward lower tariff rates than nondemocratic developing countries.*

I compare democratic developing countries and non-democratic developing countries in this model. The regime type, “democracy” is a hypothetical treatment effect. I concern what would happen to the average tariff rate of a given developing country as a result of a hypothesized “treatment (democracy)” in the observational studies. I understand this is not a randomized experiment at all, and the treatment and control groups in nature are not similar. I hope that I can to some extent address this problem through proper modelling and statistical adjustment.

Table 2 shows that compared to non-democratic countries, democratic countries on average tend to have lower tariff rates. In all of the regressions, the estimators of *democracy* have expected negative signs that support my main hypothesis. In equation (6), the main model here with three covariates, we know that the average tariff rates fall by 2 percent if a country changes from non-democracy to a democracy.

Table 2: Multiple Linear Regression Model: Results

	<i>Dependent variable:</i>					
	Tariff Rates					
	(1)	(2)	(3)	(4)	(5)	(6)
Democracy	-4.287 (2.681)	-1.748 (2.780)	-1.832 (2.760)	-1.882 (2.793)	-2.003 (3.087)	-1.992 (2.768)
GDP per capita		-2.746** (1.085)	-2.855*** (1.079)	-2.739** (1.088)	-2.712** (1.105)	-2.853** (1.080)
Economic Crisis			-10.600 (7.124)			-11.150 (7.158)
GATT/WTO Member				2.900 (3.745)		3.391 (3.724)
Years of Office					-0.040 (0.204)	
Constant	22.890*** (1.459)	41.020*** (7.300)	42.830*** (7.346)	38.600*** (7.954)	41.160*** (7.381)	40.110*** (7.940)
Observations	81	81	81	81	81	81
R <sup>2</sup>	0.031	0.105	0.130	0.112	0.105	0.139
Adjusted R <sup>2</sup>	0.019	0.082	0.096	0.077	0.070	0.094
Residual Std. Error	11.020	10.660	10.570	10.690	10.720	10.590
F Statistic	2.557	4.569**	3.831**	3.231**	3.021**	3.074**

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

In regression modelling, *lmrob* is a robust statistics, which could be considered to handle unusual observations and outliers that do not follow the general trend of the data set. ‘Robust’ means that the tests have good power for a wide variety of distributions even if some of the assumptions used to justify the estimation method are not applicable. The *lmrob* output from the *lmrob* function provides ‘robustness weights’ to give us information about



the outliers in the observations by evaluating their weights and summarize statistics from the remaining observations. It uses a fitting criterion that is not as vulnerable as least squares to unusual data. The estimators in the *lmrob* robust regression are weighted-least-squares estimates.

In all the models, the estimators' signs are negative, as expected. Compared to the lm models, the magnitudes of the treatment effects are higher. In equation (6), we can see that the average tariff rates fall by 3.5 percent if a country changes from non-democracy to a democracy. The estimators for GDP per capita and economic crisis also have the expected negative signs.

Table 3: Multiple Linear Regression Model (lmrobust): Results

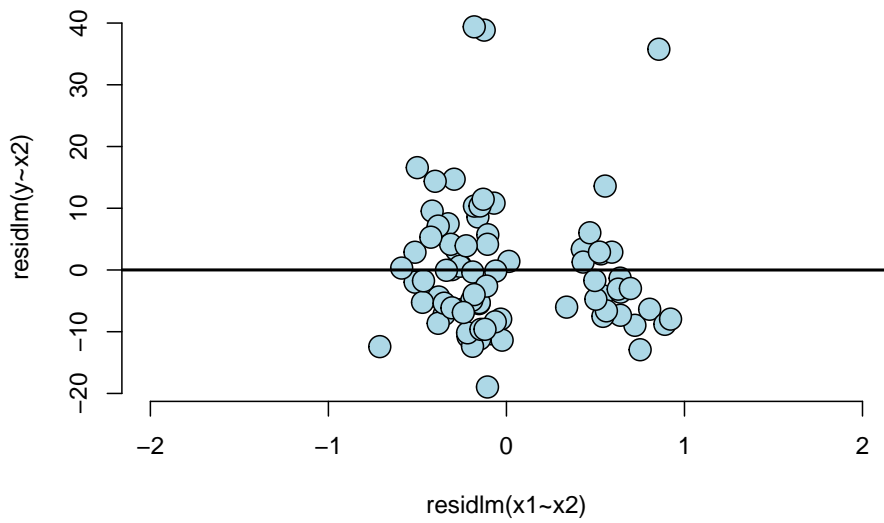
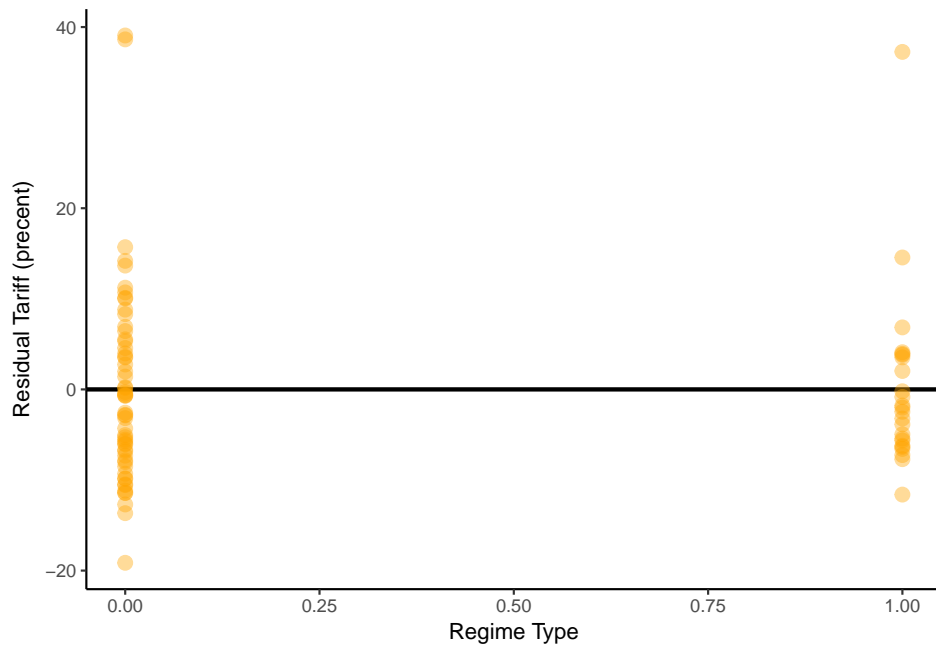
	<i>Dependent variable:</i>					
	Tariff Rates					
	(1)	(2)	(3)	(4)	(5)	(6)
Democracy	-4.580*** (1.581)	-3.427** (1.707)	-3.479* (1.748)	-3.443** (1.719)	-2.657 (2.108)	-3.507* (1.769)
GDP per capita		-1.210 (0.924)	-1.266 (0.935)	-1.213 (0.926)	-1.307 (0.947)	-1.272 (0.938)
Economic Crisis			-3.419 (3.614)			-3.565 (3.660)
GATT/WTO Member				0.313 (2.532)		0.532 (2.588)
Years of Office					0.107 (0.129)	
Constant	21.170*** (1.266)	29.160*** (6.313)	29.960*** (6.537)	28.940*** (6.751)	28.740*** (6.329)	29.590*** (6.874)
Observations	81	81	81	81	81	81
R <sup>2</sup>	0.077	0.103	0.110	0.103	0.110	0.110
Adjusted R <sup>2</sup>	0.065	0.080	0.075	0.068	0.076	0.063
Residual Std. Error	7.189	7.290	7.259	7.365	7.109	7.365

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## Check the assumption of linearity

What assumptions do we have about the unmeasured forces influencing the outcome variables? In the linear regression, we assume linearity as a function form of the measured regressors and the dependent variable. However, all it requires is the independent variables should not be *perfectly* collinear (Achen, 1982). According to Achen (1982), if these assumptions are met, the coefficient estimates from the OLS are consistent, meaning they will be very near to their true value in a large sample almost all the time. In other words, in general conditions, regression will produce fairly accurate estimates if we set up the model correctly. Great resilience is a strength of ordinary linear regression (Achen, 1982). That is probably why we should analyze our data with this powerful tool in the first place.



I use the residual plot to check the linearity assumption. Residuals are the vertical distance between the outcomes and the fitted regression line. It is difficult to see its pattern when the independent variable is a binary one. However, linearity is satisfied because the independent variable only has two possibilities (0 or 1). Therefore, the estimator here is the differences of means between the treated and controlled group.

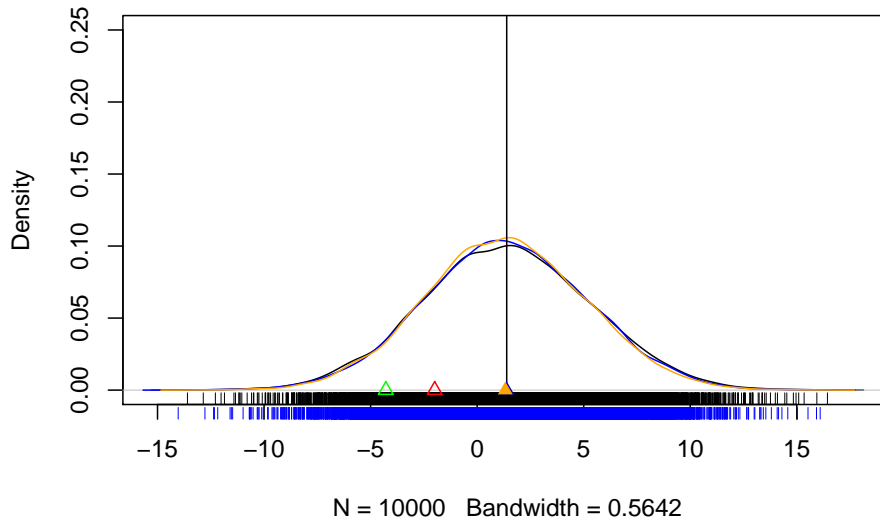
I am concerned about the partial relationship between tariff rate and democracy score. I am interested in seeing what percent of the variation in the full multiple regression model (Model 2 in Table 2) cannot be explained by the independent variable but can be explained by the rest of the confounders. It is easier to show the partial relationship when I only have one independent variable and one control variable so I can remove one of them at a time. In this plot, I can see by removing one my covariate, “log(GDP\_c)”, there is still residual variability left after accounting for it. After removing the linear relationships between the potential confounder and both the explanatory and the outcome, the residual plot shows there is still a linear relationship between tariff rates and regime type.

### Checking unbiasedness and consistency of the estimates in the OLS model

	bestATE	unbiasedATE	bestATE2
trueATE	1.39	1.39	1.39
sampmeans	1.39	1.38	1.32
bias	0.00	0.01	0.07
sd	3.96	3.85	3.76
MSE	15.64	14.84	14.12

Table 4: Simulation results from different estimates for OLS model

### Simulation results from estimates for different OLS models



The *bestATE* is the estimator in a *lm* function with all the relevant covariates in the model. The *unbiasedATE* is the estimator in a *lm* function with no covariates in the model at all. The *bestATE2* is the estimator in a residual-based function.

To assess biasedness, I compare sample means of the estimators with the true mean I created in the simulation test. Simply, I can compare the third row, the values of bias for each estimator, and choose the smallest one (bias is denoted as the absolute positive values: the differences between trueATE and the means of the estimates). All of the three estimators are close to the true mean (1.3866) I created in the simulation test. This suggests all three of them are (pretty much) unbiased. In fact, the *bestATE* has the lowest bias out of the three. Its absolute distance to the true mean is the smallest one, with bias equals to 0.0031. The *bestATE2* is slightly biased (with biase equalling to 0.0673).

Mean Squared Error (MSE) is the mean of the squared differences between the estimated value and the true value. It combines bias and efficiency to measure the closedness of the estimator to the ture parameter. To assess efficiency, the MSEs for *bestATE* and *bestATE2* are relatively the same, but the *bestATE2* behave slightly better ( $14.1233 < 15.6412$ ). The three estimators *bestATE*, *unbiasedATE* and *bestATE2* are all efficient and consistent, suggesting that these three estimate are very converging to the true mean with relatively low variance. From this simulation test, because the *bestATE2* has the lowest MSE, this estimate is preferred.

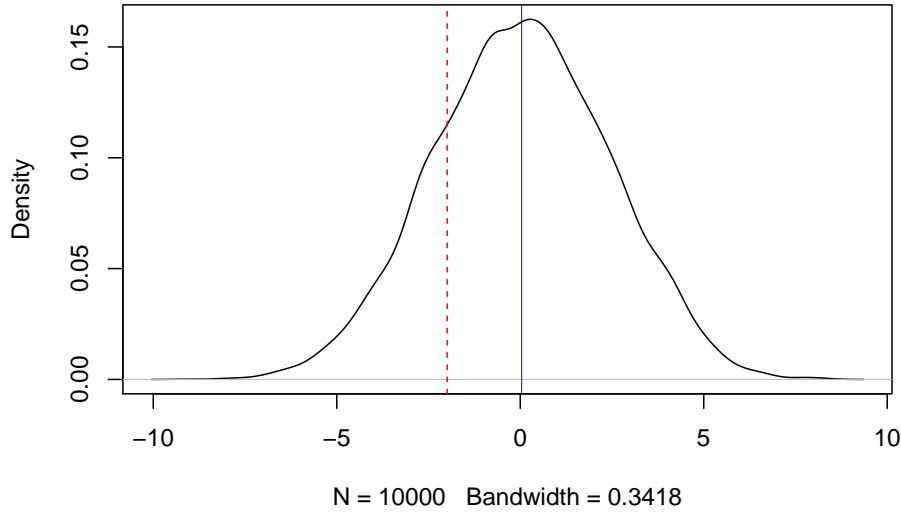
## Section III Statistical Inference and Assessment II

In this section, I present the first method of doing statistical inference: permutation: When we have conducted an experiment, and present the readers the treatment effects, we are worried about whether there is a real treatment effect or we get the result by chance. We can surely redo the experiment, but we do not have to. Permutation gives us a way to compute the resampling distribution for test statistics when we assume there is no treatment effect on the outcome of the null hypothesis. If the null hypothesis is true, data after shuffling (the relationships between treatment and outcome are broken) should look like the observed data. P-value gives us information about the probability of getting the observed or more extreme data assuming the null hypothesis is true.

### P-values from the permutation test

P-value tells us how likely I can get the observed treatment effect from my experiment under the no treatment effect null hypothesis. After I have done the hypothetical experiment, I would do a hypothesis testing. Here, in this study, the hypothetical experiment is that countries are “randomly assigned” to be democratic or non-democratic on average over the 1980s and 1990s. The worrisome is the Fisher’s sharp null hypothesis: there is a possibility of no effect for all the units in this hypothetical experiment. Instead, I just observe the differences in means by chance. My null hypothesis is there is no treatment effect between the treated and control groups for each unit. In other words, the null hypothesis is there are no differences in the tariff rates between democratic and non-democratic countries.

### Sampling distribution under null hypothesis and Observed Statistic

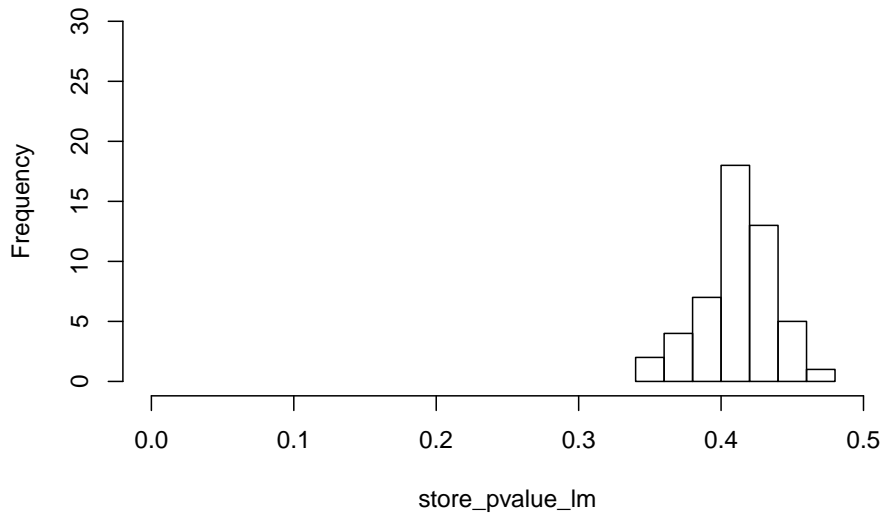


First, I use a test statistic to summarize my observed data from the experiment: -1.9923 is my test statistic of mean difference in tariff rates. Second, I set the null hypothesis of no effects: I create a new experiment and shuffle the labels of countries. Here, I break the relationships existing in the data structure to create an experiment of no effect. Then, by using the computing power, I replicate the experiments of no effect in the testing on the computer as if I run the experiments 10000 times. Then I observe the differences-in-means 0.0381 under the null hypothesis, and we compare how likely it is to get the differences-in-means greater than or equal to the observed data. This probability is the p-value 0.409. It means we have 0.409 (around 1 in 5 replications of the no effect experiment) to produce the values as large as or greater than the estimators in the *lm* function. The p-value here is the probability that value as extreme or more extreme will be observed under the null hypothesis. This probability gives me the information that I may not have much evidence to against the null effect hypothesis, which is the difference between the observed treatment effect and the effect under the null hypothesis is not due to chance.

I use the permutation test to obtain the p-value where the test statistic is calculated based on the data itself under the null hypothesis. The key advantage of this test does not rely on any assumptions of the distribution. In the canned *lm* function, the standard assumption that the statistic follows a t-distribution gives a p-value of 0.4739 (by default). This is in quite good agreement with the p-value I obtained in the permutation test 0.409. But I would not necessarily know beforehand that the two p-values would agree. The following figure shows the null distribution obtained from using the data itself is close to a t-distribution. This can explain why the p-value from the CLT+IID justified test and the p-value from the permutation test is similar.

[1] 0.384

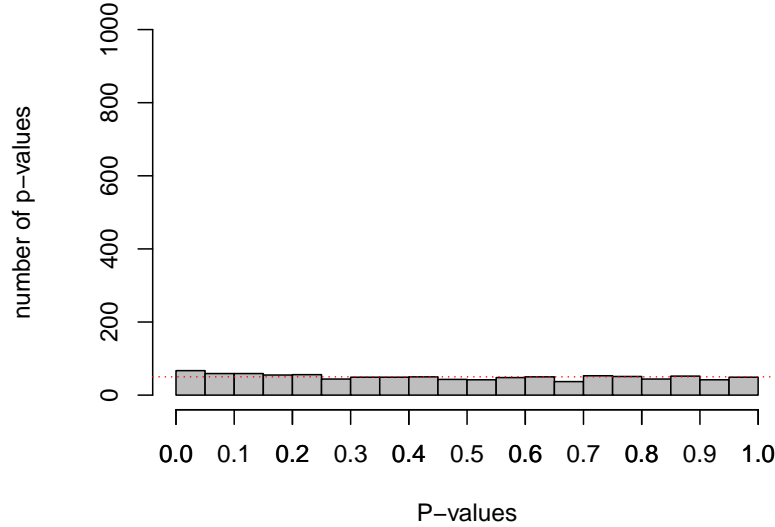
### P-value distribution with Permutation in 50-times simulation



After I calculated the p-value 0.409 from a permutation test, I replicated this process for 50 times and calculated 50 different p-values generated from the same process. From this histogram, we can see that the p-values are distributed around from 0.22 to 0.26.

To check the error rate of the *lm function*, I create a null effect in the error rate test knowing that my null hypothesis is true. If the falsely positive rate is 0.05, this means in 5 out of 100 times, the test falsely rejects the nulls (knowing the null is true but I still reject it). If the false positive rate is close to 0.05, it means the test fulfils its promises. The false positive rate is 0.051. If we run it a couple of times, the false positive rates are slightly different, but they are around 0.05. The p-value from the built-in *lm* function has a similar false positive rate to the nominal false positive rate (0.05).

### P-value Distribution under Null Effect in 1000-times Simulation



In this plot, we know that when there is no true effect, p-values are what is called ‘uniformly distributed under the null’. The p-value distribution is basically flat. Every p-value is equally likely when the null hypothesis is true, and every bar in the graph will contain 5% of all the p-values (as indicated by the dotted red line). The first bar is the false positive rate, which is slightly higher than but it is very close to 0.05.

I also followed the same procedures to calculate the p-value and false positive rates of the *lmrob* function. I summarize the results in the following table.

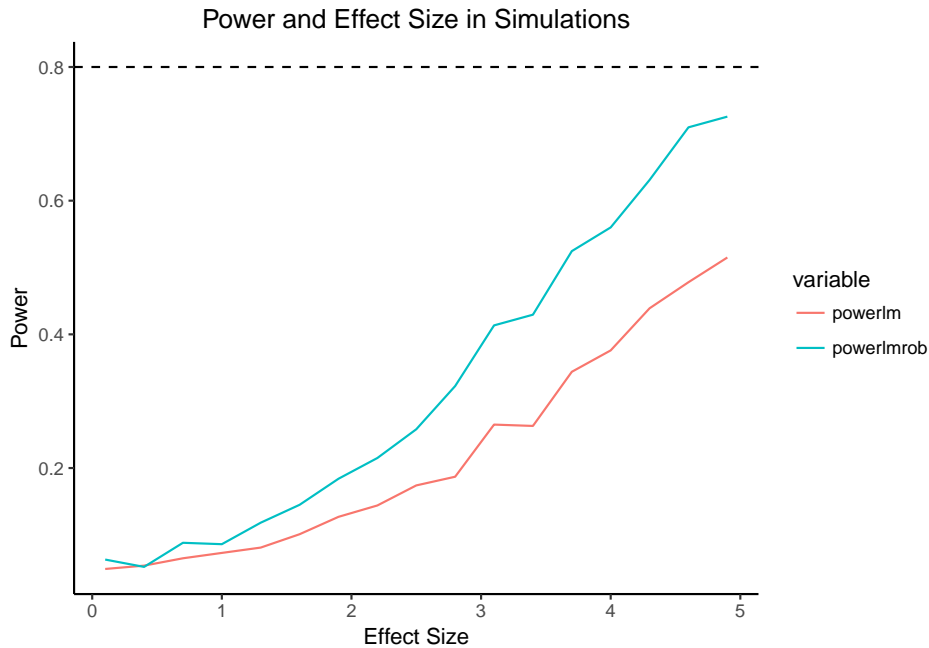
	Permutation	t-distribtuion	False Positive Rate
lmp	0.41	0.47	0.05
lmrobp	0.06	0.05	0.06

Table 5: P-values obtained from simulation, t-distribution and their error rates

### Power and effect size

When my study has effects, I hope that the test has the power to detect the true effect when the null hypothesis is false. Increasing the power of the test requires bigger sample sizes, or studying larger effects. Here, my sample size is 81, and I want to check which test (of *lm* or *lmrob*) has higher power for different effect sizes.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
effectsize	0.10	0.40	0.70	1.00	1.30	1.60	1.90	2.20	2.50	2.80	3.10	3.40	3.70	4.00
somepower	0.05	0.05	0.06	0.07	0.08	0.10	0.13	0.14	0.17	0.19	0.26	0.26	0.34	0.38
somepower_lmrob	0.06	0.05	0.09	0.09	0.12	0.14	0.18	0.21	0.26	0.32	0.41	0.43	0.52	0.56



I can use simulations to estimate the statistical power of a model. The statistical power is the probability of observing a statistically significant result, if there is a true effect. When there is an effect, I hope that my statistical test is able to detect it. This denotes to high power in my study.

Cohen describes effect size as the degree to which the null hypothesis is false. In this simulation test, I generate different hypothetical effect sizes (from 0.1 to 5), and I calculate the number of p-values that are lower than 0.05 (“reject the null”) when I know there is a true effect (the null is false). When the effect size increases, the powers in both functions also increase.

For a given sample size, the *lmrob* model has larger statistical power given an effect size. As effect size increases, the power of the *lmrob* model is also increasing faster than that of the *lm* model. To achieve an ideal 80% statistical power, the *lmrob* model requires an effect size larger than 5. 80% statistical power essentially means when there is a true effect, there is 80 percent that I will observe a significant effect. For this *lm* model, I need a bigger effect size to achieve the same level of power as *lmrob* model requires. The tests may have lower power so I may not be able to identify a significant result where there is a real effect.



## Citation

- Achen, Christopher H. 1982. *Interpreting and Using Regression*. Newbury Park, CA: Sage.
- Angrist J., Pischke J-S. 2015. *Mastering 'Metrics: The Path from Cause to Effect*
- Baccini, L., & Urpelainen, J. (2014). International Institutions and Domestic Politics: Can Preferential Trading Agreements Help Leaders Promote Economic Reform? *The Journal of Politics*, 76(1), 195-214. <https://doi.org/10.1017/S0022381613001278>
- Campbell. "Reforms as Experiments," *American Psychologist*, 1969, pp. 409-429.
- Gelman, A., and J. Hill. 2007. *Data Analysis Using Regression and Multilevel/ Hierarchical Models*. Cambridge University Press.
- Kruger, Anne. 1997. Trade Policy and Economic Development: How We Learn. *American Economic Review* 87(1): 1-22.
- Mansfield E.D, Eric Reinhardt. 2003. "Multilateral Determinants of Regionalism: The Effects of GATT/WTO on the Formation of Preferential Trading Arrangements," *International Organization* 57:4, 829-862.
- Mansfield, E. D., Milner, H. V., & Rosendorff, B. P. (2002). Why Democracies Cooperate More: Electoral Control and International Trade Agreements. *International Organization*, 56(3), 477-513.
- Rodrik, Dani. 1994. "The Rush to Free Trade in Developing World: Why So Late? WHY Now? Will it Last?" In *Voting for Reform*, edited by Stephan Saggard and Steven Webb, 1457-94. New York: Oxford University Press.
- Shadish W., Cook T. and Campbell, Eperimental and Quasi-Experimental Designs. 2002.
- Vreeland, James Raymond. 2007. *The International Monetary Fund: Politics of Conditional Lending*. New York: Routledge

## Appendix

Table 6: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Year	3,580	1,990.0	5.8	1,980	1,985.0	1,994.0	1,999
Tariff	897	20.8	15.0	0.0	11.0	27.0	102.2
Openness	1,860	0.4	0.5	0	0	1	1
Regime	2,292	-1.1	7.1	-10	-7	7	10
Democracy	2,950	0.4	0.5	0	0	1	1
Dictator	2,960	5.0	2.9	1	2	8	8
Years of Office	2,385	8.4	8.3	0	2	12	44
GATT/WTO	3,173	0.5	0.5	0	0	1	1
Economic Crisis	2,379	0.1	0.3	0	0	0	1
GDP PC 95	2,737	2,912.0	4,450.0	84.7	447.8	3,140.0	37,841.0
IMF	2,836	0.1	0.4	0	0	0	1

```

## To make the pdf file do
## render("exploration4.Rmd",output_format=pdf_document())

require(knitr)
opts_chunk$set(
  tidy=FALSE,      # display code as typed
  size="small",    # slightly smaller font for code
  echo=TRUE,
  results='markup',
  strip.white=TRUE,
  fig.path='figs/fig',
  cache=FALSE,
  highlight=TRUE,
  width.cutoff=132,
  size='footnotesize',
  out.width='.9\\textwidth',
  message=FALSE,
  comment=NA)

##First, just setup the R environment for today:
if(!file.exists('figs')) dir.create('figs')

options(SweaveHooks=list(fig=function(){
  par(mar=c(3.5, 3, 1.1, 0),
      pty="s",
      mgp=c(1.5,0.5,0),
      oma=c(0,0,0,0)),
  echo=function(){options(continue=" ") ##Don't show "+" prompts,
  options(prompt=" ")
  }},
  digits=4,
  scipen=8,
  width=132
)
options(error=function(){options(prompt="> ",continue="+ ");NULL})

getwd()

#To prepare the environment
library(readstata13)
library(dplyr)
library(ggplot2)
#install.packages("ggrepel")
library(ggrepel)
library(stargazer)

#install.packages("devtools")

```

```

library(devtools)

library(foreign)
#install.packages("gplots")
library(gplots)

library(lmtest)
library(sandwich)

library(tidyverse)

library(MASS)
library(robustbase)

#install.packages("rmngb")
library(rmngb)
library(here)

devtools::install_github("ropenscilabs/gramr")
library("gramr")
#write_good_ip()
getwd()
library("here")
set_here(path=".", verbose=T)
here()

list.files(path=".")
load("mydata_2fulldataset.rda")
load("mydata_5collapseyear.rda")

#library(readstata13)
mydata <- read.dta13(file.choose())
#View(mydata)
#Subset the large dataset
mydata2 <- mydata[,c(2,3,14,7,15,5,40,17,4,8,10,16), drop=F]

#View(mydata2)
colnames(mydata2) <- c("Country", "Year", "Tariff", "Openness", "Regime", "Democracy", "Dictator", "Year")
mydata_2 <- subset(mydata2, Year > 1979) #Select the years after 1979.

save(mydata_2, file="mydata_2fulldataset.rda")

#The change of the statutory tariff rate
m1980 <- mean(mydata2$Tariff[mydata2$Year==1980], na.rm = T)
m1999 <- mean(mydata2$Tariff[mydata2$Year==1999], na.rm = T)
mperc<- (m1999-m1980)/m1980

```

```

mperc #-0.58

result_m_tariff <- data.frame(matrix(nrow=20, ncol=2))
colnames(result_m_tariff) <- c ("Year", "Tariff")
for (i in 1:20){
  result_m_tariff[i,1] <- i+1979
  mean <- mean(mydata2$Tariff[mydata2$Year== i+1979], na.rm = T)
  result_m_tariff[i,2] <- mean
}
result_m_tariff
#plot(result_m_tariff)

o1980 <- sum(mydata2$Openness[mydata2$Year==1980 & mydata2$Openness == 1], na.rm = T) #14
o1999 <- sum(mydata2$Openness[mydata2$Year==1999 & mydata2$Openness == 1], na.rm = T) #62
operc <- (o1999-o1980)/o1980

result_m_regime_2 <- data.frame(matrix(nrow=30, ncol=2))
colnames(result_m_regime_2) <- c ("Year", "PolityScore")
for (i in 1:30){
  result_m_regime_2[i,1] <- i+1969
  mean <- mean(mydata2$Regime[mydata2$Year== i+1969], na.rm = T)
  result_m_regime_2[i,2] <- mean
}
result_m_regime_2

#plot(result_m_regime_2)

result_m_regime_3 <- data.frame(matrix(nrow=30, ncol=2))
colnames(result_m_regime_3) <- c("Year", "Democracy")

for (i in 1:30){
  result_m_regime_3[i,1] <- i+1969
  counter <- sum (mydata2$Democracy[mydata2$Year== i+1969 & mydata2$Democracy ==1], na.rm=T)
  result_m_regime_3[i,2] <- counter
}

result_m_regime_3

#plot(result_m_regime_3)
#summary(mydata_2)
#str(mydata_2)
#dim(mydata_2)

#####Explore the data
#Check the missing values

```

```

#table(mydata_2$Tariff, useNA="ifany") #The missing values: NA = 2683
sum(mydata_2$Tariff == " ", na.rm=T) #if there is any missing valus in the blank
sum(mydata_2$Tariff == "0", na.rm=T) #20
is.na(mydata_2)[1,] #From this, we can see only the variable "Tariff" has missing values

#Check the missing values of Tariff before 1980
mydata2_3<- subset(mydata2, Year <= 1979)
#table(mydata2_3$Tariff, useNA="ifany") #The missing values: NA = 2683

#describe the data
library(psych)
#subset the data without country name
mydata_2_des_noc <- mydata_2[,!names(mydata_2)%in%c("Country")]
#basic descriptive Statistics
data2_des <- describe(mydata_2_des_noc, na.rm=T, check=T, skew=F, quant=c(.25,.5,.75))
data2_des

#install.packages("stargazer") If you haven't already installed. Install stargazer.
library(stargazer)
#Now use stargazer to format the linear model
stargazer(mydata_2_des_noc, header = FALSE, iqr=T,
          title = "Summary Statistics", digits = 1, out = "table1.txt",
          no.space = TRUE, font.size = "footnotesize")

#table(mydata_2$Tariff, useNA="ifany") #The missing values from 1980 to 1999: NA = 2683
sum(mydata_2$Tariff == " ", na.rm=T) #if there is any missing valus in the blank
sum(mydata_2$Tariff, na.rm=T)
2683/18628

is.na(mydata_2)[1,] #From this, we can see only the variable "Tariff" has missing values

##---

mydata2_3<- subset(mydata2, Year <= 1979) #before 1979
table(mydata2_3$Tariff, useNA="ifany") #The missing values: NA = 1780

table(mydata2$Openness, useNA="ifany") #The missing values of openness: NA = 2580
mydata3 <- mydata[,c(2,3,14,7,31,11,41,28,49,13,21,34), drop=F]
colnames(mydata3) <- c("Country", "Year", "Tariff", "Openness", "Regime(t-1)", "Democracy(t-1)", "Dicta
                      "GATT/WTO(t-1)", "Economic Crisis(t-1)", "GDP PC 95(t-1)", "IMF(t-1)")

save(mydata3, file="mydata3_fulldatasetlagged1year.rda")

Regime_l1 <- mydata3[, "Regime(t-1)"]
WTO_l1 <- mydata3[, "GATT/WTO(t-1)"]
Y0_l1 <- mydata3[, "Years of Office(t-1)"]
EC_l1 <- mydata3[, "Economic Crisis(t-1)"]

```

```

IMF_l1 <- mydata3[, "IMF(t-1)"]
GDP_l1 <- mydata3[, "GDP PC 95(t-1)"]
lnGDP_l1 <- log(mydata3[, "GDP PC 95(t-1)"])

mydata4 <- mydata3
Regime_l1 <- mydata3[, "Regime(t-1)"]

## Recode this variable into 1= "democracy (polity score is 6 or more) or more" versus 0="non-democracy"
mydata4 <- mutate(mydata3, Regime_l1_Bi = 1*(Regime_l1 >= 6))
#Check the recoding
with(mydata4, table(Regime_l1_Bi, Regime_l1, useNA="ifany"))
#I turned Polity's Regime variable into a dichotomous variable with countries scoring below 6 as non-democracy

save(mydata4, file="mydata4.rda")

#####
row.has.na <- apply(mydata4, 1, function(x){any(is.na(x))})
sum(row.has.na)

mydata4_small <- mydata4[!row.has.na,]
summary(mydata4_small)
nrow(mydata4) #5370
nrow(mydata4_small) #662

Regime_l1_Bi_s <- mydata4_small$Regime_l1_Bi
GDP_l1_s <- mydata4_small$`GDP PC 95(t-1)`
EC_l1_s <- mydata4_small$`Economic Crisis(t-1)`
WTO_l1_s <- mydata4_small$`GATT/WTO(t-1)`
YO_l1_s <- mydata4_small$`Years of Office(t-1)`

save(mydata4_small, file="mydata4_NAremoved.rda")

#getwd()
#list.files()
load(file="C:/Users/Lucie Lu/Documents/531-explorations/Finalpaper/Data/mydata4_NAremoved.rda")

mydata_5 <- mydata4_small %>%
  group_by(Country) %>%
  summarize(Tariff_c = mean(Tariff),
            Regime_c = mean(`Regime(t-1)`),
            YearOffice_c = mean(`Years of Office(t-1)`),
            WTO_c = mean(`GATT/WTO(t-1)`),
            EC_c = mean(`Economic Crisis(t-1)`),
            GDP_c = mean(`GDP PC 95(t-1)`),
            IMF_c = mean(`IMF(t-1)`))

```

```

#the number of observations in each cluster
mydata4_small %>%
  group_by(Country) %>%
  summarize(no_obs=n())

mydata_5 <- mutate(mydata_5, Regime_Bi_c = 1*(Regime_c >= 6))
mydata_5$countryF<-factor(mydata_5$Country)

save(mydata_5,file="mydata_5collapseyear.rda")

plot (mydata2$Regime, mydata2$Tariff, col = "darkblue", pch = 20, xlab = "Regime Type (Polity Score)",
      main = "Bivariate Relationship between \n Regime Type and Tariff Rate")

OLS_1 <- lm(Tariff ~ Regime, data=mydata2)
abline(OLS_1)

r_tr <- cor(mydata2$Tariff, mydata2$Regime, use = "pairwise.complete.obs")

rr <- mydata2$Regime[!is.na(mydata2$Regime)]
rt <- mydata2$Tariff[!is.na(mydata2$Tariff)]

text(quantile(rr, .95), quantile(rt, .99), paste("r=",round(r_tr,2)), font=2)
Regime <- mydata2$Regime
Tariff <- mydata2$Tariff
lnGDP <- log(mydata2[, "GDP PC 95"])
EC <- mydata2[, "Economic Crisis"]
WTO <- mydata2[, "GATT/WTO"]
YO <- mydata2[, "Years of Office"]

matrix_cor <- data.frame(Tariff, Regime, lnGDP, EC, WTO, YO)
#as.matrix(matrix_cor)

rcor<-cor(matrix_cor,use="pairwise.complete.obs")
for(i in 1:6){
  for(j in 1:6){
    if(j<i){rcor[i,j] <- NA}
  }}
round(rcor,2)

#install.packages("xtable",repos = "http://cran.us.r-project.org")
library(xtable)
names <- c("Tariff rate", "Regime (Polity Score)", "GDP per capita (log)", "Economic Crisis", "GATT/WTO", "Years of Office")
rownames(rcor) <- names
TAB1<-round(rcor,2)
xtable(TAB1)

```



```

print(xtable::xtable(rcor, caption = "Correlation Matirx: Correlations among Variables and Confor
      html.table.attributes="border=1")
mydata4_small$countryF<-factor(mydata4_small$Country)

countryMeans<-with(mydata4_small,tapply(Tariff,countryF,mean))
countrydiff<-var(countryMeans)
yeardiff<-mean(with(mydata4_small,tapply(Tariff,countryF,var)), na.rm = T)
#This code gives the mean of within-country (grouped by countryF) variance

#Across-country variance is around 1.7 times of the within-country-years variance. It means that
countrydiff #123.7
yeardiff #70.71

#Interclass correlation (Gilman & Hill, p. 258): the relative values of country- and year-level
yeardiff/(countrydiff+yeardiff) #0.36
countrydiff/(countrydiff+yeardiff) #0.63

#install.packages("ICC")
library(ICC)
ICC_wc <- ICCbare(y=Tariff,x=countryF,data=mydata4_small)
ICC_wc #0.63

# It is a method of Simple Estimation of the Intraclass Correlation Coefficient. The ICC score d
# High intraclass correlation means values(y) from the same group tend to be similar. The varian

## to visualize variances accross countries

plotCF <- ggplot(mydata4_small, aes(x=countryF, y=Tariff)) + geom_boxplot()

plotCF + theme_classic() +ggtitle("Tariff variations across countries") + theme(
  plot.title = element_text(hjust = 0.5),
  axis.title.x=element_blank(),
  axis.text.x=element_blank(),
  axis.ticks.x=element_blank()) + labs (X = "Country", y= "Tariff Rates (in percentage)")
## to visualize variances accross years

Year.factor <- factor(mydata4_small$Year)
plotyears <- ggplot(mydata4_small, aes(x=Year.factor, y=Tariff)) + geom_boxplot()

plotyears +
  theme_classic() +
  ggtitle("Tariff variations across Years") +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs (X = "Year", y= "Tariff Rates (in percentage)")

```

```

plot_bi5 <- ggplot(mydata_5, aes(x= Regime_c, y = Tariff_c, label=countryF)) + geom_point()

### geom_label_repel

plot_bi5 +
  geom_text_repel(data = filter(mydata_5, (Regime_c < -3 & Tariff_c >30) | (Regime_c > 6 & Tariff_c > 30)),
    theme_classic() +
    ggtitle("Bivariate Relationship between \n Regime Type and Tariff Rate (with Country Labels)",
    theme(plot.title = element_text(hjust = 0.5)) +
    labs (x = "Regime Type (Polity Score)",
          y= "Tariff Rate (in percentage)")

#comparing democracy and non-democracy
OLS_5_0Bs <- lm(Tariff_c ~ Regime_Bi_c, data=mydata_5)
summary(OLS_5_0Bs)$coefficient #-4.287

OLS_5_1Bs <- lm(Tariff_c ~ Regime_Bi_c + log(GDP_c), data=mydata_5)
summary(OLS_5_1Bs)$coefficient #-1.748

OLS_5_2Bs <- lm(Tariff_c ~ Regime_Bi_c + log(GDP_c) + EC_c, data=mydata_5)
summary(OLS_5_2Bs)$coefficient #-1.83

OLS_5_3Bs <- lm(Tariff_c ~ Regime_Bi_c + log(GDP_c) + WTO_c, data=mydata_5)
summary(OLS_5_3Bs)$coefficient #-1.89

OLS_5_4Bs <- lm(Tariff_c ~ Regime_Bi_c + log(GDP_c) + YearOffice_c, data=mydata_5)
summary(OLS_5_4Bs)$coefficient #-2

OLS_5_5Bs <- lm(Tariff_c ~ Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=mydata_5)
summary(OLS_5_5Bs)$coefficient #-2

library(stargazer)
stargazer(OLS_5_0Bs, OLS_5_1Bs, OLS_5_2Bs, OLS_5_3Bs, OLS_5_4Bs, OLS_5_5Bs, header = FALSE, df =
  title = "Multiple Linear Regression Model: Results",
  covariate.labels = c("Democracy", "GDP per capita", "Economic Crisis", "GATT/WTO Member"),
  dep.var.labels = "Tariff Rates",
  column.sep.width = "0.3pt", no.space = TRUE, font.size = "footnotesize")

set.seed(2000814)

#3 observations are outliers
lmrob_5_0Bs <- lmrob(Tariff_c ~ Regime_Bi_c, data=mydata_5, control = lmrob.control(max.it = 100))
summary(lmrob_5_0Bs)$coefficient #-4.58

lmrob_5_1Bs <- lmrob(Tariff_c ~ Regime_Bi_c + log(GDP_c), data=mydata_5, control = lmrob.control(max.it = 100))
summary(lmrob_5_1Bs)$coefficient #-3.18

```

```

lmrob_5_2Bs <- lmrob(Tariff_c ~ Regime_Bi_c + log(GDP_c) + EC_c, data=mydata_5, control = lmrob.
summary(lmrob_5_2Bs)$coefficient #-3.27

lmrob_5_3Bs <- lmrob(Tariff_c ~ Regime_Bi_c + log(GDP_c) + WTO_c, data=mydata_5, control = lmrob
summary(lmrob_5_3Bs)$coefficient #-3.20

lmrob_5_4Bs <- lmrob(Tariff_c ~ Regime_Bi_c + log(GDP_c) + YearOffice_c, data=mydata_5, control =
summary(lmrob_5_4Bs)$coefficient #-2.50

lmrob_5_5Bs <- lmrob(Tariff_c ~ Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=mydata_5, control =
summary(lmrob_5_5Bs)$coefficient #-3.3

library(stargazer)
stargazer(lmrob_5_0Bs, lmrob_5_1Bs, lmrob_5_2Bs, lmrob_5_3Bs, lmrob_5_4Bs, lmrob_5_5Bs, header =
  title = "Multiple Linear Regression Model (lmrobust): Results",
  covariate.labels = c("Democracy", "GDP per capita", "Economic Crisis", "GATT/WTO Member
  dep.var.labels = "Tariff Rates",
  column.sep.width = "0.3pt", no.space = TRUE, font.size = "footnotesize")

#Residual plot

g_re <- ggplot(data.frame(x=mydata_5$Regime_Bi_c, y=resid(OLS_5_1Bs))
  , aes(x=x, y=y))
g_re <- g_re + geom_hline(yintercept=0, size=1)
#g_re <- g_re + geom_point(size=5, colour="black", alpha=0.4)
g_re <- g_re + geom_point(size=3, colour="orange", alpha=0.4)
g_re <- g_re + xlab("Regime Type") + ylab("Residual Tariff (precent)")
g_re + theme_classic()
OLS_5_1Bs <- lm(Tariff_c ~ Regime_Bi_c + log(GDP_c), data=mydata_5)
OLS_5_1Bsyres <- resid(OLS_5_1Bs)
OLS_5_1x1x2res <- resid(lm(Regime_Bi_c ~ log(GDP_c),data=mydata_5))
OLS_5_1yx2res <- resid(lm(Tariff_c ~ log(GDP_c),data=mydata_5))

plot(OLS_5_1x1x2res, OLS_5_1yx2res, frame = FALSE, col = "black", bg = "lightblue", pch = 21, ce
abline(lm(I(OLS_5_1x1x2res) ~ I(OLS_5_1yx2res)), lwd = 2)

OLS_5_0Bs <- lm(Tariff_c ~ Regime_Bi_c, data=mydata_5)
summary(OLS_5_0Bs)$coefficient[2,1] #-4.29

OLS_5_5Bs <- lm(Tariff_c ~ Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=mydata_5)

summary(OLS_5_5Bs)$coefficient[2,1] #-2.84

#####

```

```

#z is the experiment, shuffle the mydata4$Regime_l1_Bi variable to create a pretend experiment

set.seed(2018)

mydata_5test <- mydata_5

#rm(mydata_5test$a)
#mydata_5test %>% select(-a)
mydata_5test$Z <- sample(rep(c(0,1), each = 1, len = 81))

#Define a potential outcome?

#Invent TrueATE

#First come up with a way for covariates to be related to the potential outcome under control (y0)
mydata_5test$y0 <- with(mydata_5test, 3 + .05*log(GDP_c) + .4*EC_c - .8*WTO_c) +
runif(n = 81, min=min(mydata_5test$Tariff_c),max=max(mydata_5test$Tariff_c))
#To create background noise...no effect at all.

lm_test2 <- lm(y0 ~ Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=mydata_5test)
summary(lm_test2)$r.squared #0.98

mydata_5test$y1 <- with(mydata_5test, y0 + .08*sd(y0))

trueATE <- mean(mydata_5test$y1) - mean(mydata_5test$y0)
trueATE #11.43
.08*sd(mydata_5test$y0) #11.43

#Now, I know the trueATE: 11.43

## Observed outcome: y1 among the treated, y0 among the controls
mydata_5test$Y <- with(mydata_5test, Z*y1 + (1-Z)*y0)

lm_test3 <- lm(Y ~ Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=mydata_5test) #also use the cov
summary(lm_test3)$r.squared #0.98

## Notice that this includes all of the true covariates, just not in the correct function, plus 1
estATEbest_5_5Bs <- coef(OLS_5_5Bs)[["Regime_Bi_c"]] #-2.84
estATEbest_5_5Bs ##-2.84

##Unbiased estimates: includes nothing (no covariates)
estATEunbiased_5_5Bs <- coef(OLS_5_0Bs)[["Regime_Bi_c"]]

estATEunbiased_5_5Bs ##-4.29

```

```

set.seed(2018)

## Bias refers to a relationship between the repeated operation of a procedure and a truth.
## So we have to invent a truth (that is why we **created** potential outcomes for all units ---
## ordinarily we would not observe these quantities).

#Repeat the experiment. Each repetition reveals a different potential outcome for each person. C
#This should show (1) bias and (2) whether one or the other is better in mean squared error term
#--- especially that one or the other converges to the truth more or less quickly as sample size.

newEstimate<-function(obsz,dat){
  newExperiment<-function(z){
    #sample(z) is permutating z
    sample(z)
  }
  #create newz
  dat$newz <- newExperiment(obsz)
  ##names(newz) <- row.names(dat)
  #create newY
  dat$newY <- with(dat,newz*y1 + (1-newz)*y0)
  #use three methods to create ATE
  theestATEbest <- coef(lm(newY~newz + Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=dat))["newz"]
  theestATEunbiased <- coef(lm(newY~newz,data=dat))["newz"]
  ## Another method of using covariates
  dat$e1 <- residuals(lm(newY~Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=dat))
  theestATEbest2 <- coef(lm(e1~newz,data=dat))[[2]]
  return(c(bestATE=theestATEbest,
           unbiasedATE=theestATEunbiased,
           bestATE2=theestATEbest2))
}

## test, maybe a good idea to test it after we create a function
## newEstimate(obsz=wrkdat$Z,dat=wrkdat)

## An unbiased estimator is one where  $E[\text{estimator}] = \text{Truth}$ . Hmmm.. Is this the right test?
set.seed(1234568)
estdists <- replicate(10000,newEstimate(obsz=mydata_5test$Z,dat=mydata_5test))

trueATE #11.36
##sampdistmeans<-apply(estdists,1,mean)
#meanATE
sampmeans <- apply(estdists[1:3,],1,mean)
#sampledistmeans: very close to the trueATE
sampmeans

```

```

#absolute difference between the estimated mean and the true ATE

bias <- abs(trueATE-apply(estdists[1:3,],1,mean)) # abs(): absolute positive values: the difference
apply(estdists[1:3,],1,function(x){ mean( abs(x - trueATE) ) } )
bias
#the standard deviation of the ATE means

##Finding the standard deviation of the estimated ATEs allows us to see if our estimators are efficient
sd <- apply(estdists[1:3,],1,sd)
sd

#the RMSE of the difference in the true ATE and the estimated ATE

##The RMSE allows us to further assess if the estimators are biased because it overestimates the true ATE
MSE <- apply(estdists[1:3,],1,function(x){ mean( ( x - trueATE)^2 ) })
MSE

#install.packages("xtable",repos = "http://cran.us.r-project.org")
library(xtable)

sim <- as.matrix(rbind(trueATE, sampmeans, bias, sd, MSE))
#rownames(rcor) <- names
simtab<-round(sim,2)
simtab
xtable(simtab)
print(xtable::xtable(sim, caption = "Simulation results from different estimates for OLS model"),
      html.table.attributes="border=1", comment=FALSE)

plot(density(estdists["bestATE",]),ylim=c(0,.25), main = "Simulation results from estimates for OLS model",
     rug(estdists["bestATE",],col="black",line=0)
     lines(density(estdists["unbiasedATE",]),col="blue")
     rug(estdists["unbiasedATE",],col="blue",line=.5)
     lines(density(estdists[3,]),col="orange")
     abline(v=trueATE)

sampdistmeans<-apply(estdists,1,mean)
points(c(sampdistmeans[1:3],estATEunbiased_5_5Bs,estATEbest_5_5Bs), rep(0,5),
       pch=c(17,17,17,2,2),
       cex=1,col=c("black","blue","orange","green","red"))

##bestATE and bestATE2 have a little bias, but it is consistent;
##unbiasedATE is unbiased, but it is very inefficient and inconsistent.

##We can reduce bias only at a potential increase in variance.

```

```

OLS_5_5Bs <- lm(Tariff_c ~ Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=mydata_5)
simlmttest_lm5Bs <- summary(OLS_5_5Bs)$coef[2,1]
simlmttest_lm5Bs ##-2.84
default_p <- summary(OLS_5_5Bs)$coef[2,4]

#Imagine we re-run the "experiment"
newExp<-function(y){
  ## A function to randomly assign treatment effect
  shuffledz <- sample(mydata_5test$Z)
  newlmttest <- coef(lm(y ~ shuffledz + log(GDP_c) + EC_c + WTO_c, data=mydata_5test))["shuffledz"]
  return(newlmttest)
}

#set.seed(19900814)
expBeta_lm5Bs <- replicate(10000,newExp(y=mydata_5test$Tariff_c))
mean(expBeta_lm5Bs) #0.053 treatment effect under null

plot(density(expBeta_lm5Bs), main = "Sampling distribution under null hypothesis and Observed Stat",
     rug(mean(expBeta_lm5Bs), 1, col="blue"))
abline(v=simlmttest_lm5Bs,lty=2, col="red")

#One-sided p-value
pvalue_lm_os <- as.numeric(as.matrix(summary(expBeta_lm5Bs > simlmttest_lm5Bs)))/length(expBeta_lm5Bs)
pvalue_lm_os #0.1134

#Two-sided p-value
pvalue_lm_ts <- 2*min(mean(expBeta_lm5Bs>=simlmttest_lm5Bs),mean(expBeta_lm5Bs<=simlmttest_lm5Bs))
#pretty close to the two-sided p-value obtained in the lm (0.28).
p_lmcan <- summary(OLS_5_5Bs)$coef[2,4]

newExp<-function(y){
  ## A function to randomly assign treatment effect
  shuffledz <- sample(mydata_5test$Z)
  newlmttest <- coef(lm(y ~ shuffledz + log(GDP_c) + EC_c + WTO_c, data=mydata_5test))["shuffledz"]
  return(newlmttest)
}

set.seed(19900814)
#expBeta_lm5Bs <- replicate(10000,newExp(y=mydata_5test$Tariff_c))

#mean(expBeta_lm5Bs)#0.053

pvalue_lm <- function(siml,obsTestStat){
  ## return a p-value
  refDistNull<-replicate(siml,newExp(y=mydata_5test$Tariff_c))

```

```

    pTwoSided <- 2*min(mean(refDistNull>=obsTestStat),mean(refDistNull<=obsTestStat))
    return(pTwoSided)
}

pvalue_lm (siml=1000, obsTestStat=simlmtest_lm5Bs)

store_pvalue_lm <- replicate(50,
pvalue_lm (siml=1000, obsTestStat=simlmtest_lm5Bs))

hist(store_pvalue_lm, xlim=c(0,0.5), ylim = c(0, 30), main=paste("P-value distribution with Permu

#Codes to assess the false positive rate of the canned lm function

err_rate<- function(y){
  shuffledz <- sample(mydata_5test$Z)
  ## since we know that there is no effect, then we can assess the false positive rate of lm
  sim_ps <- summary(lmrob(y ~ shuffledz + log(GDP_c) + EC_c + WTO_c, data=mydata_5test))$coef[1]
  return(sim_ps)
}

set.seed(19900814)
results_lm5Bs <- replicate(1000,
err_rate(y=mydata_5test$Tariff_c))

mean(results_lm5Bs <= .05, na.rm=T) #0.04637

falsepos_lm5Bs <- as.numeric(as.matrix(summary(results_lm5Bs < .05))[3,1])/length(results_lm5Bs)

falsepos_lm5Bs

set.seed(800814)
nSims <- 1000

results_lm5Bs <- replicate(1000,
err_rate(y=mydata_5test$Tariff_c))#get the p-value and store it
#Check power by summing significant p-values and dividing by number of simulations

#The frequency of p-values, or the type 1 error is 0.043. When there is no effect (in the simula

bars<-20

op <- par(mar = c(5,7,4,4)) #change white-space around graph
hist(results_lm5Bs, breaks=20, xlab="P-values", ylab="number of p-values\n", axes=T,

```



```

    main=paste("P-value Distribution under Null Effect in 1000-times Simulation"),
    col="grey", xlim=c(0,1), ylim=c(0, nSims))
axis(side=1, at=seq(0,1, 0.1))
abline(h=nSims/bars, col = "red", lty=3)

set.seed(19900814)
lmrob_5_5Bs <- lmrob(Tariff_c ~ Regime_Bi_c + log(GDP_c) + EC_c + WTO_c, data=mydata_5, control =
summary(lmrob_5_5Bs)$coefficient #-3.31

simlmrobtest_5Bs <- summary(lmrob_5_5Bs)$coefficient[2,1]
simlmrobtest_5Bs #-3.31

#Imagine we re-run the "experiment" by shuffling the country/treatment
newExp_lmrob<-function(y){
  ## A function to randomly assign treatment effect
  shuffledz <- sample(mydata_5test$Z)
  newlmtest <- coef(lmrob(y ~ shuffledz + log(GDP_c) + EC_c + WTO_c, data=mydata_5test))["shuffledz"]
  return(newlmtest)
}

set.seed(19900814)
expBeta_lmrob_5Bs <- replicate(10000, newExp_lmrob(y=mydata_5test$Tariff_c))
mean(expBeta_lmrob_5Bs)

plot(density(expBeta_lmrob_5Bs)) #plot the null distribution
rug(mean(expBeta_lmrob_5Bs), 1, col="blue")
abline(v=simlmrobtest_5Bs,lty=2, col="red")

#One-sided p-value
pvalue_lm <- as.numeric(as.matrix(summary(expBeta_lmrob_5Bs < simlmrobtest_5Bs)))/length(expBeta_lmrob_5Bs)
pvalue_lm #0.0325

#Two-sided p-value
pvalue_lmrob_ts <- 2*min(mean(expBeta_lmrob_5Bs>=simlmrobtest_5Bs),mean(expBeta_lmrob_5Bs<=simlmrobtest_5Bs))

plmrob_can <- summary(lmrob_5_5Bs)$coefficient[2,4]
#different from the two-sided p-value obtained in the lmrob canned function (0.04).

#Codes to assess the false positive rate of the canned lmrob function

err_ratetlmrob<- function(y){
  require(robustbase)
  shuffledz <- sample(mydata_5test$Z)
  sim_psrob <- summary(lmrob(y ~ shuffledz + log(GDP_c) + EC_c + WTO_c, data=mydata_5test))$coef

```

```

return(sim_psrob)
}

set.seed(23358243)

nSims <- 1000

results_lmrob5Bs <- replicate(nSims, err_ratetlmrob(y=mydata_5test$Tariff_c))#get the p-value and

falsepos_lmrob5Bs <- mean(results_lmrob5Bs <= .05, na.rm = T) #0.055

falsepos_lmrob5Bs
#The Type 1 error, the false positive is 0.055.

#Check power by summing significant p-values and dividing by number of simulations

#The frequency of p-values, or the type 1 error is 0.055. When there is no effect (in the simulation)

bars<-20

op <- par(mar = c(5,7,4,4)) #change white-space around graph
hist(results_lmrob5Bs, breaks=20, xlab="P-values", ylab="number of p-values\n", axes=T,
      main=paste("P-value distribution under null effect in the simulation in lmrob"),
      col="grey", xlim=c(0,1), ylim=c(0, nSims))
axis(side=1, at=seq(0,1, 0.1))
abline(h=nSims/bars, col = "red", lty=3)

#install.packages("xtable",repos = "http://cran.us.r-project.org")
library(xtable)

lmp <- round(c(pvalue_lm_ts, p_lmcan, falsepos_lm5Bs),3)

lmrobp <- round(c(pvalue_lmrob_ts, plmrob_can, falsepos_lmrob5Bs),3)

pvalues <- as.matrix(rbind(lmp, lmrobp))

names_pvalues <- c("Permutation", "t-distritbuion", "False Positive Rate")
colnames(pvalues) <- names_pvalues

xtable(pvalues)

```

```

print(xtable::xtable(pvalues, caption = "P-values obtained from simulation, t-distribution and t
      html.table.attributes="border=1", comment = F)

## doParallel will not work on all machines.
library(foreach)
library(doParallel)
cl <- makeCluster(2) ## 2 cores on my machine
registerDoParallel(cl)

set.seed(800814)

power_fn<- function(outcome,effect){
  shuffledz <- sample(mydata_5test$Z)
  ## since we know that there is no effect, then we can assess the false positive rate of lm
  newoutcome <- outcome - shuffledz*effect
  sim_p<- summary(lm(newoutcome ~ shuffledz + log(GDP_c) + EC_c + WTO_c, data=mydata_5test))$co
return(sim_p)
}

#effectsize <- seq(0.1,5, by=0.3)
#results_sim_p <- replicate(1000, power_fn(outcome=mydata4_small$Tariff, effect=0.5))

#somepower <- mean(results_sim_p <= .05) #power
#somepower

#####

effectsize <- seq(0.1,5, by=0.3)
somepower <- sapply(seq(0.1,5, by=0.3), function(effectSize){
  results_sim_p <- replicate (1000, power_fn(outcome=mydata_5test$Tariff_c, effect=effectSize))

  power <- (mean(results_sim_p <= .05))
  return(power)
})

rbind(effectsize, somepower)

## doParallel will not work on all machines.
library(foreach)
library(doParallel)
cl <- makeCluster(2) ## 2 cores on my machine
registerDoParallel(cl)

set.seed(800814)

```

```

power_fn_lmrob<- function(outcome,effect){
  shuffledz <- sample(mydata_5test$Z)
  ## since we know that there is no effect, then we can assess the false positive rate of lm
  newoutcome <- outcome - shuffledz*effect
  sim_p<- summary(lmrob(newoutcome ~ shuffledz + log(GDP_c) + EC_c + WTO_c, data=mydata_5test))
  return(sim_p)
}

#effectsize <- seq(0.1,5, by=0.3)
#results_sim_p <- replicate(1000, power_fn(outcome=mydata4_small$Tariff, effect=0.5))

#somepower <- mean(results_sim_p <= .05) #power
#somepower

#####

effectsize <- seq(0.1,5, by=0.3)

somepower_lmrob <- sapply(seq(0.1,5, by=0.3), function(effectSize){
  results_sim_p <- replicate (1000, power_fn_lmrob(outcome=mydata_5test$Tariff, effect=effectSize))

  power <- (mean(results_sim_p <= .05, na.rm = T))
  return(power)
})

somepower_lmrob

#tablepowers <- rbind(effectsize, somepower, somepower_lmrob)

#install.packages("xtable",repos = "http://cran.us.r-project.org")
library(xtable)

tablepowers <- as.matrix(rbind(effectsize, somepower, somepower_lmrob))
xtable(tablepowers)

print(xtable::xtable(tablepowers, caption = "Simulation results of Power and Effect Size for lm a",
  html.table.attributes="border=1", width = "\\textwidth", comment=FALSE))

powersdata <- data.frame(effectsize, somepower, somepower_lmrob)

stacked <- with(powersdata,
  data.frame(
    power = c(somepower, somepower_lmrob),
    variable = factor(rep(c("powerlm", "powerlmrob"),
      each = NROW(powersdata))),
    effectsize = rep(effectsize, 2)))

```

```

ggplot(data=stacked, aes(x=effectsize, y=power, colour=variable))+
geom_line()+
geom_hline(yintercept=0.8, linetype="dashed", color="black")+
  labs(title="Power and Effect Size in Simulations", x="Effect Size", y = "Power") +
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5))

cat("\\\\newpage")
cat("\\\\newpage")
#install.packages("stargazer") If you haven't already installed. Install stargazer.
library(stargazer)
#Now use stargazer to format the linear model
stargazer(mydata_2_des_noc, header = FALSE, iqr=T,
          title = "Summary Statistics", digits = 1, out = "table1.txt",
          no.space = TRUE, font.size = "footnotesize")
cat("\\\\newpage")

```