The Effect of Oil Wealth on Political Violence

Wanran Zhao, Cole Silva, Tony Zong

I. Introduction

BUILDING on the Malthusian theory that resource scarcity triggers conflict, many scholars argue that possession of resources, especially oil, makes developing countries vulnerable to political violence. For instance, oil-rich countries like Angola and Chad witnessed high incidences of civil conflicts. In response to this argument, (Cotet and Tsui 2013) examined whether oil abundance causes political violence measured by civil wars, coup attempts, irregular leadership transitions, and military spending. They showed that oil abundance does not affect political violence in general, whereas oil abundance increases military spending among non-democratic countries.

In our paper, we use the Augmented Inverse Propensity Weighted Estimator (AIPW) to replicate the findings of (Cotet and Tsui 2013). We examine the effect of two treatments, changes in oil reserves and oil discoveries, on numerous measures of political violence outcomes. Our results confirm what (Cotet and Tsui 2013) found in their studies, which is that there is no causal effect of oil abundance on conflict.

This paper will proceed as the following. Section II discusses our data source, choice of treatments and outcomes, as well as variable transformations. Specifically, we talk about how our data fit the exogeneity and i.i.d. assumptions. We also provide a preliminary analysis to check the general correlation between our treatment and outcomes. Section III provides the mathematical background for our model and how our model satisfies the Overlap and No Unobserved Confounders Assumptions. In particular, we use ATT instead of ATE estimates to relax the Overlap Assumption. We also argue that despite that we may not exhaust all confounders, our estimates should provide an upper bound for the interested casual effect. Section IV presents our results. Section V includes sensitivity analysis and estimation results after taking out USA and USSR as outlier countries. We give our conclusion in Section VI.

II. Data Sources and Assumptions

The data cover oil reserves, oil exploration, measures of domestic conflict and violence, and a set of control variables of 103 countries over the 1930-2003 period. The data on oil reserves is from the Association for the Study of Peak Oil and Gas (ASPO); the domestic conflict data is from Gleditsch's revision of the Correlates of War (for war status), the Center for Systemic Peace (for military coup attempts), Archigos dataset of political leaders (for irregular leadership transitions), and Stockholm International Peace Research Institute (for defense burden). Additionally, we have a set of control variables from various sources extracted by (Cotet and Tsui 2013). A detailed list of controls is given in table II and further discussed in section III.B.

A. Treatment Choice

We choose two treatments for their different advantages: oil reserves and oil discoveries. We have two measures for oil reserves: simply oil reserves, or value of oil reserves (oil reserves times inflation-adjusted crude oil price). Oil discovery is measured in thousand million barrels. All treatments are on a per capita basis. On one hand, as a stock variable, oil reserves represent a country's wealth in oil and therefore serve as a close proxy for oil resource abundance that we are interested in. However, oil reserves suffer from reverse causality. Political violence and stability affect oil reserves through many factors, like oil exploration decisions, oil sales, and oil spills' lack of public maintenance. Therefore, oil reserves may violate the exogeneity assumption.

On the other hand, oil discovery is a flow variable and does not fully capture a country's oil abundance. Nevertheless, oil discoveries are only affected by political stability through oil exploration decisions [1]. After the exploratory borehole is drilled, the success of oil discovery is a relatively random hit-or-miss business that only depends on geography. For this reason, we assume that oil discovery is exogenous, conditional on exploration efforts and geographical conditions. In our analysis, we include the number of wildcat[2] drilled, the share of mountainous land in the territory, and many other country-specific conditions listed in section III.B. Hence, while oil reserves are a better measure of oil abundance, oil discovery plausibly satisfies the identification assumption of exogeneity.

B. Variable Transformation

Time series data is often auto-correlated. Therefore, we cannot treat observations from the same country in different years as independent and identically distributed. To reduce the correlation between across-time observations, for time-variant variables, we transform them into their percentage yearly change for each country. We then use the relative changes, instead of the absolute magnitude, as our variables in our machine learning models.

The variables for which we calculate the percentage increase include oil reserves, crude oil price (adjusted for inflation), value of oil reserve, population, population density, democracy index, natural disaster, and the number of wildcats drilled (a proxy for oil exploration). For the defense burden, which is the military spending out of the country's GDP, since it's already a proportion, we decide to directly calculate the difference

[1] For instance, according to (Robinson, Torvik, and Verdier 2006), political leaders tend to be myopic and heavily discount the future during irregular leadership shifts. Therefore, in politically unstable circumstances, leaders tend to overly extract oil resources.

[2] Wildcat is the first exploratory borehole in a new area of oil exploration.

between consecutive years. Since the treatment variables (oil reserve and value of oil reserve) are continuous, we also binarize them, with 1 being an increase of 10% or more than the previous year for each country and 0 otherwise. We choose 10% as the cutoff because smaller cutoffs, such as 0, can undermine the potential treatment effect since a hypothetically extremely small increase in oil reserve would have little effect on domestic violence in the country.

For other country-fixed-effects variables including share of mountainous land, ethnicity, religion, and language fractionalization, we don't transform them, because they are relatively time-invariant for each country.

C. Outcome Choice

No single measure can capture all types of political violence, so we chose to have a list of outcomes against which we could test our different treatments. For oil reserves, outcomes include war onset, military coup attempts, defense burden, and irregular leadership transitions. For oil discovery, we expect a more lagged effect, since it would likely require some time for the new discovery of oil to have an impact on any sort of political violence. Therefore, we include two additional outcomes of future violence for oil discovery. They are within-2-year change in war status and within-3-6year change in war status.

For preliminary analysis, Table I reports the results of t-tests under the hypothesis that the outcomes from treatment and control groups have the same mean. The only significant[3] mean difference comes from irregular leadership transition correlated to value of oil reserves. Following the results in the original paper, we expect to see no significant results between either pair of our treatments and outcomes.

TABLE I: Preliminary t-test. We include mean differences (treatment mean-control mean) divided by 1e-02 and p-values in parenthesis from t-tests. Our treatments are binarized oil reserves, the value of oil reserves, and oil discovery. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

	oilres	**valoilres**	**newdiscovery**
war onsets	0.234	0.532	1.180
	(0.730)	(0.900)	(5.271)
military coup attempts	-0.806	-0.951	0.687
	(0.637)	(0.415)	(0.329)
irregular leadership transitions	-0.818	5.688***	-5.020
	(0.700)	(0.000)	(2.107)
defense burden change	230.527	39.479	8.138
	(0.325)	(0.250)	(0.548)
war status change			0.952
(in 2 years)			(0.377)
war status change			1.821
(in 3-6 years)			(0.184)

III. Modeling for Treatment Effect Estimation

A. Estimands and Estimators

In this section, we will explore how we arrived at our estimates for the Average Treatment Effect (ATE) and Average Treatment Effect on the Treated (ATT) of oil abundance on conflict. First, we will provide some background on the estimation methods used. ATE is the difference in means between the group that is treated (above average oil abundance) and the group that is untreated (below average oil abundance), with the estimand for ATE being

$$\mathrm{ATE} = \mathrm{E}[\mathrm{Y} \mid \mathrm{do}(\mathrm{A}=1)] - \mathrm{E}[\mathrm{Y} \mid \mathrm{do}(\mathrm{A}=0)].$$

However, the presence of the do operator makes this formula inapplicable to observational data, since we don't have access to the counterfactual world. Instead, we can derive a statistical estimand for ATE with quantities that can be estimated from observational data. In this case, we use the Adjustment with No Unobserved Confounders Theorem, where we arrive at

$$\mathrm{ATE} \ = \mathrm{E}[\mathrm{E}[\mathrm{Y} \mid \mathrm{A}=1, \mathrm{X}]] - \mathrm{E}[\mathrm{E}[\mathrm{Y} \mid \mathrm{A}=0, \mathrm{X}]].$$

Nevertheless, the Overlap Assumption for this estimand requires a sufficient amount of random variation in treatment assignments. Namely, we need

$$0 < P(A=1 \mid X=x) < 1$$

which essentially means that there must be some heterogeneity in whether the sample gets treated or untreated based on their covariates. Namely, the covariates shouldn't identify the treatment. In practice, this assumption is even less forgiving, as values close to 0 or 1 can still drastically inflate the estimates. This happens to be the case in our data, where we have many observations (sometimes as many as 2000) that have a very small probability (< 0.05) of receiving treatment. Luckily, we can adjust our estimand, and instead use the Average Treatment Effect on the Treated (ATT), defined as

$$\mathrm{ATT} = \mathrm{E}_{\mathrm{X} \mid \mathrm{A}=1}[\mathrm{E}[\mathrm{Y} \mid \mathrm{X}, \mathrm{do}(\mathrm{A}=1)] - \mathrm{E}[\mathrm{Y} \mid \mathrm{X}, \mathrm{do}(\mathrm{A}=0)]].$$

The ATT is useful when we care more about the effect of the treatment on those who were actually treated than on the entire population, along with the fact that it helps to relax our Overlap Assumption. It is useful for our dataset since we care more about the effect of oil reserves and oil discovery on political violence among those countries that do produce oil. Using the ATT relaxes our Overlap Assumption to

$$P(A=1 \mid X=x) < 1,$$

which our data satisfies. If the No Unobserved Confounders Assumption also holds, then we can causally identify ATT as

$$\hat{\tau} = E_{X \mid A=1}[E[Y \mid A=1, X] - E[Y \mid A=0, X]]$$

With the formula above, we use a double machine learning estimator, known as the Augmented Inverse Probability of Treatment Weighted (AIPTW) estimator, defined for ATT as

$$\tau^{ATT-AIPTW} = \frac{1}{n} \sum_i \frac{A_i}{P(A=1)} \left(Y - \widehat{Q}\left(0, X_i\right) \right) - \frac{\left(1-A_i\right) g(X)}{P(A=1)(1-g(X))} \left(Y - \widehat{Q}\left(0, X_i\right) \right).$$

where $\widehat{Q}$ is the estimate for the conditional expected outcome, $\widehat{g}$ is the estimate for the propensity score function, X is the set of covariates, Y is the outcome, and A is the treatment.

[3]in the main body of this paper, we use the significance level $\alpha = 0.05$.

Usually, reusing our data for our Q, g, and AIPTW models would lead to overfitting. To remedy this, we elected to use k-fold cross validation to prevent these issues. By having an estimator that combines both g and Q, we can make an estimate that is consistent as long as either one of $\widehat{Q}$ or $\widehat{g}$ is consistent, also known as a doubly robust estimator. On top of this, the AIPTW estimator is non-parametrically efficient, meaning it has the smallest possible variance out of any estimator that avoids making parametric assumptions. The estimator for the variance of $\widehat{\tau}$ is

$$\widehat{V}[\widehat{\tau}] = \frac{1}{n} \sum_i \phi \left(X_i; \widehat{Q}, \widehat{g}, \widehat{\tau} \right)^2 ,$$

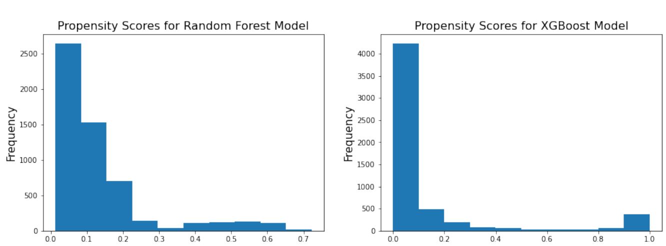
with ϕ being our influence curve function. With the estimates for $\hat{\tau}$ and its variance, we can conduct hypothesis tests to see the statistical significance of $\hat{\tau}$.

B. *Model Assumptions*

In section 2, we have talked about exogeneity assumption for treatments and i.i.d. observation assumption. In this section, we will discuss the assumptions that must hold true for the model results to be interpreted as causal. We have two assumptions that need to hold true if our estimate for $\widehat{\tau}$ to be equal to the ATT, which are the Overlap Assumption and the No Unobserved Confounders Assumption.

1) Overlap Assumption: As discussed in the previous section, our Overlap Assumption is relaxed by using ATT instead of ATE. In our Random Forest models, we find no observations with $\hat{g} > .95$, so we can safely say that the Overlap Assumption is not violated. However, in the XGBoost models, we do have some observations with $\hat{g} > .95$, even after adjusting some of the regularization parameters (max depth, number of iterations) on our model, as shown in Figure 1. Instead of dealing with inflated estimators, we solve this issue by dropping the observations where the Overlap Assumption is violated. While this does change the interpretation of our estimand, we can compare our results to that of the Random Forest models to look for any large variations in our estimate for $\widehat{\tau}$.

FIGURE 1: Histograms for propensity scores. As shown above, the XGBoost model violates the Overlap Assumption unless extreme observations are dropped, whereas the Random Forest model has all propensity scores $< .95$.



2) No Unobserved Confounders Assumption: The No Unobserved Confounders Assumption states that any covariate that affects both the treatment and the outcome needs to be included within the model. Table II lists our covariates.

TABLE II: Covariate list. For some covariates, we take the yearly percent difference to reduce intertemporal dependence between same-country observations, as discussed in section II.B.

Covariates in Percentage Change	**Untransformed Covariates**
1. Crude Oil Price	2. OPEC Membership Dummy
2. Wildcats Drilled	2. Log Share of Mountainous Land
3. Natural Disasters	3. Ethnic Fractionalization
4. GDP per Capita	4. Language Fractionalization
5. Population	5. Religious Fractionalization
6. Population Density	6. Legal British Origin Dummy
7. Democracy Index	7. Decade Fixed Effect

While we cannot exhaust all the confounders that may affect political violence, we try to encapsulate countries' natural, historical, social, cultural, and economic conditions in our covariates as much as the data allow. Because the untransformed covariates (2)-(6) are very country-specific and mostly time-invariant, we did not include an additional country-fixed effect at the cost of further violating the Overlap Assumption. To corroborate this, adding country code to our models had an insignificant impact while creating more observations with an extreme propensity score[4]. We also had a decade-fixed effect (untransformed covariates, 7) to capture the historical trend of violence on a global level.

Some of our covariates are common causes of oil abundance and political violence. For example, low GDP growth per capita (percent change covariates, 4) may render a country more vulnerable to wars while triggering it to further exploit oil resources. Some variables primarily affect the outcome, instead of oil abundance, and we include them to restrict variation in outcome for stabler ATT estimation. For example, we expect religious fractionalization (untransformed covariates, 5) to ignite civil conflicts.

While we cannot assert that there are no unobserved confounders, we believe our results provide an upper bound on the effect of oil abundance on political violence. Based on existing literature, we find that most unobserved confounders affect oil abundance and political violence in the same direction. For instance, unfavorable institutions, like weak property rights, may de-incentivize extraction (Bohn and Deacon 2000) and are prone to civil conflicts. Thus, our estimates, if biased, are most likely to be biased upwards and thus serve as an upper bound for the true causal effect.

Furthermore, as discussed in section II.A, we may view oil discovery as conditionally random given our set of covariates. Therefore, at least one of our treatments is very likely, if not truly, to satisfy the No Unobserved Confounders Assumption.

IV. RESULTS

A. *Oil Reserves Effects*

When running our models, we use both Random Forest Classifiers and XGBoost Classifiers for our g and Q models in order to compare the results. We choose to use multiple models because we are testing many hypotheses, so we can

[4]We performed our analysis with country fixed effect as an additional covariate. Because it does not change our results substantially, we leave it out of this paper. Interested readers may find our analysis code with country fixed effect on GitHub (see section VII).

cross reference significant values for our ATT against the other model to see if we obtain similar results. Using the AIPTW estimator outlined in Section III.A, we are able to find estimates for our ATT for all of our treatments and outcomes. Shown below are the impact of oil reserves (oilres) and value of oil reserves (valoilres) as treatments with war onset (war onsets), number of military coup attempts (coup), percent leadership transitions in following 20 years (leadership), and military spending as a fraction of GDP (Δ defense burden).

As can be seen in Table III, none of the p values are significant for both the Random Forest model and the XGBoost model. For example, the impact of oil reserves on the number of coups appeared to be significant from our XGBoost model ($p = .0481$), but then it showed to be insignificant in our XGBoost model. Similarly, we find that the effects of the value of oil reserves on coup attempts ($p = .027$) and irregular leadership shifts ($p = .014$) diminish in significance when we shift from Random Forest to the XGBoost model. Given that we have done 16 ATT estimations on sub-samples in total, it should not be surprising that we have a few statistically significant results. Thus, we argue that we cannot conclude that oil abundance affects political violence in general.

This, along with the established conclusion from the original paper, points us toward the qualitative result that oil abundance has no impact on political violence, given the information that we have.

TABLE III: ATT estimates for oil reserves and value of oil reserves. In the model column, rf refers to the Random Forest model, and xgb refers to the XGBoost model. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

model	treatment	outcome	$\hat{\tau}$	τ std.	p value
rf	oilres	coup	-0.0234	0.02159	0.139
xgb	oilres	coup	-0.0481**	0.02750	0.040
rf	valoilres	coup	-0.0246**	0.01286	0.027
xgb	valoilres	coup	0.13969*	0.10730	0.096
rf	oilres	Δ defense burden	3.32819	2.92982	0.128
xgb	oilres	Δ defense burden	3.74020	2.97670	0.104
rf	valoilres	Δ defense burden	0.52978*	0.41046	0.098
xgb	valoilres	Δ defense burden	0.09476	0.12695	0.227
rf	oilres	war onsets	-0.00181	0.00682	0.395
xgb	oilres	war onsets	-0.00171	0.00794	0.414
rf	valoilres	war onsets	-0.00009	0.00484	0.492
xgb	valoilres	war onsets	-0.00881	0.01666	0.298
rf	oilres	leadership	-0.00342	0.00687	0.309
xgb	oilres	leadership	-0.00350	0.01073	0.372
rf	valoilres	leadership	-0.01274**	0.00584	0.014
xgb	valoilres	leadership	-0.00580	0.01250	0.321

B. *Oil Discovery Effects*

As aforementioned, oil discovery also serves as an intuitive way to measure a country's natural oil wealth, which relates back to our central question of whether or not natural resource abundance leads to political instability. Since estimated effects of oil discovery on war onsets, military coups, irregular leadership transitions, and defense burden are similar to those of oil reserves, Table IV only displays the effects on additional outcomes: 2-year change in war status (Δ war status (2 yr)) and a 3-6 year change in war status (Δ war status (3-6 yr)).

TABLE IV: ATT estimates for oil discovery. In the model column, rf refers to the Random Forest model, and xgb refers to the XGBoost model. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

model	treatment	outcome	$\hat{\tau}$	τ std.	p value
rf	discovery	war onsets	0.00478	0.00363	0.093
xgb	discovery	war onsets	0.00474	0.00641	0.229
rf	discovery	coup	-0.03628***	0.00997	0.000
xgb	discovery	coup	-0.01015	0.02397	0.335
rf	discovery	leadership	-0.01579***	0.00498	0.000
xgb	discovery	leadership	-0.01756**	0.01049	0.047
rf	discovery	Δ defense burden	0.00369	0.14001	0.489
xgb	discovery	Δ defense burden	0.03495	0.10932	0.374
rf	discovery	Δ war status (2 yr)	0.00256	0.01281	0.420
xgb	discovery	Δ war status (2 yr)	-0.01817	0.02404	0.224
rf	discovery	Δ war status (3-6 yr)	-0.00717	0.01556	0.314
xgb	discovery	Δ war status (3-6 yr)	0.00097	0.03708	0.489

As can be seen in the results above, we arrive at a very similar conclusion to the previous section. We find no significant increase or decrease in future violence based on new discoveries of oil. One exception is irregular leadership transition. The effects of oil discovery on this outcome is consistently significant and negative in both models. However, we will see that the significance largely vanishes when we split the sample into democratic and non-democratic sub-samples. In general, we hold the similar qualitative reasoning, which is that new oil discovery does not lead to political instability in the following years.

C. *Heterogeneous Effects: Democracies vs. Non-Democracies*

While our results have been insignificant thus far, this could be due to heterogeneous effects in different types of countries. Particularly, we believe there could be a greater effect of oil abundance on political instability in non-democratic countries since the wealth that comes with oil abundance could lead to more hostile political actions if a democratic system is not in place. We define democratic countries as countries with a democracy index strictly over .005 at the year of observation.

Our treatment distribution is balanced in the democratic and non-democratic sub-samples. Similar to the previous case, both subgroups violate the Overlap Assumption under the XGBoost model, so we discarded the extreme observations. Under the Random Forest model, the Overlap Assumptions hold for both subgroups.

From Table V, in the democratic sub-sample, under the XGBoost model, the value of oil reserves has a significantly negative impact on defense burden ($p = .022$). Under the Random Forest model, oil reserves have a significantly negative impact on war onsets ($p = .037$). From Table VI, in the non-democratic sub-sample, all ATT estimates are not significant, while oil reserves and the value of oil reserves consistently have positive impacts on defense burden.

We have two interesting findings. First, as in the original paper, we find that oil abundance positively affects the defense burden only among non-democratic countries. However, the estimates are insignificant (though they are consistent and have p-values close to 0.1). Second, as in our previous analysis, our model almost fails to spot any significant effects of oil abundance in either sub-sample. The only exceptions are the effect of the value of oil reserves on war onsets and defense burden among democratic countries. However, we argue that we cannot conclude that oil abundance affects political violence in general.

We expect all our outcome variables to be closely related to measures of political violence. While we admit nuanced differences between these incidences in specific political discourses, it is hard to conjecture a global explanation for why oil abundance only diminishes defense burden and war onsets among democratic countries, but does not alter irregular leadership shifts and coup attempts. Furthermore, even if oil abundance indeed reduces defense burden and war incidences, its aggregated effect on political violence should not be very large.

The effects of oil discoveries among democratic and non-democratic countries are very similar. We report them in Table VII and VIII. Notice that the effects on leadership transition spotted in the whole sample (from Table V) becomes less consistent and largely dissolves except for in the Random Forest model in non-democratic countries.

TABLE V: ATT estimates for oil reserves and value of oil reserves among democratic countries. In the model column, rf refers to the Random Forest model, and xgb refers to the XGBoost model. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

model	treatment	outcome	$\hat{\tau}$	τ std.	p value
rf	oilres	coup	0.012426	0.013874	0.185
xgb	oilres	coup	-0.008449	0.029862	0.388
rf	valoilres	coup	0.011234	0.010210	0.135
xgb	valoilres	coup	0.055809*	0.043246	0.098
rf	oilres	Δ defense burden	0.894443	0.886181	0.156
xgb	oilres	Δ defense burden	7.164805	7.195176	0.159
rf	valoilres	Δ defense burden	-0.010718	0.022415	0.316
xgb	valoilres	Δ defense burden	-0.174507**	0.086964	0.022
rf	oilres	war onsets	-0.005476**	0.003083	0.037
xgb	oilres	war onsets	-0.023302*	0.015872	0.071
rf	valoilres	war onsets	-0.000528	0.003076	0.431
xgb	valoilres	war onsets	-0.005328	0.004251	0.105
rf	oilres	leadership	0.005464*	0.004043	0.088
xgb	oilres	leadership	-0.007567	0.009049	0.201
rf	valoilres	leadership	-0.001331	0.004297	0.378
xgb	valoilres	leadership	-0.008859	0.007592	0.121

TABLE VI: ATT estimates for oil reserves and value of oil reserves among non-democratic countries. In the model column, rf refers to the Random Forest model, and xgb refers to the XGBoost model. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

model	treatment	outcome	$\hat{\tau}$	τ std.	p value
rf	oilres	coup	-0.04051	0.04190	0.166
xgb	oilres	coup	-0.05755	0.05234	0.135
rf	valoilres	coup	-0.02650	0.02341	0.128
xgb	valoilres	coup	0.015975	0.03145	0.305
rf	oilres	Δ defense burden	5.79781	5.41978	0.142
xgb	oilres	Δ defense burden	4.71506	4.79327	0.162
rf	valoilres	Δ defense burden	0.99018	0.94079	0.146
xgb	valoilres	Δ defense burden	1.23434	1.03988	0.118
rf	oilres	war onsets	0.003498	0.01514	0.408
xgb	oilres	war onsets	-0.00026	0.01818	0.494
rf	valoilres	war onsets	0.002855	0.00886	0.373
xgb	valoilres	war onsets	-0.02484	0.03930	0.263
rf	oilres	leadership	0.005301	0.01263	0.337
xgb	oilres	leadership	-0.01212	0.01755	0.244
rf	valoilres	leadership	-0.01052	0.00986	0.143
xgb	valoilres	leadership	-0.01456	0.01691	0.194

TABLE VII: ATT estimates for oil discovery in democratic countries. In the model column, rf refers to the Random Forest model, and xgb refers to the XGBoost model. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

model	outcome	tau_hat	tau_std	p_value	
rf	discovery	Δ war status (2 yr)	0.00615	0.01491	0.339
xgb	discovery	Δ war status (2 yr)	0.00313	0.01979	0.437
rf	discovery	Δ war status (3-6 yr)	-0.01928	0.02086	0.177
xgb	discovery	Δ war status (3-6 yr)	-0.05428	0.04856	0.131
rf	discovery	war onsets	0.00441	0.00361	0.111
xgb	discovery	war onsets	-0.00026	0.00586	0.481
rf	discovery	coup	-0.00371	0.00778	0.316
xgb	discovery	coup	0.01494	0.02893	0.302
rf	discovery	leadership	0.00461	0.00381	0.113
xgb	discovery	leadership	-0.00490	0.00794	0.268
rf	discovery	Δ defense burden	-0.04507	0.02407	0.030
xgb	discovery	Δ defense burden	-0.04663	0.03474	0.089

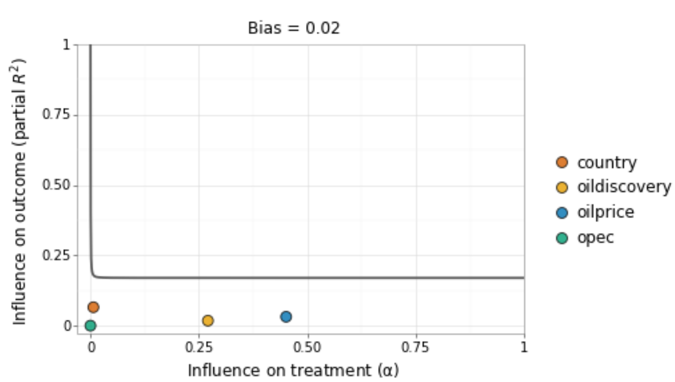
TABLE VIII: ATT estimates for oil discovery in non-democratic countries. In the model column, rf refers to the Random Forest model, and xgb refers to the XGBoost model. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

model	treatment	outcome	tau_hat	tau_std	p_value
rf	discovery	Δ war status (2 yr)	0.00213	0.02031	0.458
xgb	discovery	Δ war status (2 yr)	-0.06370*	0.04458	0.076
rf	discovery	Δ war status (3-6 yr)	0.01774	0.02284	0.218
rf	discovery	Δ war status (3-6 yr)	0.02743	0.03860	0.238
rf	discovery	war onsets	0.01026**	0.00592	0.041
xgb	discovery	war onsets	0.01675***	0.00634	0.004
rf	discovery	coup	-0.05179***	0.01831	0.002
xgb	discovery	coup	-0.02331	0.04299	0.293
rf	discovery	leadership	-0.01924***	0.00805	0.008
xgb	discovery	leadership	-0.04808	0.02167	0.013
rf	discovery	Δ defense burden	-0.00383	0.37019	0.495
xgb	discovery	Δ defense burden	0.39297	0.57878	0.248

V. DISCUSSION

Figure 2 shows the Austen plot from a sample sensitivity analysis showing how strong an unobserved confounder would need to be to induce a bias of 0.02, the ATT estimate (though not significant), in studying the effect of the value of oil reserves on the war status.

FIGURE 2: Austen plot from a sample sensitivity analysis showing how strong an unobserved confounder would need to be to induce a bias of 0.02 in studying the effect of the value of oil reserves on the war status



From the Austen plot, we can see that it would take a large effect in order to introduce a bias of .02. This instills further confidence in our results since the odds that there exists an unobserved confounder with an effect greater than any observed variable seem to be low from our qualitative assessment.

Furthermore, in both models of the effect of oil abundance and oil discovery on political violence, we show that excluding the United States (where, unlike most other countries, ownership of mineral resources is private) and the USSR/Russia (another major power and top oil producer) in the sample does not change the conclusions. Therefore, the specific economic mechanisms for USA and USS/Russia as potential outliers do not largely alter the effects of oil abundance.

Indeed, as shown in Table IX we have a few significant estimates: Under the Random Forest model, the value of oil reserves negatively affects coup attempts ($p = .022$) and irregular leadership transition ($p = .027$). Under the XGBoost model, the value of oil reserves negatively affects irregular leadership transition ($p = .013$). We have no time-lagged significant effects with oil discoveries as the treatment (see Table X). To conclude, we should still refrain from rejecting the hypothesis that oil abundance affects substantial political violence in general, as the majority of our estimates do not display statistical significance. Among all our the treatment-outcome pairs, the most likely casual relation comes from oil discovery and leadership, as shown in Table V. This effect is less obvious when we split the sample into non-democratic and democratic sub-samples.

Even if we believe that there exists a significant causal effect of oil abundance on political violence, the effect should be negative, instead of positive as many scholars claimed in the introduction. Moreover, we have mentioned in section III.B that our results are most likely biased upwards, not downwards. So, our findings are counter-evidence to the commonly believed positive causal relationship between oil abundance and political violence.

TABLE IX: ATT estimates for oil reserves and value of oil reserves among countries except for USA and USSR/Russia. In the model column, rf refers to the Random Forest model, and xgb refers to the XGBoost model. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

model	treatment	outcome	$\hat{\tau}$	τ std.	p value
rf	oilres	coup	-0.02356	0.02149	0.136
xgb	oilres	coup	-0.03376	0.0283	0.117
rf	valoilres	coup	-0.02646**	0.01318	0.022
xgb	valoilres	coup	0.021666	0.03065	0.239
rf	oilres	Δ defense burden	3.22234	2.94886	0.137
xgb	oilres	Δ defense burden	3.05623	2.82551	0.139
rf	valoilres	Δ defense burden	0.562839*	0.42549	0.093
xgb	valoilres	Δ defense burden	0.580299	0.48289	0.114
rf	oilres	war onsets	-0.00356	0.00696	0.304
xgb	oilres	war onsets	-0.00338	0.00774	0.331
rf	valoilres	war onsets	-0.00172	0.00508	0.366
xgb	valoilres	war onsets	-0.08754	0.09628	0.181
rf	oilres	leadership	-0.00127	0.00603	0.416
xgb	oilres	leadership	-0.00108	0.00885	0.451
rf	valoilres	leadership	-0.01119**	0.00584	0.027
xgb	valoilres	leadership	-0.02168**	0.00983	0.013

TABLE X: ATT estimates for oil discoveries among countries except for USA and USSR/Russia. In the model column, rf refers to the Random Forest model, and xgb refers to the XGBoost model. Again, we left out the effects of oil discoveries on other outcomes, because they are similar to results in Table IX when we used oil abundance as the treatment. $p < .1 =$*, $p < .5 =$**, $p < .01 =$***.

model	treatment	outcome	$\hat{\tau}$	τ std.	p value
rf	discovery	Δ war status (2 yr)	0.00676	0.01220	0.289
xgb	discovery	Δ war status (2 yr)	-0.01764	0.0240	0.247
rf	discovery	Δ war status (3-6 yr)	0.01318	0.01529	0.194
xgb	discovery	Δ war status (3-6 yr)	-0.01290	0.04681	0.391

VI. CONCLUSION

In this project, we used non-parametric methods to replicate a study that used only parametric methods to determine if the results of the previous paper held using more advanced causal techniques. We found that the results found by the original authors were very similar when using these non-parametric methods, with the level of oil abundance in a country having no significant impact on their level of conflict. In addition, we found that new discoveries of oil also have no significant impact on future events of political violence. Admittedly, there are a few significant ATT estimates depending on the model specification and choice of variables. However, the significance of these effects is usually not consistent across models and outcomes. Hence, we suspect that this most likely comes from the sheer number of estimations we have done. Even if we admit that oil abundance affects political violence, the effects should be very limited in magnitude and negative.

Additionally, our analysis stood in contrast to the original paper in regard to the heterogeneous effects in democratic vs. non-democratic countries. While the original paper found that oil abundance had a significant impact on defense spending, our results found no significance even after splitting on democracy. An interesting next step in this project could be to explore a similar dataset from more recent years to see if the insignificant results still hold on data from the 21st century.

VII. CODE AVAILABILITY AND ACKNOWLEDGMENTS

Code, data, and analysis for this project are all contained in the GitHub repo located at https://github.com/wanranzhao/STAT27420-final-code. We would also like to thank Prof. Veitch for his insightful and helpful advice on choosing covariates and making model assumptions. Along with this, we would like to thank him for a great, informative quarter.

REFERENCES

Bohn, Henning and Robert T. Deacon (June 2000). "Ownership Risk, Investment, and the Use of Natural Resources". en. In: *American Economic Review* 90.3, pp. 526–549. ISSN: 0002-8282. DOI: 10.1257/aer.90.3.526. URL: https://www.aeaweb.org/articles?id=10.1257/aer.90.3.526 (visited on 12/10/2022).

Cotet, Anca M. and Kevin K. Tsui (2013). “Oil and Conflict: What Does the Cross Country Evidence Really Show?” In: *American Economic Journal: Macroeconomics* 5.1. Publisher: American Economic Association, pp. 49–80. ISSN: 1945-7707. URL: https://www.jstor.org/stable/43189930 (visited on 12/10/2022).

Robinson, James, Ragnar Torvik, and Thierry Verdier (2006). “Political foundations of the resource curse”. In: *Journal of Development Economics* 79, pp. 447–468.