Natural Resources and Democracy: Not Blessing*

Xiaoyu Li[+]

Abstract

In Haber and Menaldo's 2011 article, they analyze the relationship between natural resources (measured by total oil income per capita) and regime type (measured by polity score), reporting support for a resource blessing. In this article, I demonstrate that the finding of a positive relationship between natural resources and democracy is to an extent a product of the unrealistic assumption that the relationship between total oil income per capita and polity score is linear. I first successfully replicate Haber and Menaldo's original analysis and critique their model specification on inclusion of unnecessary variables and implausible functional form. I then respecify the model under more plausible assumptions by performing logarithmic transformation on the highly skewed independent variable. Results show that natural resources actually has nonexistent effects on democracy.

* This is a course project for GOV2001 - Advanced Quantitative Research Methodology - instructed by Gary King. My thanks to Stephen Haber and Victor Menaldo for sharing the dataset and code files, and Gary King, Stephen Pettigrew, Sole Prillaman, Peter Kim, and Sean Paul Ashley for constructive comments, and Omer Orsun for helpful discussion and programming assistance.

[+] Department of Social and Behavioral Sciences, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02215 (xil288@mail.harvard.edu).

Introduction

In a recent article in American Political Science Review, Haber and Menaldo (2011) analyze the relationship between natural resources and regime type. Their goal is to assess the resource curse — the claim that natural resources fuel authoritarianism. This idea could be traced back to Mahdavy (1970), who notes that the external rents from petroleum revenues in Middle Eastern countries are directly captured by the governments and rendering them unaccountable to citizens. As Luciani (1987)states, no representation without taxation.

The authors worry that previous research might suffer omitted variable bias caused by unobserved country-specific and time-invariant heterogeneity. To address the concern, they develop a unique historical dataset (covering 168 countries from 1800 to 2006) to test if there is a long-run relationship between resource reliance and democracy. They do so by specifying and estimating a series of error correction mechanism (ECM) regressions with country and time fixed effects.

They find that there is no association between resource reliance and authoritarianism. Instead, in many specifications, they find evidence for a resource blessing: increases in natural resources are associated with increases in democracy. This is a noteworthy claim since a plethora of previous research has demonstrated the opposite – a resource curse. Strong evidence of a resource blessing would mean that the previous conclusions are wrong.

In what follows, I do not evaluate the reliability of Haber and Menaldo's data and assume their choice of models are generally justified. Rather, I make three specific

points. First, as it turns out, Haber and Menaldo's results could be replicated relatively easily with a few acceptable exceptions. Second, I critique the specification of their original model on inclusion of unnecessary variables and inappropriate functional form. Finally, I show that the resource blessing the authors find is a product of their assumption that the relationship between resource reliance and democracy is linear. A more feasible specification of the model renders the effects nonexistent.

Replication of Haber and Menaldo's Analysis

Original Model Specification

In this article I focus on Haber and Menaldo's most favored model (Table 5, Column 1) (2011, page 19), which is an error correction model for the impact of total oil income per capita on polity score with country and time fixed effects for the full panel. They specify the model as follows:

$$\Delta Y_t = \Delta X_{1t}\beta_1 + X_{2(t-1)}\beta_2 + \delta\left(Y_{t-1} - X_{1(t-1)}\gamma\right) + u_t \ , \qquad (1)^1$$

where Greek letters are parameters to be estimated; Y_t, the Polity Score normalized to range from 0 to 100 at time t; ΔX_{1t}, a vector of differenced predictors at time t, including Total Oil Income Per Capita (in thousands of 2007 US dollars), Log(per Capita Income in thousands of 2007 US dollars), Regional Democratic Diffusion (the percentage of democracies in a country's geographic-cultural region), and Global Democratic Diffusion (the percentage of democracies in the world); $X_{1(t-1)}$; a lag of

[1] The original specification of the model by Haber and Menaldo's is: $\Delta Y_t = \Delta Y_{t-1}\rho + \Delta X_t\beta_1 + \Delta X_{t-1}\beta_2 + \cdots + \Delta X_{t-k}\beta_1 + \delta(Y_{t-1} - X_{t-1}\gamma) + u_t$. Their final model does not include everything in this generalized form. I thus simplify it.

all variables in the vector X_{1t}; $X_{2(t-1)}$; a lag of a control dummy indicating an ongoing civil war.

The ECM model estimates both the long-run total impact on polity score made by a permanent change in the level of predictors and the short-run effects. Take total oil income per capita for example, an immediate change in polity score induced by a change in total oil income per capita that take a year's time to elapse is captured by the associated coefficient in β_1 on the differenced predictor; and change in total oil income per capita produces a change in polity score that disrupts the long-term equilibrium relationship between the two. Polity score will therefore gradually return to the path traced by total oil income per capita, registering a total change equal to γ. The δ term is the error correction rate: a δ proportion of this discrepancy/error is corrected by a movement in the polity score each subsequent period."

Successful Replication

Table 1 presents the results. The first two columns are from Haber and Menaldo's article and the last two columns are replicated results. As could be seen, the original estimation could be replicated successfully, with a few acceptable exceptions. Note that the replicated results do not have coefficients for the differenced global democratic diffusion and the lag of the same variable. Reasons will be outlined in the next section.

Table 1: Error Correction Models for the Impact of Total Oil Income on Polity Score (Country Fixed Effects) [Column 1 in Table 5][a]

	Original		Replicated	
Variable	**Estimate**	**Robust t-Stat[b]**	**Estimate**	**Robust t-Stat**
Polity in Levels t − 1 (Error Correction Term)	−0.087	[11.55]***	-0.087	[-11.35]***
Total Oil Income t − 1	0.055	[2.90]***	0.055	[2.85]***
Total Oil Income Long-run Multiplier (LRM)[c]	0.634	[3.06]***	0.634	[3.06]***
ΔTotal Oil Income	−0.02	[0.97]	-0.02	[-0.95]
Log(Per Capita Income) t − 1	−0.279	[0.88]	-0.28	[-0.86]
Civil War t−1	0.065	[0.15]	0.065	[0.14]
Regional Democratic Diffusion t − 1	0.025	[3.49]***	0.025	[3.43]***
Global Democratic Diffusion t − 1	0.038	[1.54]		
ΔLog(Per Capita Income)	1.289	[0.74]	1.288	[0.73]
ΔRegional Democratic Diffusion	0.375	[5.37]***	0.375	[5.27]***
ΔGlobal Democratic Diffusion	−0.244	[2.34]**		
Country Fixed Effects	Yes		Yes	
Year Fixed Effects	Yes		Yes	
Observations	10,195		10,195	
Number of Groups	163		163	
R-squared	0.1		0.1	

Notes:
a. Separate country and year intercepts estimated but omitted from table.
b. Robust-statistics in brackets (Driscoll–Kraay standard errors estimated with Newey–West adjustment with one lag length).
c. LRM esimated by −1(b(Total Oil Incomet −1)/b(Polityt −1)); standard errors estimated using the delta method.
* Significant at 10%; ** 5%; *** 1%.

Critique of Original Model Specification

Inappropriate Inclusion of the Global Democratic Diffusion Variables

As Table 1 shows, in replication I leave out the differenced and the lagged global democratic diffusion variables, which has not substantively changed the results. Haber and Menaldo include these two variables since "one might also argue that increased

democratization in resource-reliant countries is influenced by world or regional trends" (2011, page 12). This is an excellent motivation. However, global democratic diffusion is calculated as the percentage of democracies in the world, which will be the same value for every country in the same year. Hence, the inclusion of these two variables would be inappropriate. In replication, I leave the differenced and the lagged global diffusion variables out because a model with time fixed effects already accounts for variables that are constant across country within the same year.

Divergence of Robust and Classical Standard Errors of Total Oil Income Per Capita

Haber and Menaldo's analysis finds a positive and statistically significant long-term impact of total oil income per capita on polity score. Though the magnitude is small. The Long-Run Multiplier (LRM) for total oil income per capita is 0.634, indicating that over the long-run, for every $1,000 increase in total oil income per capita, polity score will increase by .635 points (on a 0 to 100 scale), holding other covariates constant. Note that the interpretation in the article does not include any measure of uncertainty, though the result table gives a robust s-statistic and corresponding significant levels. A more informed interpretation could have been possible if measures of uncertainty are explicitly included.

That said, from estimation I see that the estimated robust Driscoll-Kraay standard error for total oil income per capita is harshly different from the classical standard error, with the former being less than a half of the latter (.211 vs. .525, respectively). The LRM is two robust standard errors away from zero but not two classical standard errors

away from zero.

Such a big divergence of the classical and robust standard errors could suggest a significant amount of heteroskedasticity and autocorrelation in the data. It could also suggest model misspecification. “If clustering and time-dependence is well modeled so that the residuals have no clustering or time series patterns, then the classical and cluster-or autocorrelation-robust standard variance matrices should be approximately the same”(King & Roberts, 2013). Since ECM models with country and time fixed effects are supposed to take care of autocorrelation at least to some extent, the divergence of classical and robust standard errors gives a clue about possible model misspecification.

Model Respecification

In this section I respecify Haber and Menaldo’s original model with the goal of better inference.

First I exclude the differenced and lagged global democratic diffusion variables, as in replication. More importantly, I employ logarithmic transformation on the key explanatory variable total oil income per capita. Most researchers in the field have been transforming income variables(e.g., Friedman, 2006; Oskarsson & Ottosen, 2010; Ross, 2001). In addition to this long-standing tradition, the logarithmic transformation is desirable for two reasons. First, taking the logarithm of distance will reduce the skew in the distribution and bring the outlying data points from the right tail towards the rest of the data, which would help fitting the variable into a model. Even if we only look at

observations with non-zero total oil income per capita, there is still extreme skewness in the distribution, as can be seen in the long right tail in the left panel of Figure 1[2]. On the other hand, the log of total oil income per capita, as shown in the right panel of Figure 1, is much more symmetric and approximately normally distributed.

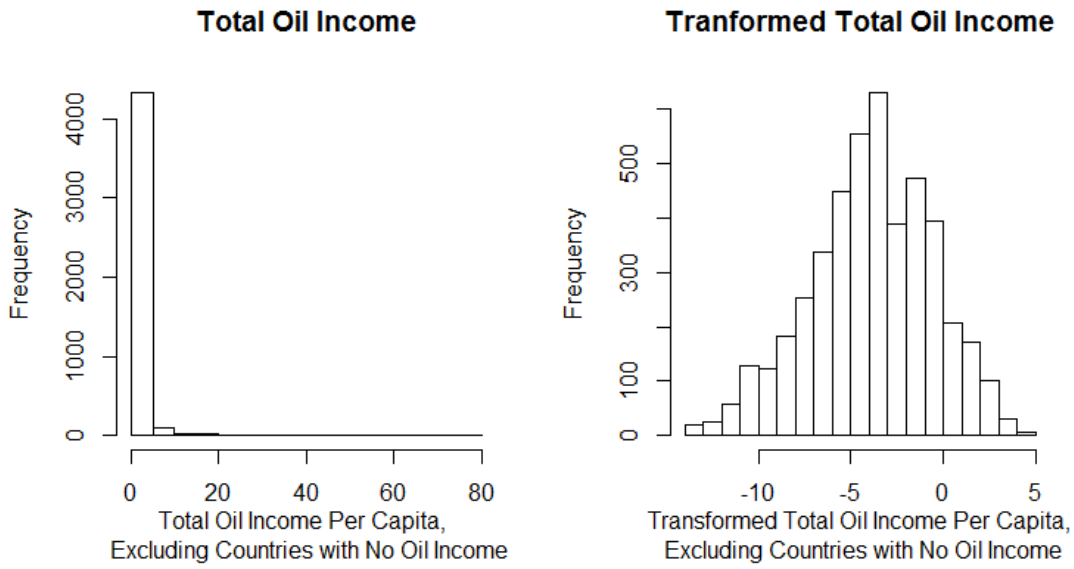

Figure 1 Distribution of the Key Independent Variable before (left) and after (right) the Logarithm Transformation.

The second reason for this transformation is that the logged form makes substantially better sense because the estimated coefficients from regressions are predicting changes in polity score from proportionate changes in total oil income per capita. Haber and Menaldo's decision simply to regress polity score against total oil income per capita as it comes, on a linear scale which assumes that the difference between one thousand dollars of total oil income per capita is the same between $0 and $1,000 as it is between $77,000 and $78,000, is apparently very implausible. To just

[2] Only 4538 out of 10195 country-year observations in the analysis sample has non-zero total oil income per capita.

give one example, an observation with a $77,000 total oil income per capita is usually much wealthier than an observation with $0 total oil income per capita. In such cases, a change of one thousand dollars is likely to have different impacts on regime type for these two observations. A plausible assumption is that this one thousand dollars would have a bigger impact on a poorer country than on a wealthier country. The log scale, in this sense, is far more sensitive to differences when values are small than when they are large.

Since the imposed linearity assumption in the original analysis is likely to bias coefficients estimated, I replace the differenced and lagged total oil income per capita variables with their log transformed counterparts in equation (1) and re-estimate the model.

Results: No blessing

Convergence of Robust and Classical Standard Errors of Log (Total Oil Income Per Capita)

Results of the reanalysis which changes the linearity assumption between polity score and total oil income per capita are presented in Table 2.

Let us first look at the (-.210) coefficient on the explanatory variable of interest: Log (Total Oil Income Per Capita) LRM. The classical standard error is now .404 and the robust standard error .308. Clearly these two standard errors have converged.

This could be indicative of a better model fit as the change of functional form bring the classical and robust standard errors more in line and thus reduce the possibility of

misspecification. Of course, a smaller difference between the two standard errors is not proof for a hypothesis of no misspecification.

Table 2: Error Correction Models for the Impact of Total Oil Income on Polity Score (Country Fixed Effects), Transformed[a]

Variable	Estimate	Robust Std.Error[b]	Robust t-Stat
Polity in Levels t − 1 (Error Correction Term)	−0.087	[0.008]	[-11.350]***
Log(Total Oil Income) t − 1	-0.018	[0.027]	[-0.682]
Log(Total Oil Income) t − 1 Long-run Multiplier (LRM)[c]	-0.210	[0.308]	[-0.68]
Δlog(Total Oil Income)	-0.037	[0.082]	[-0.454]
Log(Per Capita Income) t − 1	-0.178	[0.304]	[-0.585]
Civil War t−1	0.067	[0.457]	[0.147]
Regional Democratic Diffusion t − 1	0.025	[0.007]	[3.386]***
ΔLog(Per Capita Income)	1.220	[1.788]	[0.682]
ΔRegional Democratic Diffusion	0.375	[0.071]	[5.268]***
Country Fixed Effects	Yes		
Year Fixed Effects	Yes		
Observations	10,195		
Number of Groups	163		
R-squared	0.100		

Notes:
a. Separate country and year intercepts estimated but omitted from table.
b. Robust-statistics in brackets (Driscoll–Kraay standard errors estimated with Newey–West adjustment with one lag length).
c. LRM esimated by −1(b(Total Oil Incomet −1)/b(Polityt −1)); standard errors estimated using the delta method.
∗ Significant at 10%; ∗∗ 5%; ∗∗∗ 1%.

Elimination of Original Results

A close look at Table 2 shows that the differences between the new results and the original results are quite striking. The coefficient on the lagged log(total oil per capita) is no longer significant (two standard errors away from zero). Correspondingly, the LRM for log(total oil income per capita) is no longer positive or statistically significant,

indicating that there is no evidence that total oil income per capita has a long-term effect on polity score. If anything, the LRM shows a negative sign which is more in line with a resource curse. But the coefficient is not two (robust or classical) standard errors away from zero.

In terms of the short-term effect of total oil income per capita, the coefficient on the differenced predictor is also not statistically different from zero, which means that there is no evidence suggesting that total oil income has an immediate effect on polity score.

Under the new model specification, which is more plausible, it is clear that total oil income per capita does not have an impact on polity score, either short-term or long-term. The resource blessing finding suggested by Haber and Menaldo is thus eliminated. Natural resources is not a blessing on democracy.

Figure 2 presents two plots of residuals from the author's model (left panel) and the altered model (right panel) with total oil income per capita or log(total oil per capita) on the horizontal axis. As shown, taking the log makes the residuals more homoskadestic as compared to the original model. But note that both models do not do a good job in predicting polity score when the country does not have oil income at all. This could be a result that more than half of the observations have a total oil income per capita of $0.

Figure 2 Residual Plots of the Original Model (Left) and the Alternative Model (Right)

Concluding Remarks

In this article, I have first replicated results in Haber and Menaldo's article and then discussed the problems in the original model specification. I have respecified a more plausible model which changes the linearity assumption on the relationship between total oil income per capita and polity score. Results from the respecified model are consistent with prior research and are contradictory to Haber and Menaldo's conclusion, indicating that there is no evidence of a resource blessing.

In short, Haber and Menaldo's finding of a positive relationship between oil income and democracy could be a methodological artifact. At least, the relationship is scale dependent. Further research is needed to improve our understanding.

One thing that worth future efforts is that multiple imputation could be used to better deal with missing data. This analysis use interpolation, which is a little outdated and not as justified as multiple imputation. Another thing is that matching could be used

for more confident causal inference. Countries could be matched based on similar histories, geographic areas, cultural traditions, etc. Also the choice of model might need further consideration. In this article I try to adhere to Haber and Menaldo's approach as close as possible. Alternative approaches might prove to be beneficial, especially when there are so many observations with no oil income at all. Analysis employ such methods will sure advance our knowledge of the problem.

References

Friedman, T. L. (2006). The First Law of Petropolitics. *Foreign Policy*. 154, 28-36.

Haber, S., & Menaldo, V. (2011). Do Natural Resources Fuel Authoritarianism? A Reappraisal of the Resource Curse. *American Political Science Review*, *105*(01), 1–26.

King, G., & Roberts, M. (2013). How Robust Standard Errors Expose Methodological Problems They Do Not Fix.

Luciani, G. (1987). Allocation versus Production States: A Theoretical Framework. In *Hazem Beblawi and Giacomo Luciani*, ed. Beblawi, H. and Luciani, C. New York: Croom Helm.

Mahdavy, H. (1970). The Patterns and Problems of Ecnomic Development in Rentier States: The Case of Iran. In *Studies in the Economic History of the Middle East*, ed. M.A. Cook. London: Oxford University Press.

Oskarsson, S., & Ottosen, E. (2010). Does Oil Still Hinder Democracy? *Journal of Development Studies*, *46*(6), 1067–1083.

Ross, M. L. (2001). Does Oil Hinder Democracy? *World Politics*, *53*(2), 325.