First Draft of Final Paper

Lucie Lu

March 19, 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 make 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 section is an empirical analysis of a classic OLS approach with a discussion of statistical inference. It will move on to employ a combination of matching and regression to make a good comparison between country with and without political reform demonstrate the treatment effect of democratization.

[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 179 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.

Data and Measures

Dedependent 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 data set. To be fair, it is notoriously difficult to measure and hard to find time-series and cross-sectional data on this measure.

Dependent Variable

Regime (Polity Score) is a 21-poinit 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).

Dictator is a measure of the variations of autocracies differentiated by single-party regime, military regime and personalist regimes. It is coded as 1-8 with various characteristics of a regime type.

Control Variables

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.

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 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 a conditionality of foreign aid.

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.

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 follows in Table 1.

Table 1: 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

Section II: Empirical analysis: The OLS Model

Our central dependent variable is a country' trade policy in year t. We want to explain the changes of the extent of openness of the trade regime of the country on average. We use two ways of measuring trade policy. The first measure is a country's unweighted average statutory tariff rate (*TARIFF*). A decline of tariff rates indicates a more liberalized trade policy of a country. The mean of the statutory tariff rate is 21% between 1980 to 1999. However, it is poorly measured. This measurement may not be valid or reliable to capture the latent variables of the trade policy concept that I am interested in. In the periods of 1980s and 1990s, I still have 2683 missing values (14% of the total). I may end up drawing inaccurate inference due to the measurement problem.

My central independent variable is the type of political regime in place in a country in year t. As noted, the Polity Score ranging from -10 to 10 captures the institutional differences between democracies and autocracies regarding the fairness and openness of political competition and the size of the electorate. The mean of *Regime (Polity Score)* is -1.1, which indicates on average over this period a majority of these developing countries have mixes of democratic and autocratic features. It is later converted as a dichotomous variable *Democracy* with countries scoring below 6 in the Polity Score as autocracies (=0) and those above 6 as democracies (=1).

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 provide evidence of a low possibility of the collinearity issues in the later regression analysis.

Bivariate Relationship bewtween Regime Type and Tariff Rate
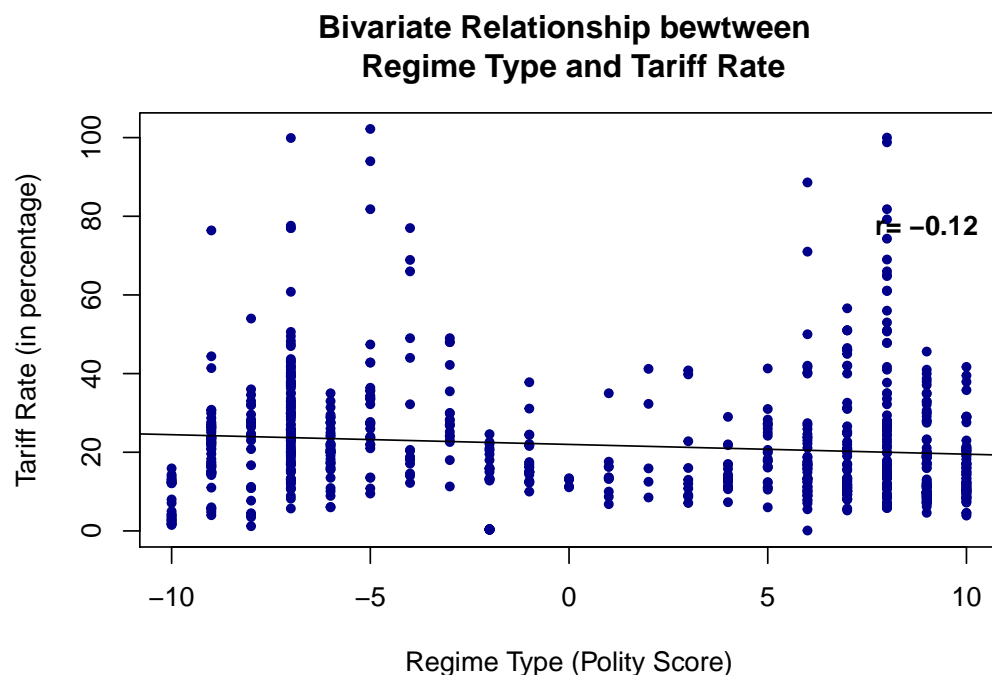
This figure tells me two important problems of the data structure in this dataset: 1) There are outliers in the

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 data set. 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.

Tariff variations across countries

Tariff variations across Years
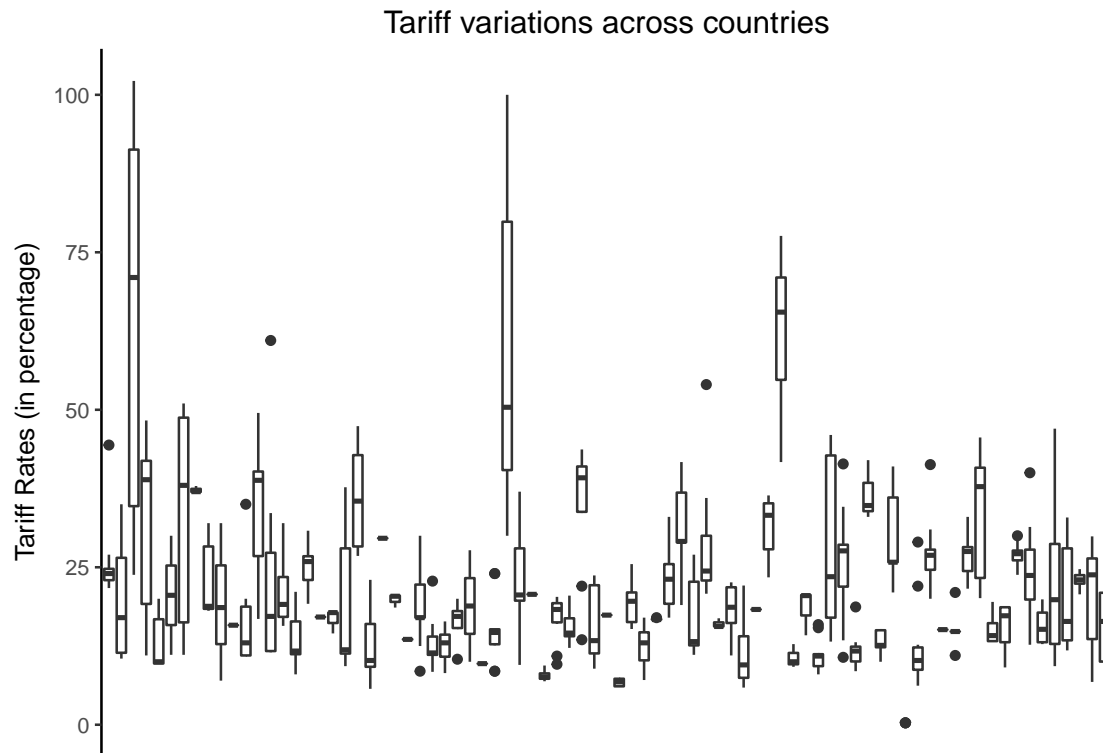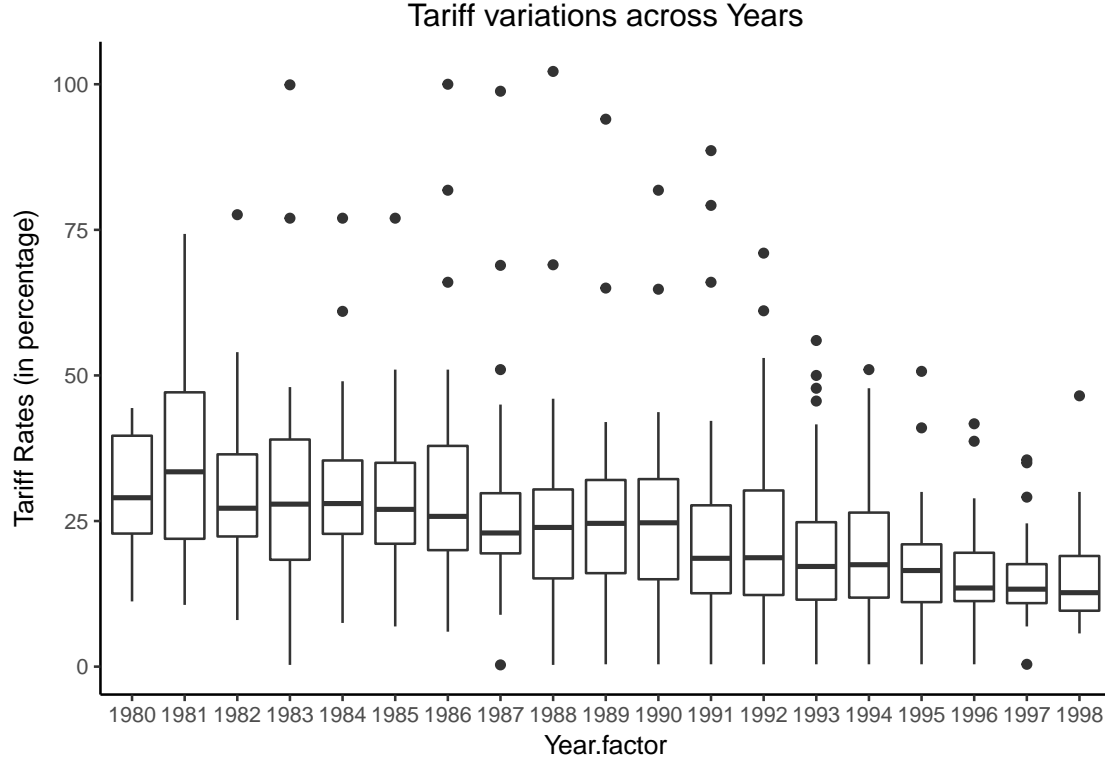
I calculated across-country variance (*123.7*) and within-country-years varianceis (*70.71*). Across-country variance is around 1.7 times of the within-country-years variance. It means that differences across countries are higher than differences across years within each country. I also calculated an ICC score: It is a method of Simple Estimation of the Intraclass Correlation Coefficient. Interclass correlation is the relative values of country- and year-level variances variances (Gilman & Hill, p. 258). The variance within countries is quite high with ICC score *0.63*. In this case, the interclass correlation *0.63* suggests that differences acorss 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 yearly tariff data for 81 developing countries over 20-year period. Tariffs variations across countries are larger than vairations across years within a country. Therefore, later on, I collapse the data to 81 data points so I can know what the mean tariff rate (and other information) is for each country. Here, country is the grouping variable.

OLS Models

According to Achen (2005)'s Rule of Three, 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 control of these relevant covariates, the treatment effect of democratization is less cofounded. I lag all of the independent variables by one year to some extent account for the endogeneity issue. The basic OLS equation estimating the relationship between democracy and trade policy is:

$tradepolicy_{i,t} = \beta_0 + \beta_1 * REGIME_{i,t-1} + \beta_2 * GDPPC_{i,t-1} + \beta_3 * ECCRISIS_{i,t-1} + \beta_4 * WTO_{i,t-1} + \beta_5 * OFFICE_{i,t-1} + \varepsilon_i$

Table 3 shows that on average more democratic regimes tend to have lower tariff rates. In almost all of the

regressions, regime type is correctly signed, as expected. In equation (5), an expected 0.13 unit decrease in tariff rate (in percentage) for every one unit increase in democracy in Polity Score. Tariff rates expect to drop from about its mean at 20.8 percent to about 20.67 percent when the regime polity score increases by point 1. Moving from an absolute autocracy (-10) to a perfect democracy (10) induces 2.6 percent decline in tariffs.

Table 2: Multiple Linear Regression Model: Results

	Dependent variable:				
	Tariff Rates				
	(1)	(2)	(3)	(4)	(5)
Regime (Polity Score)	−0.095	−0.075	−0.159**	−0.225**	−0.130*
	(0.072)	(0.074)	(0.074)	(0.093)	(0.075)
GDP per capita	−4.935***	−4.932***	−4.716***	−4.793***	−4.756***
	(0.431)	(0.440)	(0.433)	(0.453)	(0.441)
Economic Crisis		−8.041***			−7.878***
		(1.632)			(1.624)
GATT/WTO Member			4.267***		3.861***
			(1.275)		(1.276)
Years of Office				−0.238***	
				(0.077)	
Constant	57.310***	58.200***	52.390***	58.620***	53.870***
	(3.116)	(3.181)	(3.427)	(3.251)	(3.472)
Observations	774	765	774	729	765
R^2	0.166	0.183	0.177	0.170	0.193
Adjusted R^2	0.163	0.180	0.174	0.166	0.189
Residual Std. Error	13.760	13.590	13.670	13.880	13.520
F Statistic	76.460***	56.970***	55.380***	49.440***	45.470***

Note: *p<0.1; **p<0.05; ***p<0.01

Panel data may have problems that violate the standard assumptions necessary for ordinary least squares, so I will do statistical inference about an unobserved population by correcting for the standard errors with bootstrap method and panel-corrected robust standard errors in the next section. Before that, I will briefly discuss the advantages and potential problems of using the OLS.

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). I have checked this assumption in the various residual plots (figures are not shown). We also assume the unmeasured error terms are random. In a linear regression model, the dependent variable is a function of the observed independent variables and some unobserved factors represented by the disturbance term (Achen, 1982). These unobserved factors in the disturbance term need to be random. A different drawing from the disturbance may produce different values of dependent variable and hence different coefficients. This assumption will be revisited in a moment. 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 at the first place.

I am also 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 4 in Table 3) cannot be explained by the independent variable but can be explained by the rest of the confounders.

Let us try a different modelling strategy.

In regression modeling, lmrob is a robust statistics, which could be considered to handle unusual observations 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. The lmrob output from the lmrob function provides 'Robustness weights' to give us information about the outliers in observations by evaluating their weight and summarize statistics from the remaining observations.

Here, the coefficient sign for Democracy has flipped. Simpson's Paradox problem might exist: For example, if you have a categorical and continuous predictor then the continuous predictor could easily flip sign if the categorical one is added and within each category the sign is different than for the overall score. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2266743/)

may need to fix it later

Table 3: Multiple Linear Regression Model: Results

	Dependent variable:					
	Tariff Rates					
	MM-type linear		*OLS*			*MM-type linear*
	(1)	(2)	(3)	(4)	(5)	(6)
Democracy	−4.566***	0.093	0.813	−0.115	−1.369	−1.368
	(0.879)	(1.176)	(1.167)	(1.191)	(1.332)	(1.046)
GDP per capita		−4.583***	−4.727***	−4.556***	−4.490***	−2.889***
		(0.507)	(0.500)	(0.508)	(0.507)	(0.454)
Economic Crisis			−8.793***			−4.993***
			(1.859)			(0.919)
GATT/WTO Member				1.682		−0.584
				(1.527)		(1.014)
Years of Office					−0.174**	
					(0.075)	
Constant	22.260***	55.780***	57.340***	54.270***	57.110***	42.540***
	(0.592)	(3.516)	(3.477)	(3.773)	(3.552)	(3.330)
Observations	662	662	662	662	662	662
R^2	0.047	0.119	0.148	0.121	0.126	0.137
Adjusted R^2	0.046	0.116	0.144	0.117	0.122	0.132
Residual Std. Error	9.564	14.340	14.110	14.340	14.290	9.429
F Statistic		44.510***	38.090***	30.090***	31.640***	

Note: *p<0.1; **p<0.05; ***p<0.01

```
plot (mydata_5$Regime_c, mydata_5$Tariff_c, col = "darkblue", pch = 20, xlab = "Regime Type (Polity Sco
       main = "Bivariate Relationship bewtween \n Regime Type and Tariff Rate")
```

**Bivariate Relationship bewtween
Regime Type and Tariff Rate**

Tariff Rate (in percentage)

Regime Type (Polity Score)