PS 531 Final Project

Ji Soo Yoo

October 24, 2020

1. Research Question

To mitigate global climate crisis, international organizations and nation-states have set goals such as reducing greenhouse gas emissions and promoting sustainable development through various international agreements. Since the late 1800s, states have created about 700 multilateral agreements and more than 1000 bilateral agreements related to environment protection (Mitchell 2003). However, some nation-states take less time to ratify multilateral and bilateral environment agreements. On the other hand, other states take longer time to ratify and/or other states never ratify international agreements. This creates a situation in which states might agree to reduce emissions but then fail to make actual reductions. This is a common challenge for all international cooperation, often interpreted in terms of the Prisoners' Dilemma. More effective treaties will take these incentives into effect. To do that, we need to understand better what factors influence nation-states' ratification with international environmental agreements.

Since more international environmental agreements are expected in the future, it is important to understand what influences the behavior of nation-states in terms of ratifying with agreements, especially (Jacobson and Weiss 1990). In understanding which factors affect ratification of nation-states on international environmental agreements, academia has been studying characteristics of international agreements, role of domestic institution, domestic audience costs, economic condition, and scientific information and technology. Among these factors, this paper will focus on the role of domestic political institution. George Tsebelis (2002) argues that the factor of political institutions in determining ratification and compliance with international agreements can be understood in the context of veto players theory. Nation-states' constitution specifies whose consensus is required to change domestic policies from the status quo. In other words, consensus of veto-players is required to ratify international agreements. Tsebelis emphasizes the role of political institution since political power distribution among veto players such as presidents, prime minister and legislators set by their political institution determines ratification and compliance process of international environment agreements. In the addition, Lauren Peritz (2020) informs that more veto players imply more actors to block ratifying international agreements. Based on the theoretical framework of veto-players theory

for understanding the role of political institution, this paper explores an idea of *whether different types of domestic political institution determine duration of ratifying the Paris Agreement (international environmental agreements)*.

2. Research Design

The main argument of this research exploration is that nation-states with a parliamentary system takes less time to ratify international environmental agreements. In other words, the objective of this research exploration is to test whether parliamentary system tends to take less time to ratify international environmental agreements. My hypotheses are following:

H0. There is no association between political insitution type and duration of ratifying the Paris agreement.

H1. There is an association between political insitution type and duration of ratifying the Paris agreement.

2.1. Data Set

To test the above hypotheses, I conduct an observational analysis with using both of empirical data and simulated data.

First, the empirical data is based on a cross-sectional country data. I obtain the data on 2016 gross domestic product ('gdp') and carbon dioxide ('CO2') emission of 180 member states from the Paris Agreement through the World Bank Open Dataset. The year 2016 is selected since this when the Paris Agreement opened for signature. *Gross domestic product* ('gdp') informs economic progress and power of a country by measuring the final goods and services produced during the specific time period; a rising gross domestic product implies a good economic performance of a country. *Carbon dioxide* ('CO2') measures emissions from burning of fossil such as coal, natural gas, and oil. Based on the actual date from when the member state signed the Paris Agreement to the actual date when the member state ratified the Agreement, I created the variable called ('Ratification Time'). This variable measures how long a nation-state takes the time to ratify the Paris Agreement. The unit of measurement is a month.

Then, I simulate the explanatory variable, *different types of political institution* ('political inst') and the outcome variable ('time'). To assign different types of political institution, a country has a presidential system if it is assigned as 0. A country has a parliamentary system if it is assigned as 1. The outcome variable ('time') is the result of the manipulated ('Ratification Time') variable by a randomly simulated error term to the outcome variable. Figure 1 presents distributions of simulated population for each variable.

In summary, different types of political institution ('political inst') is the explanatory variable. Duration for ratifying the Paris Agreement ('time') is the outcome variable.
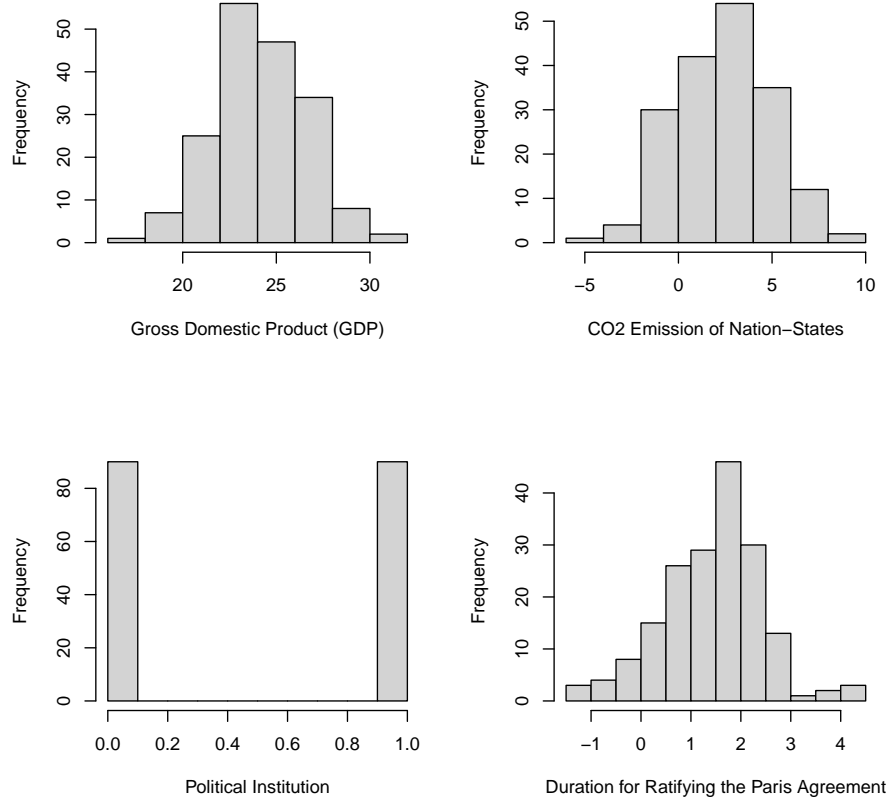
Figure 1: Distribution of the Simulated Data

3. Advantages and Disadvantages of Research Design

As stated in the previous section, this research conducts an observational analysis with cross sectional and simulated data. Although randomly assigned experiments are known as the gold standard, the observational study has few advantages. First, observational studies are quick, cheap, and simple to organize (Mariani and Pêgo-Fernandes 2014). Second, using the cross-sectional data of gross domestic product and carbon dioxide emission of countries captures the reality and provide appropriate contexts to conduct the observational analysis. Lastly, the observational analysis is applicable in studying an association between explanatory variables and an outcome variable, which can be considered as the first step to develop cause and effect analysis for the future study.

On the other hand, this research design of the observational study has some significant disadvantages. As I mentioned above, the observational study is useful to check an association between explanatory variables and an outcome

3

variable, but the study does not differentiate between cause and effect. In this sense, the study does not provide the information of causality. However, the observation study serves the purpose of this research note since its objective is to explore and check whether different types of political institution determines duration of ratifying the Paris Agreement. Second, this study has an internal validity problem. Compared to observational studies, randomized experiment balances on observed and unobserved confounding factors (Green and Gerber 2012) and allows researchers to claim that there is no influence of confounders on treatment effects. Hence, the study cannot conclude that observed and unobserved confounding factors are balanced. Lastly, the study does not capture the reality of political institution. In this research, political institution is a binary variable, 0 or 1, depending on whether a nation-state has a presidential or parliamentary system. However, types of political institution vary within a presidential system and parliamentary system in the reality. For example, nation-states with a parliamentary system can be a bicameral parliament or a unicameral parliament while nation-states with a presidential system can have multiple political parties or two political parties, which influences legislature powers of veto players. Hence, the simple binary assignment does not capture the reality of complicated political institutions.

4. Identification Strategy

The question of "what is your identification strategy?" can be rephrased as "what research deign and assumptions are you using?" (Keele 2015). It is important for observational study to address identification strategy and identify a causal estimand to claim that a statistical estimate can be considered as causal interpretations (Keele 2015). This is because observational study encounters *fundamental problem of causal inference* (Holland 1986) and generates identification problem based on counterfactual variables that researchers cannot observe (Keele 2015). To resolve identification problems and claim statistical to be given a causal interpretations, assumptions are needed and identified and then check if assumptions are held.

Accordingly, I provide the main identification assumption for the research design. I assume that observations are independent and identically distributed and the treatment assignment is randomly assigned conditional on observable. This assumption correctly holds because this research design uses the simulated data and the binary variable for assigning types of political institution was randomly assigned in the simulated data. This identification assumption is important because it allows researchers to use statistical adjustment method such as regression and matching if the assumption is held (Keele 2015).

Since the identification assumption holds correctly due to the simulated data, the research design looks at a simple OLS regression first for the statistical adjustment. The fundamental assumption of the regression model is that confounding relationships between variables are linear. Confounding relationship is the widely known threat to identification because confounding lead to a spurious relationship. If the second assumption holds, I can adjust for variables

4

such as 'emission' and 'gdp', also known as control variables. Adjusting for the control variables removes the confounding relationships not only between control variables and outcome variables but also between control variables and explanatory variables. If the linear relationship is removed among variables, the result of linear regression models reflects the effects of different types of political institution on duration of ratifying the Paris Agreement.

However, this also means that adjusting for the control variables fails to remove the confounding relationship between the variables if the assumption is not held and possibility of bias caused by unobserved (omitted) variables still remains even if the assumption is correct. In other words, the linear relationship is not the correct relationship between variables. Hence, it is critical for the second assumption to be held to claim that my research design is theoretically clear and relevant.

To demonstrate that the assumption holds, I create plots to visualize the distributions between variables (Figure 2). The Figure 2 shows the strong linear relationship between outcome variable ('time') and explanatory variable ('political inst'). On the other hand, the Figure 2 shows the weak linear relationship between explanatory variables ('political inst') and other independent variables ('gdp' and 'emission').
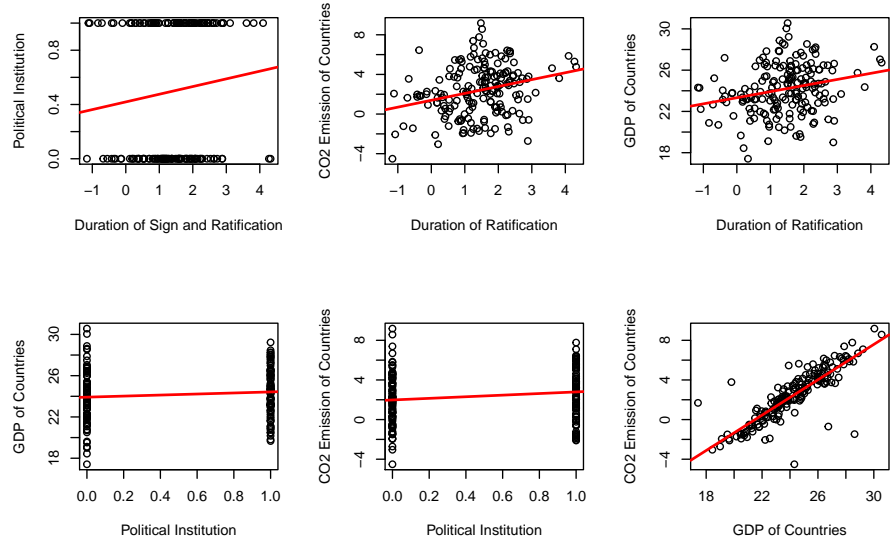
Figure 2: Linear Relationship between Variables

Also, the simple OLS regression model assumes that error variance is linear. This means that error variance is constant (or all error terms have the same variance). I check two conditions of whether the red line is approximately horizontal, which shows the average value of residuals is constant, and whether spread of values around the red line does not change with the fitted values. The Figure 3 shows that the red line is almost horizontal.
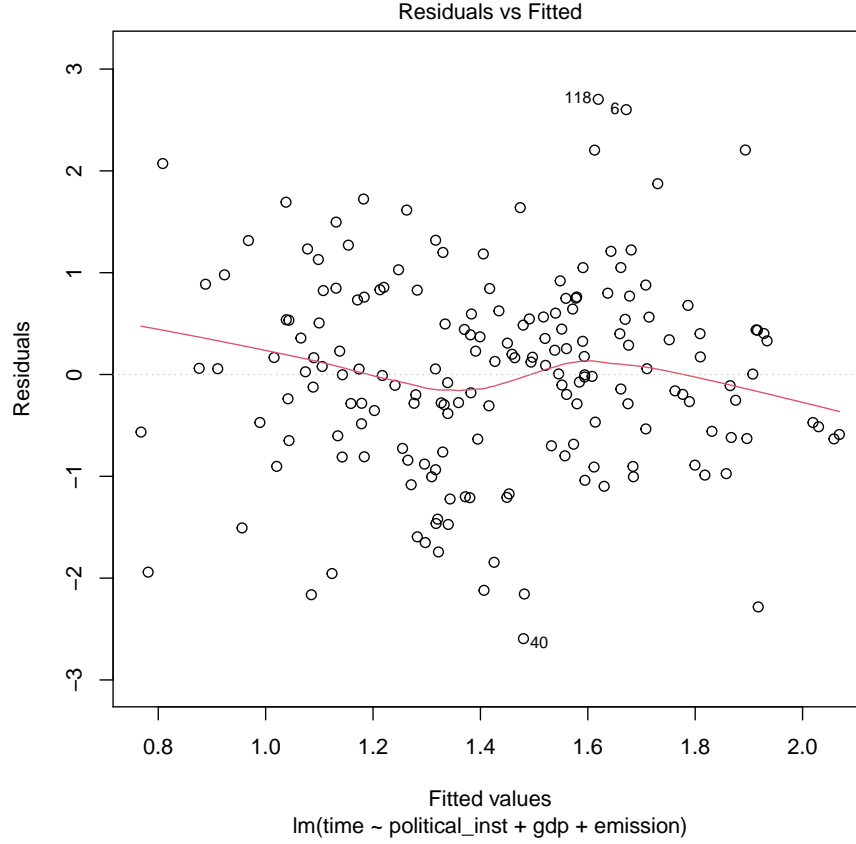
Figure 3. Residual Plot

To check the second condition, I conduct a Breusch-Pagan Test. The null hypothesis is homoskedasticity (the error variances are all equal/constant) and Table 1 shows that I fail to reject the null hypothesis at alpha=0.05 since the p-value is 0.7 (Table 1).

Table 1: **Results of Breusch-Pagan Test**

BP	df	P-Value
1.6	3	0.7

These analyses raise a concern of whether assumptions of the simple OLS regression are held and confounders are adjusted "enough". Hence, the second statistical adjustment method, propensity score matching, is conducted (please see the Appendix for codes). Propensity score matching reduces bias caused by covariates and allows researchers to compare outcomes between treatment

7

group and control group, which enables to estimate the effects of cause. Based on the similar propensity score, researchers can pair/match subjects in the control and treatment condition, which replicates randomized experiment (Rubin and Thomas 2000). Among various propensity score matching procedure, I choose near neighbor matching. The procedure pairs/matches subjects from the control and treatment condition based on closeness. I will discuss the results of matching to see if my adjustment strategy succeeds in the following section.

5. Judge the Success of the Adjustment Strategy

To judge the success of matching, I conduct balance testm which verifies the success of adjument (Rosenbaum 2010). The 'Balance Test' produces a standardized differences in means test between the control group (nation-states with a presidential system) and the treatment group (nation-states with a parliamentary system) on covariates within strata.

To judge whether matching worked (reduced bias and enabled the two treatment groups comparable by removing the association between covariates and treatment assignment), I check whether I can reject the null hypothesis of the balance test. The null hypothesis of Chi Square test is that there is no association between covariate and treatment assignment. Based on Table 2, I cannot reject at at alpha=0.05 the null hypothesis of that there is no association between the covariate and control groups and I cannot claim that performing a full match made the two treatment groups comparable. I cannot reject the null hypothesis since the p-value (0.076) is larger than the significance level (alpha=0.05).

Table 2: **Balance Test**

	Chi Square Statistic	Degrees of Freedom	P-Value
Full Match Data	5.2	2	0.076

This led me to compare the p-value of chisquare before and after matching. I found that the p-value of unmatched data is equal to the matched data (Table 3). Figure 4 and Figure 5 show that there is no difference between distribution of raw data and matched data. I do not know why matched did not work. My suspect is that I have the wrong covariates added to the model. However, I do not think this is the case and they are are possible confunders because economy scale and level of emission also influences nation-states when ratifying the Paris Agreement

Table 3: **Balance Test Before Matching**

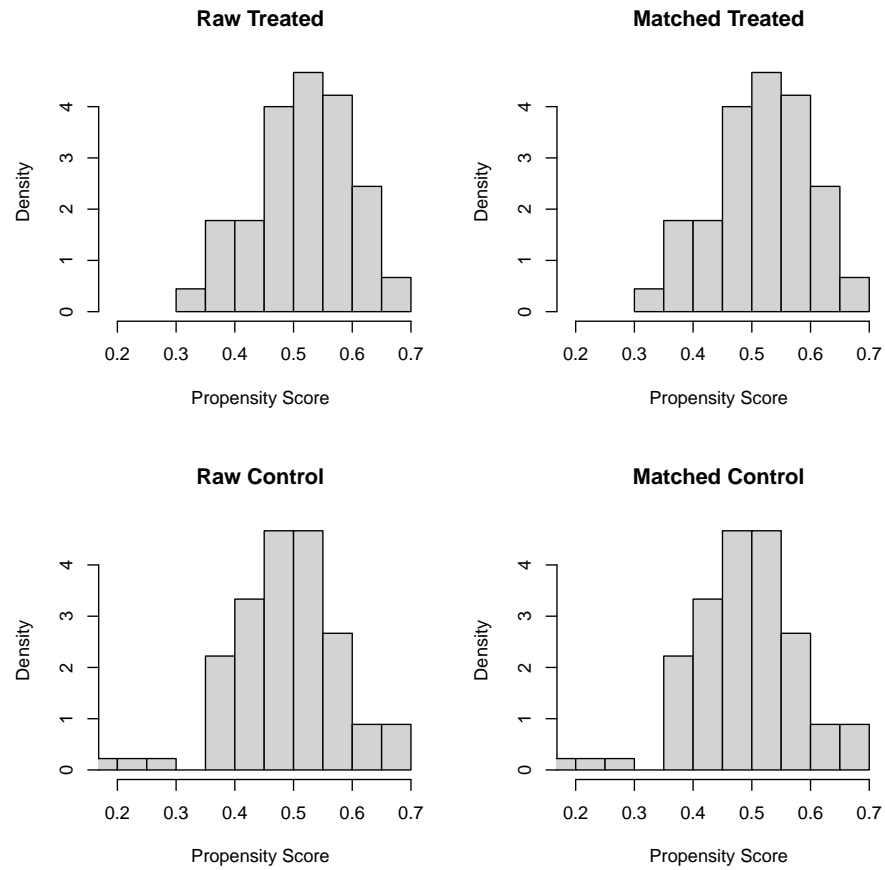|	Chi Square Statistic	Degrees of Freedom	P-Value
Unadjusted Data	5.2	2	0.076

Figure 4: Distribution of Control and Treatment Group Before and After Matching

Propensity Score Before Matching   **Propensity Score After Matching**
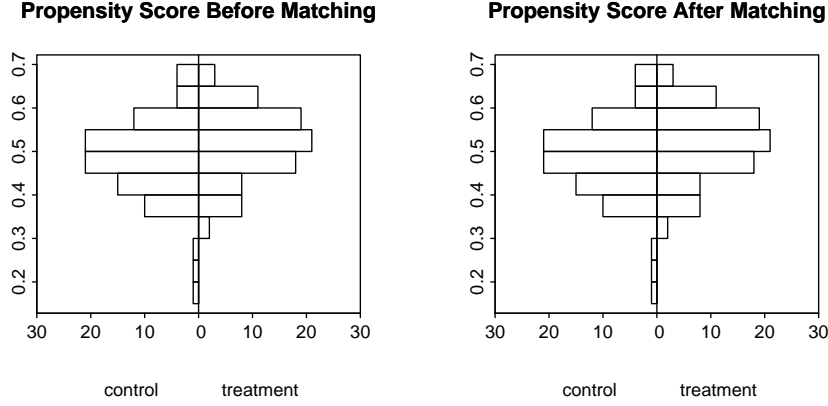

Figure 5: **Distribution of Control and Treatment Group Before and After Matching**

6. Statistical Tests

Statistical tests are used to determine whether a explanatory variable has a statistically significant relationship with an outcome variable by testing hypotheses or estimate the difference between two treatment groups. For the purpose of this research design, I use the null hypothesis at a significance level of alpha = 0.05 to estimate the difference between two treatment groups. In other words, I will reject the null hypothesis if the result is less than or equal to alpha=0.05. This significant level of alpha=0.05 is not a principle but it is a standard practice in the discipline of political science. First, I will use an independent sample z-test for hypothesis testing, which compares the means for two groups. Z-test is appropriate for the larger sample. For z-test, I assume that assumptions of normal distribution and i.i.d. are held for t-test since the data for this research design is simulated. The z-score of z-test informs the difference

10

between the mean average change in time for the treatment and control group; the difference between the two treatment groups gets larger as the z-score increases. Then, I will perform 1000 simulations and calculate the test statistic. Simulation is useful to assess the power of statistical test, the probability that the test will reject the null hypothesis when an alternative hypothesis ('H1'), is true.

7. Judge the Performance of Test Statistics

I will run false positive rate tests and power to judge the performance of each test statistics. False positive rate refers to the probability of rejecting the null hypothesis when the null hypothesis is true, also know as the Type I error. To claim an appropriate conclusion from the p-values of tests, the false positive rate should be less than or equal to the significant level of alpha=0.05. It is not appropriate or wrong to claim that I reject the null hypothesis at the significant level of alpha-0.05 based on the data since the false positive rate is greater than 0.05. In other words, the results of test statistics are misleading.

On the other hand, power refers to the probability of rejecting the null hypothesis when the alternative hypothesis is true. As the power of test increases, the type II error, which refers to the probability of accepting the null hypothesis when the alternative hypothesis is true. Based on false positive rate and power, I will claim that the test generates the correct conclusions if the false positive rate is less than the significant level at alpha=0.05 and the power of test is greater than 0.80.

8. Testing False Positive Rate and Power

The 'diagnose design' function from Declare Design package in R is a function to calculate bias, RMSE, coverage, and power through diagnose design. Declare Design simulates random-resampling of population by drawing new sample with replace and repeating the data generating population. The null hypothesis of this research design is that there is no relationship between types of political institution and duration of ratifying the Paris agreement. To check type I error rate (false positive rate) and type II error rate (false negative rate), I perform 1000 simulations of the data generating process. To check whether I reject the null hypothesis falsely based on the result of the test, I can look at the value of coverage in the result of diagnose design. Since I use two statistical estimators(linear regression estimator and robust linear regression estimator), I need to look at the coverage of both estimators. For a canned linear and robust linear model, the coverage value is the same (0.93). The value of coverage informs the probability of how many times I obtain a confidence interval that contains the true value if we repeat the data generating process. At the significance level of alpha=0.05, I cannot claim that the test is unlikely to cause a Type 1 error (but it is very close to the significance level). Hence, the tests are highly unlikely to reject the null hypothesis falsely. The value of power ranges between 0 and 1. As power gets closer to 1, I can claim that a estimator over another estimator is better to detect the false null hypothesis. However, for both test is very low

that is 0.15. The value of power indicates that I will reject the null hypothesis with the probability of 0.15 when the alternative hypothesis is true.

Power of a statistical estimator is affected by sample size, variance, and research design. As sample size gets larger, power of the statistical estimator gets better. Since the small sample size refers to around 30 objects, I think that the sample size (n=180) is not so small to decrease the power of the both statistical estimators. As a result, I highly suspect that the low rate of power is caused by variance and research design. In the previous sections, I demonstrate whether my research design adjusted enough for confounders to reduce variance based on the assumption of the linear relationship between variables. Then, because the result indicate that it is not adjusted "enough", I took the further step to reduce variance through matching. However, matching did not reduce bias and the p-value increased after matching strangely. Due to the failure to adjust for confounders, I think that I got the lower rate of power. I believe that the simulated data should not be the cause since it simulates random re-sampling of the population. This means that there is a problem with my raw data. Hence, I think that improving raw data could increase the rate of power.

Table 4: **Coverage and Power of Tests**

Estimator	Coverage	Power
Estimator 1. canned lm	0.93	0.15
Estimator 2. lm robust	0.93	0.15

Lastly, I wanted to check whether adding variables to adjust for decrease the rate of power when the variables are not really confounders. Hence, I re-diagnose the tests by removing confounder variables ('emission' and 'gdp') in linear regression model (estimator 3) and robust regression model of estimators (estimator 4). The estimators that exclude the original confounder variables increased the value of power by 0.01 for the both estimators, which means more possibility to reject the null hypothesis when the null hypothesis is false. However, this decreased coverage for estimaotr 3 by 0.01 and for estimator 4 by 0.02. This led me to think the problem of omitted variables. What variables am I missing to adjust for? However, as Achen's a rule of three (2002), adding confounding variables more than three has disadvantages and is absolutely not a solution. Hence, I have to re-think possible of confounder variables and explanations of choosing them as such variables.

Table 5: **Coverage and Power of Tests**

Estimator	Coverage	Power
Estimator 3. canned lm	0.92	0.16
Estimator 4. lm robust	0.91	0.16

12

9.Statistical Estimand

The objective of this research is to test whether some nation-states takes longer or less time to ratify the Paris Agreement, depending on their type of political institution. In other words, this test is to check whether there is an association between political institution type and duration of ratifying the the Paris Agreement. Hence, the target of estimation is B1 (a difference in duration of ratifying the Paris Agreement (**time**), depending on political institution type of nation-states (**political inst**) by simple linear regression.

$$Y_{time} = \beta_0 + \beta_1 * X_{political_inst} + \epsilon$$

Table 6: **Estimand 1**

	Estimand Label	Estimand
Political Inst	Political Inst	0.2257

With the above formula, I use two estimators. The first estimator is a simple OLS regression model. By default, the simple OLS regression estimate the mean difference in Ytime between nation states with a presidential system and nation-states with a parliamentary system. The estimator 1 includes covariates such as GDP and CO2 emission to adjust the confounding problems. The second estimator is robust linear regression. I use the robust linear regression to reduce the effect of outliers and heteroscedasticity since the Breusch-Pagan conducted above shows that my model has a heteroscedasticity problem. The robust linear regression allows researchers to capture the true pattern of the data by minimizing the absolute difference of influential observations. The second estimator also includes GDP and CO2 emission to adjust for confounders.

10. Performance of Estimators

To investigate the performance of these estimators, I evaluate the performance in terms of bias and root mean square error (RMSE). First, an evaluation of the performance is based on certain number of simulations. This means that it repeats the data generating process and calculate the estimate with each sample.

Second, bias is the difference between the estimator's expected prediction and the true parameter value. This means that well-performed estimators have the value of bias equal to 0. When the bias of estimators is equal to 0, this means that the peak of the distribution of samples indicates the true parameter value, the true treatment effect. Hence, I can claim that the estimator performs well if the estimate is centered on the true parameter of political institution.

Observed values of the data is used to predict or estimate unknown values within the range of the data (interpolation). In order to estimate the unknown values correctly, the data fitted to the model well is required, which is based on the correct assumption about the relationship between variables in the data. To check whether the data is fitted well in the model, I can check the value

13

of RMSE. RMSE calculates the stand deviation of the residuals (error terms). As the value of RMSE gets larger, the poorer performance of the estimators is. Hence, I can claim that my estimators perform well if they have the lower values of RMSE, which indicates the better fit, and the ideal RMSE is 0.

11. Bias and RMSE of Estimators

To calculate the value of Bias and RMSE, I use the 'diagnose design' function from Declare Design package in R. The evaluation for the performance of the estimate with two estimators is based on 1000 simulations. Bias of the estimate with each estimator 1 and estimator 2is measured by calculating the difference between the true treatment effect and the mean estimate of treatment effect. The bias for each estimator 1 and estimator 2 is -0.08. This means that the estimators of my model tend to underestimate the parameter. Hence, I cannot claim that the estimate with estimator 1 and estimator 2 reflects the true value of population parameter.

As mention in the previous section, the value of RMSE tells whether the data is fitted well in the model. The ideal of RMSE is 0, which means that observed value of the data is equal to the value predicted by the model. However, Table7 shows that the estimate with estimate 1 and estimate is 0.17. Although the value of RMSE is relatively small, I cannot confidently make claims following: 1) the data is well fitted in model, 2) the model assumes the relationship between variables correctly, and 3)the the estimate with two estimators reflect the true relationship.

Table 7: **Coverage and Power of Tests**

Estimand	Estimator	Bias	RMSE
Political Inst	canned lm	-0.08	0.17
Political Inst	m robust	-0.08	0.17

12. Interpretation of the Z-test Statistic Results

I perform the Z-test statistic (coefficient divided by its standard error) based on 1000 simulations. Table 8 shows the result of the Z-test statistic of estimator 1 and estimator 2. The mean test statistic of estimator 1 is 1.01 and the mean p-value is 0.37. Strangely, the mean test statistic for estimator 2 is the same to estimator 1; the mean statistic of estimator 2 is 1.01 and the mean p-value is 0.37 as well; at a significance level, alpha=0.05, 0.37 is larger than 0.05. Hence, I cannot reject the null hypothesis of that there is no associations between type of political institution and duration of ratifying the Paris Agreement. Again, the result of the z test statistic has to do with the failure of propensity score matching between covariates.

Figure 6 shows the distribution of Z-test statistic results of estimator 1 while Figure 7 shows the distribution of Z-test statistic results of estimator 2. The red line is the mean of Z Test Statistic. The According to Figure 6 and Figure

	Table 8: **Mean of Z Test Statistic Results**		
Estimand	**Estimator**	**Z Test Statistc**	**P-Value**
Political Inst	canned lm	1.01	0.37
Political Inst	lm robust	1.01	0.37

7, the Z-test statistic of estimator 1 and estimator 2 is normally distributed. This implies that the assumptions to use Z test statistic, as mentioned above, are held correctly. Hence, I do not know why my p-value is large. Again, this might has to do with the failure of matching, which I do not understand since my data is generated based on simulations.
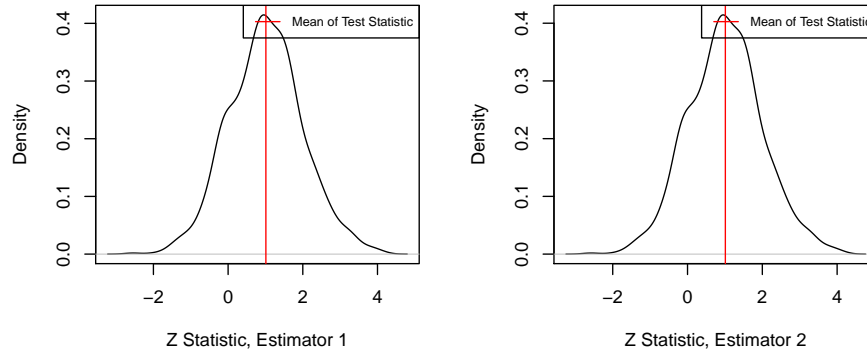
Figure 6: Z test statistic, Estimator 1 and Estimator 2

13. Interpretation of Results of Estimators

Table 9 shows the regression results of estimator 1 and estimator 2. The estimator 1 is based on canned linear model. The estimated coefficient of the

Table 9: Regression Results

	Dependent variable:	
	time	
	OLS	robust linear
	(1)	(2)
political_inst	0.146	0.208
	(0.146)	(0.141)
emission	0.084	0.053
	(0.054)	(0.052)
gdp	0.023	0.039
	(0.056)	(0.054)
Constant	0.594	0.259
	(1.264)	(1.214)
Observations	180	180
R^2	0.081	
Adjusted R^2	0.065	
Residual Std. Error (df = 176)	0.968	0.911
F Statistic	5.145*** (df = 3; 176)	
Note:	*p<0.1; **p<0.05; ***p<0.01	

16

canned linear model (estimator 1) is 0.146. This is the difference in the duration of ratifying the Pairs Agreement associated with type of political institution. In terms of the substantive question, the canned linear model tells that a nation-state with a parliamentary system (assigned as 1 for "political inst" variable) takes 0.146 (measured in a month) longer to ratify the Paris Agreement.

The estimator 2 is based on the robust linear regression model. The estimated coefficient of the robust linear regression model (estimator 2) is 0.208. This is the difference in the duration of ratifying the Paris Agreement associated with type of political institution. In terms of the substantive question, the robust linear regression model tells that a nation-state with a parliamentary system (assigned as 1 for "political inst" variable) takes 0.208 (measured in a month) longer to ratify the Paris Agreement.

As I mentioned in the previous section, the mean of p-value for estimator 1 and estimator 2 is larger than the significance level at alpha=0.05 and as a result, I cannot reject the null hypothesis of that there is no association between type of political institution and duration of ratifying the Paris Agreement.

14. What I Learned about the Substantive/Theoretical Topic

The objective of this research note is to explore whether there is an association between type of political institution and duration of ratifying the Paris Agreement, which is my null hypothesis. This idea was driven mainly by Robert Putnam's Two Level Game Theory and George Tseberlis' Veto Players' Theory. These theories emphasizes domestic politics and the role of domestic political institution. Generally, these theories argue that different types of political institution creates different number of veto players who decides to change policy and legislation process of signing and ratifying international agreements. However, I cannot claim that there is an association between type of political institution and duration of ratifying the Paris Agreement because the result of my statistical fails to reject null hypothesis. Hence, I need to explore what went wrong and how to revise my research design for the further exploration, which I will discuss in the next section.

15. What I Learned about the Methods and Research Design

For the purpose of this research note, I simulated data based on the real data of GDP, CO_2 emission, and time that each member of the Paris Agreement took to ratify the agreement after the nation-state signed the agreement. Then, I randomly assigned the treatment effect (type of political institution), a presidential system as 0 and a parliamentary system as 1 to explore whether there is an association between type of political institution and duration of ratifying the Paris Agreement—the null hypothesis.

However, as I mentioned in the previous section, the results of my statistical model failed to reject the null hypotheses. The power of the statistical model was very low and the p-value was large compared to the significance level, alpha=0.05. I suggest possible area, where it went wrong. First, we learned that it is important for researchers to address what assumptions are made for a sta-

tistical model, why we make such choice, and how we justify our choices. To use the OLS estimators, I assumed normal distribution and i.i.d among variables to use the OLS estimators since my data is simulated and the treatment effect was randomly assigned. To check whether my assumptions are held correctly, I check the distribution of the data and linear relationship among variables. Because the linear relationship between certain variables is weak, I tried to adjust the issue by using the propensity score matching. We learned that the propensity score matching is supposed to reduce bias and make variables comparable by matching variables based on the similar propensity score between variables on covariates. However, I found that there was no difference between unmatched and matched data; it did not adjust the issue even after matching. Hence, I think that this is the area, where it went wrong but I have no idea what I did wrong.

From my intuition, I think there is something wrong with raw data to begin with. Although the power of the statistical model improves as a sample size increases, this is not possible given that I wanted to look at member states of the Paris Agreement. Hence, I need to focus and examine why matching did not work.

16. Include a code appendix and a link to the github repository for this paper.

GitHub Repository: https://github.com/jisoosy2/Quant2

```
library(knitr)
library(tidyverse)
library(coin)
library(readxl)
library(DeclareDesign)
library(DesignLibrary)

#import data set.
dat1 <- read_excel("~/Desktop/2020 Fall /Quant2_Final/FinalProject_Yoo.xlsx")

#clean data set
drop<-c("Signature", "Ratification")
df=dat1[,!(names(dat1)%in%drop)]
df1 <- df %>% drop_na()
df1$emission=log(df1$CO2)
df1$gdp=log(df1$GDP)
df1$CO2 <- NULL
df1$GDP <- NULL

#create a population based on the real data set that I imported.
pop <- declare_population(df1)
```

18

```r
assignment <- declare_assignment(assignment_variable = "political_inst")
outcomes <-declare_potential_outcomes(time~political_inst*tau+.05*emission+.05*gdp+
                                    rnorm(N, mean=0, sd=1), assignment_variables = "pol:
treatment_outcomes <- declare_reveal(outcome_variables = "time",assignment_variables = "pol:
design <-pop+assignment+outcomes+treatment_outcomes
design1 <- redesign(design, tau=0.15)

set.seed(12345)
df2 <- draw_data(design1)


#plot the distribution of the simulated data
par(mfrow=c(2,2))
hist(df2$gdp, main="", xlab = "Gross Domestic Product (GDP)")
hist(df2$emission, main = "", xlab = "CO2 Emission of Nation-States")
hist(df2$political_inst, main = "", xlab = "Political Institution")
hist(df2$time, main = "", xlab = "Duration for Ratifying the Paris Agreement")

library(ggplot2)

#show linear relationship between variables
par(mfrow=c(3,3))
plot(df2$time,df2$political_inst,
    xlab= "Duration of Sign and Ratification",
    ylab= "Political Institution",
    )
abline(lm(political_inst~time, data=df2), col="red", lwd=2)
plot(df2$time, df2$emission,
    xlab="Duration of Ratification",
    ylab="CO2 Emission of Countries")
abline(lm(emission~time, data=df2), col="red", lwd=2)
plot(df2$time, df2$gdp,
    xlab= "Duration of Ratification",
    ylab= "GDP of Countries")
abline(lm(gdp~time,data=df2), col="red", lwd=2)
plot(df2$political_inst,df2$gdp,
    xlab= "Political Institution",
    ylab= "GDP of Countries",
    )
abline(lm(gdp~political_inst, data=df2), col="red", lwd=2)
plot(df2$political_inst,df2$emission,
    xlab= "Political Institution",
    ylab= "CO2 Emission of Countries",
    )
abline(lm(emission~political_inst, data=df2), col="red", lwd=2)
```

```r
plot(df2$gdp,df2$emission,
     xlab= "GDP of Countries",
     ylab= "CO2 Emission of Countries",
     )
abline(lm(emission~gdp, data=df2), col="red", lwd=2)


#check homskedasticity
lme1 <- (lm(time~political_inst+gdp+emission, data=df2))
library(lmtest)
bptest(lme1)


#plog the residual plot
plot(lm(time~political_inst+gdp+emission, data=df2), which=1)


library(MASS)
library(optmatch)
set.seed(12345)
#match on treatment assignment and covariates

#calculate propensity score
ps <- glm(political_inst~emission+gdp, data=df2, family=binomial())

#addd the predicted propensity score to the data
df2$psvalue <- predict(ps, type="response")

library(MatchIt)

#estimate near neighbor using the 1:1 ration (default)
#match using near~neighbor
m.nn <- matchit(political_inst~emission+gdp, data=df2, method="nearest", ratio=1)
summary(m.nn)
match.data=match.data(m.nn)



#showing covariate imbalance before matching
treated <- (df2$political_inst==1)
cov <- df2[,2:9]
std.diff <- apply(cov, 2, function(x) 100*(mean(x[treated])-mean(x[!treated]))/(sqrt(0.5*(va

library(RItools)
```

```r
#balance talbe before matching
xBalance(political_inst~emission+gdp, data=df2, report=c("chisquare.test", "p.values"))

#balance table after matching
xBalance(political_inst~emission+gdp, data=match.data, report=c("chisquare.test", "p.values"

#plot the distribution of the matching dataset.
plot(m.nn, type="hist")

library(Hmisc)

#plot the distribution of before and after matching to compare.
par(mfrow=c(2,2))
histbackback(split(df2$psvalue, df2$political_inst), main="Propensity Score Before Matching"
histbackback(split(match.data$psvalue, match.data$political_inst), main="Propensity Score Af

set.seed(12345)
df2 <- draw_data(design1)

#estimating the effect of type of political institution
make_estimands <-function(data){
  bs <- coef(lm(time~political_inst, data=df2))
  return(data.frame(estimand_label=c('political_inst'),
                  estimand=bs[c('political_inst')],
                  stringsAsFactors = FALSE))
}
estimand <- declare_estimands(handler = make_estimands,
                              label = "Pop_Relationships")

design1_plus_estimands <- design1 + estimand
draw_estimands(design1_plus_estimands)

#declare estimators
estimator1 <- declare_estimator(time~political_inst+emission+gdp,
                                model=lm,
                                term=c('political_inst'),
                                estimand=c('political_inst'),
                                label="canned_lm")
design1_lm <-design1_plus_estimands +estimator1
design1_lm

reg1<-lm(time~political_inst+emission+gdp, data=df2)
```

```r
#declare estimator2
estimator2 <- declare_estimator(time~political_inst+emission+gdp,
                                model=lm_robust,
                                term=c('political_inst'),
                                estimand=c('political_inst'),
                                label="lm_robust")
design2_robust<- design1_plus_estimands+estimator2
design2_robust

reg2<-rlm(time~political_inst+emission+gdp, data=df2)


#simulate and perform diagnose tests.
designtotal <- design1_plus_estimands+estimator1+estimator2
diagtwo <- diagnose_design(designtotal, sims=1000)
diagtwo


#declare additional estimators to see if changes makes any differences for the value of pow
estimator3 <- declare_estimator(time~political_inst,
                                model=lm,
                                term=c('political_inst'),
                                estimand=c('political_inst'),
                                label="lm")

estimator4 <- declare_estimator(time~political_inst,
                                model=lm_robust,
                                term=c('political_inst'),
                                estimand=c('political_inst'),
                                label="lm_robust")

#then, perform the dignose test for the additional estimators.
designtotal2 <- design1_plus_estimands+estimator3+estimator4
diagthree <- diagnose_design(designtotal, sims=1000)
diagthree


diagtwo


<result1>>
```

```
#make a subset data to plot
newdata<-subset(sim, estimator_label=="canned_lm", select=c(statistic))
newdata2<-subset(sim, estimator_label=="lm_robust", select=c(statistic))


#plot the distribution of Z-test statistic of estimator 1 and estimator 2.
par(mfrow=c(2,2))
plot(density(newdata$statistic), main="", xlab="Z Statistic, Estimator 1")
abline(v=mean(newdata$statistic), col="red")
legend("topright", c("Mean of Test Statistic"), lty=1, col="red", cex=0.75)
plot(density(newdata2$statistic), main="", xlab="Z Statistic, Estimator 2")
abline(v=mean(newdata2$statistic), col="red")
legend("topright", c("Mean of Test Statistic"), lty=1, col="red", cex=0.75)


library(stargazer)
#stargazer(reg1, reg2)
```

17. Include references with appropriate in-text citations.

References

Gerber, Alan S., and Donald P. Green. *Field experiments: Design, analysis, and interpretation.* WW Norton, 2012.

Holland, Paul W. "Statistics and causal inference." *Journal of the American Statistical Association* 81.396 (1986): 945-960.

Jacobson, Harold K., and Edith Brown Weiss. "Implementing and complying with international environmental accords: A framework for research." *annual meeting of the American Political Science Association, San Francisco.* Vol. 30. 1990.

Keele, Luke. "The statistics of causal inference: A view from political methodology." *Political Analysis* (2015): 313-335.

Mariani, Alessandro Wasum, and Paulo Manuel Pego-Fernandes. "Observational studies: why are they so important?." *Sao Paulo Medical Journal* 132.1 (2014): 01-02.

Mitchell, Ronald B. "International environmental agreements: a survey of their features, formation, and effects." *Annual review of environment and resources* 28.1 (2003): 429-461.

Peritz, Lauren. "When are International Institutions Effective? The Impact of Domestic Veto Players on Compliance with WTO Rulings." *International Studies Quarterly* 64.1 (2020): 220-234.

Rosenbaum, P. R. (2010). Design of observational studies. *Springer series in statistics.*

Rubin, Donald B., and Neal Thomas. "Combining propensity score matching with additional adjustments for prognostic covariates." *Journal of the American Statistical Association* 95.450 (2000): 573-585.

Tsebelis, George. *Veto players: How political institutions work.* Princeton University Press, 2002.