FUEL SUBSIDIES LIMIT DEMOCRATIZATION: EVIDENCE FROM A GLOBAL SAMPLE, 1990-2014

Matthew D. Fails
Associate Professor
Department of Political Science
Oakland University
fails@oakland.edu

Forthcoming at *International Studies Quarterly*

Abstract

Oil wealth tends to impede democracy, but there is little consensus as to why or under what conditions. This paper helps answer such questions by evaluating the field's foundational theory of the rentier state, which claims that oil wealth finances generous societal benefits that reduce citizens' demands for representation and hinder the emergence of democratic regimes. I create a new measure of such benefits, focusing specifically on the size of domestic gasoline subsidies in dollars per-capita. I then use a global sample from 1990 through 2014 to demonstrate that greater spending on these subsidies significantly reduces the likelihood of a transition toward democracy. The impact on democratization is as consequential in practical terms as are large increases in the rate of economic growth. Moreover, including the measure of fuel subsidies helps account for the autocratic effect of oil income. I conclude by highlighting how this fuel subsidy data can shed light on a number of other political economy questions.

Pages 1-25 – main manuscript
Pages 26-48 – supplemental material

A large literature concludes that oil wealth can impede democracy (Lall 2017; Ross 2015; Ahmadov 2014; Wiens, Poast, and Clark 2014). Nevertheless, there is little consensus as to when or why this political resource curse materializes. Important studies by Haber and Menaldo (2011), Liou and Musgrave (2014), and Menaldo (2016) use novel methodological approaches to suggest that oil does not always hinder democratization, and in some instances can act as a resource blessing when compared to oil-poor states. Other arguments suggest that oil's impact on democracy is conditional on various societal or political institutions (Dunning 2008; Jones-Luong and Weinthal 2010).

Even when scholars agree on the direction of the impact, there is little cohesiveness amongst the theories linking oil wealth to lower probabilities of democratization. Some emphasize the ability of oil-wealthy incumbents to co-opt regime insiders (Wright, Geddes, and Frantz 2015). Alternatively, Girod, Steward, and Walters (2016) argue that oil-rich autocracies quell popular protests through repression. By contrast, Ross (2012) argues that oil revenues enable autocracies to provide more societal benefits than they collect through taxation, thus forestalling mass demands for democratization. Similarly, Paler (2013) draws on experimental data from Indonesia to suggest that citizens in oil-wealthy states are less willing to monitor government expenditures and exercise oversight.

How should we view these contradictions? Methodological differences explain some of the empirical debate. Andersen and Ross (2014), for instance, argue that Haber and Menaldo's (2011) evidence of a resource blessing is a function of their sample and regression design. On the theoretical side, it is likely that some of these effects are operating in different subsets of resource-wealthy countries, as Herb (2017) has recently argued, and nothing precludes multiple pathways from being present in a single case. Nevertheless, acknowledging this causal complexity and more precisely measuring specific pathways are not mutually exclusive goals. As Smith (2017) notes, maturation of the resource curse literature depends on bridging the gap between the theoretical sophistication of existing arguments and the operational terms in which they are evaluated.

This article attempts to bridge this gap by re-examining the field's foundational theory of rentier states. The original rentier state thesis emerged from the study of a small number of oil producing countries and, in general terms, emphasized the easy access to massive revenues as the root cause of the oil's autocratic effects (Luciani 1987; Crystal 1995; Chaudhry 1997). With little need to tax citizens, rulers of rentier states generated legitimacy through spending programs that resulted in fewer societal demands for representation and the absence of proto-democratic institutions (Vandewalle 1998; Karl 1997). In the large-N context, research on the rentier thesis falls into “demand side” and “supply side” versions (Ulfelder 2009, 997). In the former, oil states substitute external rents for tax revenue thus circumventing a key avenue by which citizens develop demands for greater representation (e.g. Schumpeter 1918/1954; Bates and Lien 1985). By contrast, the supply side variant holds that resource wealth enables governments to prevent the emergence of domestic political challengers. Most studies here invoke a “spending effect” (Ross 2001, 333) whereby resource rents fund patronage and other societal benefits that increase political support for existing leaders.

A number of studies provide empirical support for various parts of these claims. On the demand side, recent studies demonstrate that increased hydrocarbon revenues correlate with lower levels of taxation (Lucas and Richter 2016), that higher taxes relative to services correlate with increases in the level of democracy (Ross 2004), and that increased oil rents correlate with reduced levels of transparency (Vadlamannati and De Soysa 2016). Similarly, innovative studies by Paler (2013), McGuirk (2013), and Mahdavi (2016) use micro-level data to demonstrate that windfall spending reduces citizen’s demands for accountability. Recent sub-national studies from the United States (Goldberg, Wibbels, and Mvukiyehe 2008), Argentina (Diaz-Rioseco 2016), and Iran (Mahdavi 2015) also provide support for the supply side versions of this rentier thesis. Each demonstrates that increased oil and gas revenues tend to reduce political competition and increase incumbency advantage when leaders have discretion over how these rents are used.

At the country level, however, evidence for this spending mechanism is rare. Part of this is explained by debate over what constitutes such benefits and the absence of data measuring their size. Consequently, the few studies that attempt to evaluate this mechanism use too blunt of instruments. Morrison (2015), for instance, focuses on aggregate levels of government spending and demonstrates that this tends to stabilize regimes. However, general government spending data are especially problematic for evaluating the behavior of oil states, where off-budget spending mechanisms mean such measures underestimate the extent of service provision, while corruption that prevents designated spending from reaching targeted populations means such measures simultaneously overestimate spending levels (Mazaheri 2017).

Nor can the absence of such evidence be explained by the field’s disinterest in the underlying logic of this argument. To the contrary, this spending emphasis is central to Ross’ (2012, 67-71) theory of how oil forestalls democratic transitions. It is also consistent with theory and evidence from a number of other research areas in the discipline. For instance, the strategic provision of goods in order to retain power is consistent with Bueno de Mesquita et al.’s (2005) selectorate model, which has recently been applied to the unique politics of oil states by Liou and Musgrave (2016). These authors show that the more visible and costly the public good, the lower the chance of any would-be challenger outbidding the incumbent, and thus the more secure the incumbent's hold on power. Viewed through this 'supply side' lens, spending on public goods reduces the chance that a domestic political opposition emerges.

It is also consistent with the broad literature on distributive politics, particularly the recent findings that programmatic policies like conditional cash transfers and agricultural subsidies generate similar electoral rewards for incumbents as discretionary, or clientelistic, spending (De La O 2013; Labonne 2013; Zucco 2013; Dionne and Horowitz 2016). Such policies are programmatic in the sense that they are not subject to direct manipulation by politicians and are targeted to a broad class of citizens rather than identifiable core supporters or potential swing voters. Nevertheless, mounting evidence suggests that such policies also benefit incumbent governments, either through a retrospective voting process (i.e. programs that improve citizens' material well-being yield political rewards) or voting reciprocity (i.e. voters select politicians based on past behaviors and outcomes, rather than promises of future policies). This emphasis on programmatic goods is most applicable to the rentier setting. Indeed many of the societal benefits oil wealth funds – such as health care, education, or as this paper describes below, subsidized gasoline – are non-excludable public goods, even if they are consumed in different quantities. These varied theoretical connections make it even more important to subject this spending mechanism to careful empirical investigation.

The remainder of this article takes up this task. I develop a new measure of one such benefit, namely the size of consumer fuel subsidies in annual dollars per capita, which is available for most countries starting in the year 1990. I then employ a number of regression approaches and demonstrate a consistently negative and significant effect of fuel subsidies on the likelihood of democratization episodes. More importantly, the impact is large; more spending on fuel subsidies is just as effective at warding off transitions in autocracies as high rates of economic growth. Including this measure also helps account for some of the autocratic effect of oil wealth generally, suggesting that the classic rentier mechanism remains a likely explanation for the political resource curse.

Fuel Subsidies as Rentier Spending: Theory and Measurement

The specific benefit I measure focuses on the fiscal resources governments must devote to keeping gasoline prices low for domestic consumption. Many petro-states – which almost invariably have nationalized oil industries – sell gasoline domestically at rates far below their international comparison. In some instances, they are almost comically low; in 2014, a liter of gasoline in Venezuela cost less than $0.02, roughly eighty-seven times lower than the world average for the year.[1] Though an extreme example, oil production is associated with cheaper gasoline generally (Cheon, Lackner, and Urpelainen 2013).

In some instances, low extraction costs enable gasoline to be sold to domestic consumers at low prices, though this decision to sell at a rate below world market prices costs governments billions in forgone revenue. In oil producing states that lack sufficient refinement capacities, governments also provide suppliers of gasoline with direct payments from the central (or in some case, sub-national or local) budgets. In Iran, for instance, the National Oil Company receives government compensation to cover the difference between the costs of imported gasoline and the (artificially low) mandated domestic price for consumers (IMF 2008). Low prices can also be maintained by central governments absorbing key costs of gasoline production directly, whether through use of stabilization funds to offset the price gaps between imported gasoline and the raw material costs, as in Brazil through the 1980s, or currently in Nigeria. Unsurprisingly, these crowd out other spending priorities and can result in crippling deficits. For instance, Yemen spends more subsidizing domestic gasoline consumption (ten percent of GDP in 2012) than on all other infrastructure and social services combined (Clements et al. 2013).

Why focus on fuel subsidies specifically, as opposed to the various other private and public goods survival-minded leaders might provide? First, as noted above, fuel subsidies are particularly easy to implement in oil-producing nations and therefore represent a useful way to evaluate the underlying theoretical argument that oil and autocracy are linked through a rentier spending mechanism. Second, end-use consumer fuel subsidies are high-profile policies that are easily attributable to governments, which is necessary if citizens are to reward incumbents with political support through the retrospective or reciprocity mechanisms emphasized above. Indeed, citizens are acutely aware of the size and impact of subsidies and changes to the subsidy regime can be politically costly. Nigeria's decision in January 2012 to raise the domestic price of gasoline to a cost-recovery level – which would have the effect of more than doubling retail prices – was met with widespread

[1] These figures are reported by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ, or German Corporation for International Cooperation), a development agency that frequently works in cooperation with the German Federal Ministry for Economic Cooperation and Development. This agency serves as the main data source described below.

social unrest and paralyzing strikes that caused the government to reverse the policy a mere two weeks later. Protests, often violent, have also accompanied recently planned subsidy reforms in at least a dozen other nations, including Sudan, Jordan, Bolivia, and Indonesia (Lockwood 2015). Even when the response is not violent, reform is always politically controversial, so it is unsurprising that governments seldom fully implement such plans (Sivaram and Harris 2016).

The focus on fuel subsidies also accords with the general logic of the selectorate model. Though fuel subsidies are a public good that all citizens can consume, ample evidence suggests that they are most frequently consumed by segments of the population whose continued loyalty is most critical for regimes seeking to remain in power. Urban, relatively well-off segments of a society are more likely to consume gasoline than are less affluent or rural constituencies (Del Granado et al. 2012; Coady et al. 2010). Energy policy scholars also recognize that these subsidies are common political tools for generating political support, particularly amongst influential societal groups (Rentschler and Bazilian 2016). Finally, there are clear pragmatic advantages; a measure of fuel subsidies can be constructed without relying on notoriously sparse (and inaccurate) data on government spending, particularly in autocracies.

My measurement of this variable follows prior scholarship on energy subsidies. Specifically, I adopt the "price-gap" approach which compares the domestic price of a good to a benchmark price the good would fetch in a competitive market (Kojima and Koplow 2015, 6). Subsidies exist whenever the domestic price is lower than the benchmark price. Though this approach does not capture all distortionary energy subsidies, such as those that increase revenues retained by resource owners or reduce production costs, it is a useful way to capture subsidies whose main goal is to reduce costs for end-use consumers (Coady et al. 2010).

The specific variable I construct is a measure of the total fiscal burden governments incur by subsidizing the cost of domestic gasoline, reported in per-capita US dollar terms. The variable (hereafter *fuel subsidies per capita*) is constructed by first estimating annual measures of the extent to which a single liter of gasoline is subsidized in a country, multiplying this by an estimate of the total liters consumed annually, and finally dividing this by population so the variable is reported in per capita terms.[2] Starting with the price-gap approach allows me to measure the fiscal burden of these

[2] Part B of the Supplementary File describes the variable construction in detail.

subsidies across a variety of subsidy-delivery methods, including forgone revenue from price ceilings, transfer payments, or direct cost absorption by governments.[3]

Constructing the first part of this variable requires measures of domestic gasoline prices and a benchmark price for comparison. Data recently collected by the IMF (Coady et al. 2010), supplemented by additional biennial data from the German Corporation for International Cooperation (*Deutsche Gesellschaft für Internationale Zusammenarbeit*, or GIZ) allows for the construction of panel data reporting annual per-liter costs of gasoline for a large sample of countries from 1990 to 2014. Prices are measured in November of each year and are reported in constant US cents per liter.[4]

I use the price per liter in the United States as the benchmark price. US prices represent a "revenue neutral" international benchmark, meaning that prices cover transportation and distribution costs along with industry margins, though do not include sufficient taxes to offset the environmental or health costs associated with fossil fuel usage (GIZ, various years). For countries that import oil, I follow the IMF's approach (e.g. Coady et al. 2010) and add 20 cents per liter to the US price to cover additional transportation and distribution costs.[5] I then subtract the domestic price paid per liter from the US-based benchmark price. When the domestic price is lower than the benchmark, the result is an estimate of the per-liter subsidy in a given country for a given year. When domestic prices are higher, I follow existing conventions and interpret this as the absence of a subsidy and set the per-liter value to 0.[6] These price-gaps are calculated in constant US cents per liter, but I divide by 100 so that the value is denominated in dollars.

---

[3] Cheon, Lackner, and Urpelainen (2015) examine the impact of national oil companies on the size of subsidies. However, their measurement of subsidies is simply the observed gasoline price in a country, so they cannot capture the fiscal resources necessary to maintain low prices.

[4] Like nearly all sources of cross-national time-series data, both sources contains significant missing values. Part B of the Supplementary Files details how these missing values were replaced with country-specific interpolated values to maximize the coverage of the final variable *fuel subsidies per capita*. This section also demonstrates that the results are robust to using multiple imputation methods as an alternative solution to the missing data issues.

[5] I use Ross' (2012) data to distinguish oil importer status.

[6] Even when prices are equal to the world benchmark price, governments are in effect subsidizing gasoline consumption, since the prices are insufficient to cover the (sizable) negative environmental

Next, I multiply these yearly per-liter subsidy values by a measure of that years' total liters of gasoline consumed, the latter of which I construct from the EIA's cross-national time-series data on fuel consumption.[7] Finally, I divide this total by population so the variable becomes weighted in per capita terms. The resulting variable, *fuel subsidies per capita*, ranges from $0 to $801 (Saudi Arabia in 2012) in the estimation sample. As the variable is highly skewed, subsequent empirical analyses using this variable employ a log transformation of the measure.[8]

Of course, artificially-low prices for gasoline encourage excessive consumption; this is the chief economic and environmental criticism of subsidies more generally. Why do I multiply the per-liter subsidy by the total volume of fuel consumption? First, such a transformation is required to measure the total value of a subsidy (Kojima and Koplow 2015, 14) and thus the total fiscal burden for governments. Relatedly, it is fair to assume that regime insiders understand how low prices drive up consumption – indeed, the logic of rentier theory suggests that policy makers fully understand and depend on this consequence. My measure, therefore, best corresponds to the theoretical concept emphasized throughout the rentier literature. Nevertheless, I show in the Supplementary File that the main results are robust to using the per liter price-gap as the measure of fuel subsidies.

Figure 1 demonstrates the relationship between my measure of fuel subsidies and a measure of oil wealth. The vertical axis reports *fuel subsidies per capita (log)*, averaged over the years 1990 to 2014, which is the full temporal span for which the variable can be constructed. The horizontal axis reports Ross' (2012) variable *oil income per capita (log)* averaged over these same years.

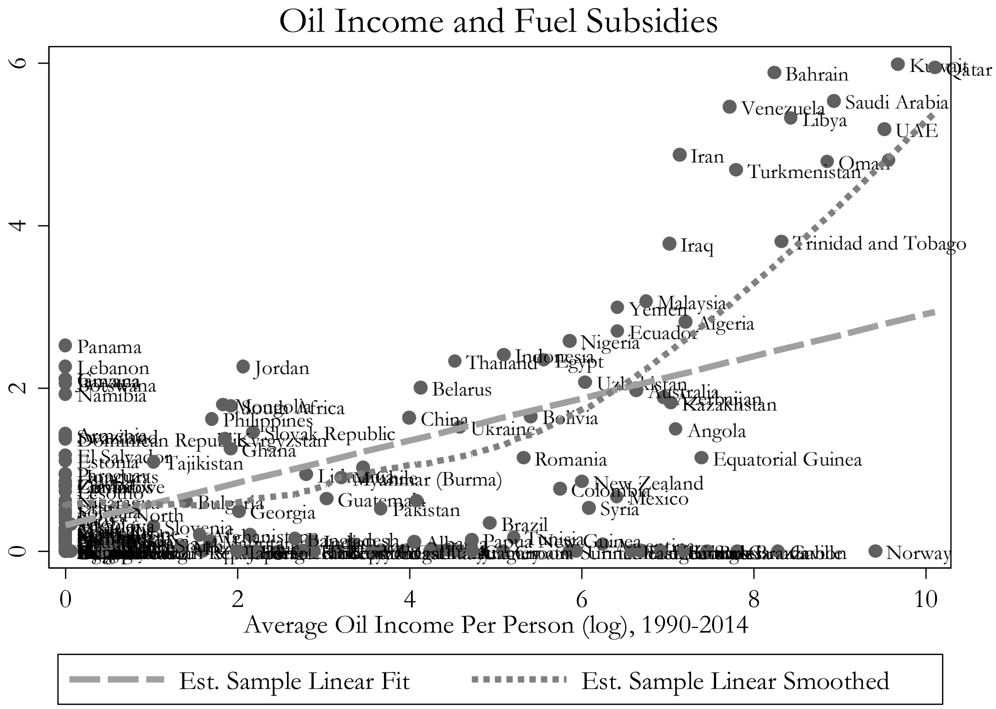
---

and public health externalities. Exploring why some countries offset the negative externalities of fuel consumption with higher taxes is an important area for future research, but outside the scope here

[7] Available from http://www.eia.gov/beta/international/data

[8] I follow Ross' (2012) and add 1 to the zero values of subsidies before computing the natural log.

Figure 1



The figure presents both a linear fit regression line and a locally-weighted smoothed line for the estimation sample. Each demonstrates that oil wealth tends to be associated with higher levels of domestic fuel subsidies. Oil wealth is neither necessary nor sufficient for subsidizing gasoline consumption; Norway is a major oil producer but does not subsidize domestic consumption of gasoline, and a number of non-oil producing countries devote considerable amounts of revenue to keeping prices low. Nevertheless, there is a strong relationship between oil wealth and the resources governments devote to keeping domestic gasoline prices low.

Empirical Approach

I now turn to evaluating whether greater spending on fuel subsidies reduces the likelihood that a country experiences a transition toward a more electorally democratic regime. I use a quantitative model with a global sample of countries from 1990 to 2014. The empirical specification draws on Bermeo's (2016) recent study, though I utilize the measure of electoral democracy from the Varieties of Democracy project (Teorell et al. 2016) as the dependent variable instead of polity scores. This measure, referred to as the *polyarchy index*, is composed of five components: the election of officials, the existence of free and fair elections, freedom of expression, associational autonomy, and inclusive citizenship. The measure is aggregated into a single index, scaled from 0.0 to 1.0.

To analyze whether fuel subsidies impact transitions toward electoral democracy, I follow Lindberg's (2015) approach and create a 5-part ordinal version of the *polyarchy index*: closed autocracies, autocracies, ambivalent, minimally democratic, and democratic. The dependent variable in this analysis is thus a binary indicator of whether a regime experienced at least a one-step increase in this ordinal variable (i.e. minimally democratic to democratic) in a given year. This coding rule means that the sample excludes country-years coded fully democratic.[9] Using this measure also avoids some of the problems that can arise when using a binary measure based on a certain size increase in the *polity2* score, such as uncertainty as to what specific component of a polity score accounted for the change. Nevertheless, the Supplementary File demonstrates that the results are robust to measuring transitions with a polity-based variable. To match the timing of this measure, I lag independent variables by one period. I also exclude from the sample any country-years where regimes are not fully autonomous. This prevents foreign-imposed regime changes from counting as democratization episodes.

Control variables include measures of *income per capita (log)* and the rate of *economic growth*. Both of these variables are taken from the World Development Indicators. I also use Ross' aforementioned measure of oil wealth, *oil income per capita (log)*, to account for the conventional findings of the resource curse. The models also include a country's lagged *polyarchy index (ordinal)* score, as higher scores mean there is less of a possibility for further positive change, and a sum of the past transitions toward democracy from 1960 onward. All models include yearly dummy variables to account for common temporal effects including changing oil prices and the international environment more broadly, regional dummy variables, and time polynomials of the years since the last increase in the *polyarchy index (ordinal)* score (age, age^2, and age^3) to account for temporal dependence (Carter and Signorino 2010).

I first estimate the model with a conventional probit regression with standard errors clustered on countries. This method, however, cannot distinguish whether differences in the level of subsidies between countries, or changes within countries over time, explain their varying

[9] The coding rule also means that country-years where regimes experienced democratic backsliding are not distinguished from those that remained stable. However, Table 7A in the supplemental file demonstrates that the results are robust to excluding these backsliding regimes from the sample.

propensities to experience transitions. Moreover, this approach may not sufficiently account for unit effects. The traditional solution to these issues is the inclusion of fixed effects, which model unit heterogeneity directly and isolate the within-country variation. However, as Wright et al. (2015) describe, including fixed effects with a binary dependent variable drops from the analysis all observations that did not experience a measured regime change, effectively biasing downward the estimates. This is particularly problematic if those stable political systems dropped from the analysis happen to have high oil wealth or generous fuel subsidies, as the theory above anticipates.

One solution is to model between-country and within-country effects concurrently by conditioning the marginal effects of the independent variables on country-specific mean values of these covariates. This approach – described as the "between-within" model – is recommended by Bell and Jones (2015) and, specific to the resource curse literature, employed in recent studies by Wright, Geddes, and Frantz (2015) and Smith (2017). Functionally, this approach replaces the single variable *fuel subsidies per capita* with country-specific measures of the mean of this variable (designated as *fuel subsidies per capita – country means*) to capture the between-country effect, and a country-year specific measure of the deviation from the country mean (designated as *fuel subsidies per capita – deviation from country mean*) to capture the within-country effect.[10] To account for unit heterogeneity, the country mean of the dependent variable is also included.

Results

Table 1 reports the results of the standard probit approach (Models 1 and 2) and the 'between-within' approach (Models 3 and 4) described above. As my argument claims that spending on fuel subsidies helps account for the political resource curse, the first model in each set (Models 1 and 3) excludes the measure of fuel subsidies. If the argument is accurate, including the measures of fuel subsidies in Models 2 and 4 should reduce the impact of the oil income variable.

The results of both regression approaches strongly support these expectations. In model 1, greater levels of oil income per capita demonstrate the expected negative impact on the likelihood of a democratization episode. However, the magnitude of this impact is considerably reduced by the inclusion of the fuel subsidy measure in Model 2. Importantly, the measure of fuel subsidies itself has a negative and statistically significant impact on the probability of a transition.

[10] The results are identical when replacing all time-variant variables with these decomposed versions. I retain the decomposed oil income variables given their centrality in the literature.

Table 1: Probit Regressions Results	(1) Standard Probit	(2) Standard Probit	(3) Between-Within	(4) Between-Within
Oil income per capita (log)	-0.11***	-0.07**	-	-
	(0.03)	(0.03)		
Fuel subsidies per capita (log)	-	-0.09*	-	-
		(0.05)		
Oil income per capita - cm	-	-	-0.10***	-0.06**
			(0.03)	(0.03)
Oil income per capita - deviation from cm	-	-	-0.08	-0.05
			(0.07)	(0.06)
Fuel subsidies per capita (log) - cm	-	-	-	-0.13*
				(0.07)
Fuel subsidies per capita (log) - deviation from cm	-	-	-	0.00
				(0.06)
Mean of DV	-	-	5.28*	4.68*
			(2.74)	(2.77)
Income per capita (log)	0.13	0.10	0.15*	0.13
	(0.09)	(0.10)	(0.09)	(0.09)
Economic growth	-0.01*	-0.02***	-0.01	-0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
Polyarchy Index	-1.92***	-1.79***	-1.85***	-1.78***
	(0.35)	(0.37)	(0.34)	(0.39)
Sum of past transitions	0.39***	0.38***	0.27***	0.28***
	(0.05)	(0.06)	(0.07)	(0.07)
Observations (clusters)	2387 (122)	2214 (120)	2497 (123)	2333 (121)

Notes: Probit regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. 'cm' = country mean. *** p<0.01, ** p<0.05, * p<0.1

Models 3 and 4 follow this same logic while employing the between-within approach. Though a relatively simple test, the results are stark: higher levels of fuel subsidies are associated with significantly lower prospects of transitions toward democracy, and this negative effect operates through the large differences in fuel subsides between countries as represented by the country-mean measures. Moreover, these effects hold even when we control for factors that may be associated with the level of fuel subsidies or the otherwise unobserved tendency of countries to experience transitions toward more electorally democratic regimes. Importantly, including the fuel subsidy measures continues to reduce the impact of oil income on democratization.[11]

Figures 2 and 3 address the practical impact of these findings in graphical form. Figure 2 plots the changes in the predicted probability of a transition, along with 90% confidence intervals, as a function of the variable *fuel subsidies per capita – country means*; that is, as a function of the variable capturing the between effect of the key independent variable. These results are based on Model 4, with all other variables held to their mean values. The figure shows a steep decline in the predicted probability, and these changes are statistically significant. For instance, a country that does not subsidize domestic gasoline consumption has a roughly four percent probability of an increase in electoral democracy in a given year, all else equal. If the level of fuel subsidies increased to 90% of the variables' distribution (a value of 2.5 on the X axis), the predicted probability is reduced by half. Though the reduction in probability is modest in absolute terms, this represents a substantial impact. The baseline probability of democratization in the sample is only three percent. Moreover, the substantive impact of fuel subsidies is comparable to other variables in the model. For instance, a shift in the annual rate of economic growth from the 10th percentile of the distribution (a rate of -2.7%) to the 90th percentile (a rate of 7.6%) would result in almost the exact same reduction in the probability of democratization: from roughly four percent to two percent.

[11] Table 1A in the supplemental material demonstrates that this main result is robust to matching sample sizes across models 1 and 2 and to excluding the measure(s) of oil income from the models.

Figure 2

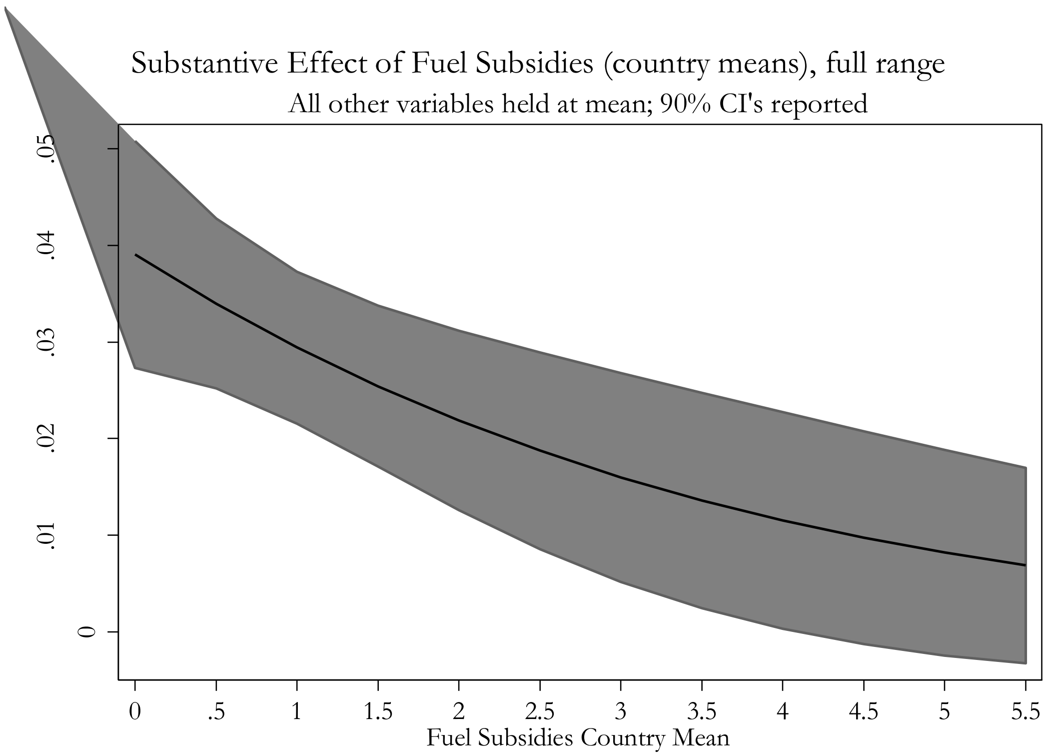
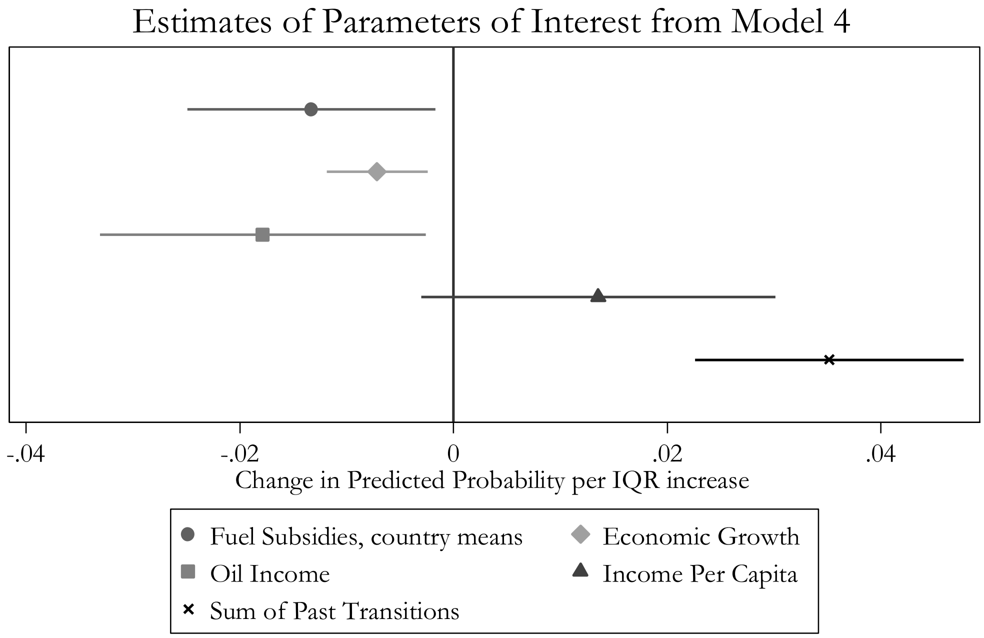


Figure 3 provides a more general view of the findings' practical significance. This figure draws on Gross' (2015) recommendations and compares five parameters of interest. Specifically, the figure displays the change in the predicted probability of a transition toward democracy resulting from an inter-quartile range (IQR) increase in the measures of *fuel subsidies*, *oil income*, *economic growth*, *income per capita*, and *past transitions*. As these are estimates of the change in predicted probability, the figure also includes confidence intervals. The figure demonstrates that an IQR increase in *fuel subsidies* has a larger negative impact than does a similar change in the rate of *economic growth*, and almost as large as impact as does the level of oil income in a country. Figure 3 demonstrates that a history of past transitions toward democracy is an important predictor of future transitions – the estimated change in predicted probability per IQR increase is about 3.5% – though note that the impact of *fuel subsidies* is more than half this size in absolute value. Put differently, a relatively small increase in the size of *fuel subsidies* (from the 25th percentile to the 75th percentile) would effectively offset the positive impact of a past transition.

Figure 3



Note: Estimates are expressed as 90% confidence intervals of the change in the predicted probability of a democratic transition per interquartile range increase in each explanatory variable. Predicted probabilities calculated while holding all other explanatory variables at mean values.

An important tradeoff to measuring this mechanism more directly than prior studies is, of course, that government spending on fuel subsidies may be endogenous to regime dynamics. The clearest example of endogeneity – an autocratic regime on the verge of a transition to democracy suddenly reducing domestic fuel prices in an attempt to thwart regime change – is less of a concern given that the results above emphasize the country-means of the subsidy variable, calculated as averages over many years. Nevertheless, the findings above warrant a more careful approach to identification.

Several prior studies have relied on instrumental variables (IV) to deal with related concerns (e.g. Ramsay 2011) and there are many factors that both explain why governments want to keep prices low and account empirically for some of the variation in fuel subsidies globally. For example, subsidies tend to be positively correlated with the price of oil; governments of resource-wealthy states ratchet up subsidies to prevent passing along the full costs of rising oil prices to their domestic consumers (Victor 2009).[12] Subsidies also compensate for episodes of high inflation or periods of low economic growth, offset the need for broader infrastructure projects in urban environments, or even emerge as a response to low state capacity (e.g. Del Granado et al. 2012; Kim and Urpelainen 2015; Cheon, Lackner, and Urpelainen 2013). Each of these factors, however, has well known conceptual and empirical links to political regimes. Thus, while such concepts are correlated with fuel subsidies, they would violate the exclusion restriction imposed on traditional IV methods. Going further afield theoretically may eventually identify an instrument that does not violate the exclusion restriction, but if it comes at the expense of a weaker instrument, then the cure may be worse than the disease, as Young (2017) has recently demonstrated.

As an alternative, I employ Lewbel's (2012) method, which identifies two-equation systems by constructing 'instruments' as a simple function of the model's data. This approach, which requires only that errors in a linear projection of the endogenous regressor on the other regressors in the main equation be heteroskedastic, can serve as a source of identification when instruments meeting the standard exclusion restriction are unavailable. For instance, consider the stylized version of the two-equation model considered here, where fuel subsidies is an endogenous (or mis-measured) predictor of democratic transitions. In the standard IV approach, Z is an instrument that is correlated with the extent of fuel subsidies but unrelated to the outcome in equation (1), democratic transitions.

$$\textit{Democratic transition} = \beta_1 X + \gamma_1 \textit{Fuel Subsidies} + \varepsilon_1 \tag{1}$$

$$\textit{Fuel Subsidies} = \beta_1 X + \gamma_2 Z + \varepsilon_2 \tag{2}$$

Lewbel (2012) demonstrates that if the errors in equation (2) are heteroskedastic, which is plausible in many applications, then the residuals from this first-stage regression can be used to construct instruments that are orthogonal to the main outcome equation error process. These

[12] More generally, Smith (2017, 601) argues that any measure of oil production is endogenous to political regimes, as oil companies, and particularly the multi-nationals responsible for production globally, must have long time horizons in order to justify the huge sunk costs involved in oil exploration and production, time horizons which are likely impacted by the perceptions of the regime in place.

instruments are described as "internal," given that they are constructed from the data used to estimate the model as opposed to some external design feature or theory. Lewbel shows that this approach can both increase the efficiency of traditional IV estimation and, critically, provide a source of identification when there is no instrument meeting the exclusion restriction typically imposed by the two-equation framework; that is, when no Z is available as a plausibly exogenous instrument for *fuel subsidies*. Since these internal instruments are constructed using heteroskedasticity in the first-stage regression, the main equation can be identified without threats to inference.

In practical terms, these internal instruments are constructed by first calculating the residuals from a first-stage regression of the endogenous regressor on all other regressors in the main model, then multiplying these residuals by each exogenous variable in the model in a demeaned version (i.e. $X_{i,t} - \bar{X}$). The net result of this is as a set of "internal" instruments (one for each exogenous variable in the model) constructed from information in the first-stage and thus capable of identifying the main outcome equation. The greater the degree of heteroscedasticity, the stronger the correlations between these generated instruments and the endogenous regressor. Importantly, this approach is still subject to conventional diagnostics on the strength of these constructed instruments (F statistics) and whether they are independent of the errors in the outcome model (J statistics), which I report below. Though relatively new, this method has been recently employed in the political science literature by Dietrich and Wright (2015) and Wilson and Wright (2015), who use the approach to assist in identification for their respective studies on foreign aid allocation and democratic change in Africa and autocratic legislatures and expropriation risk.

Table 2 uses these "internal" instruments to reanalyze two of the models discussed above. I follow the specifications in Models 2 and 4. As these are two-stage regression models, each column reports a number of diagnostic test statistics, including measures of the first-stage correlation (the F statistic) and the J statistic confirming that the instruments generated through Lewbel's (2012) process are independent of the errors in the outcome model and sufficiently correlated with the endogenous measure of fuel subsidies.[13]

[13] Table 5A in the Supplementary Files reports the results of Pagan-Hall tests of heteroscedasticity in the models – a necessary component for the Lewbel (2012) approach. Each model generates a test statistic far larger than necessary to reject the null hypothesis of homoscedastic errors.

Table 2: IV estimates using Lewbel (2012) instruments

	(5) Standard	(6) Between-Within
Oil income per capita (log)	-0.005*	-
	(0.003)	
Fuel subsides per capita (log)	-0.013*	-
	(0.007)	
Oil income per capita (log) -cm	-	-0.005*
		(0.003)
Oil income per capita (log) - deviation from cm	-	-0.001
		(0.005)
Fuel subsidies per capita (log) - cm	-	-0.015**
		(0.007)
Fuel subsidies per capita (log) - deviation from cm	-	0.001
		(0.005)
Mean of DV	-	0.619*
		(0.350)
Income per capita (log)	0.012	0.016
	(0.010)	(0.010)
Economic growth	-0.003**	-0.003**
	(0.001)	(0.001)
Polyarchy Index	-0.199***	-0.196***
	(0.034)	(0.032)
Sum of past transitions	0.048***	0.034***
	(0.006)	(0.008)
Kleibergen-Paap rk Wald F statistic	16.97	30.68
10% max relative bias	11.32	11.27
J statistic (p value)	25.6 (0.65)	33.2 (0.45)
Observations (clusters)	2214 (120)	2333 (122)

Notes: Linear Probability model regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. *Fuel subsides per capita (log)* and *Fuel subsidies per capita (log) - country means* are instrumented using Lewbel internal instruments. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. 'cm' = country mean
*** p<0.01, ** p<0.05, * p<0.1

One critical difference in the modeling reported in Table 2 is that these regressions are estimated via linear probability models. Current implementation of Lewbel's (2012) procedure is limited to this approach (e.g. Baum and Schaffer 2012), as opposed to the probit models presented in Table 1. Consequently, the estimated regression coefficients will differ from those in the probit model, thought note that Table 1A in the Supplementary Files provides a source of comparison by re-estimating the regressions in Table 1 via linear probability models.

However, the main findings are substantively identical to those in Table 1. In Model 5, increasing fuel subsidies has a statistically significant negative impact on the likelihood of a regime transition in the following year, while Model 6 demonstrates that this is driven by the between-country differences and persists even when the model accounts for unit heterogeneity. The coefficients of interest are roughly the same size as those reported in the naïve linear probability models reported in Table 2A in the Supplementary Files, as are the marginal effects.

Finally, note that these results are robust to a number of measurement and model specification changes as reported in the Supplemental File, part A. Table 3A replicates the core findings using changes in a country's *polity2* score as the measure of regime transition. Table 4A demonstrates that these results are robust to replacing the measure of *fuel subsidies per capita* with a simple measure of per-liter price gaps. Table 6A re-estimates the main while including a measure of total social expenditures, using data described in Lucas and Richter (2016), in order to probe the specific consequences of fuel subsidy spending versus more general levels of expenditures. Data limitations restrict the sample considerably, but the negative effects of fuel subsidies directly persists. Finally, section III of Supplementary Files (Part B) demonstrates that the results are consistent even when using multiple imputation to deal with a variety of missing data issues that are common in this type of analysis generally and that Lall (2017) shows are particularly important in this literature specifically.

Conclusions

Further maturation of the resource curse literature requires matching the theoretical sophistication of the field with careful empirical analysis that can evaluate and compare the impact of various mechanisms. This note approached this challenge by measuring the fiscal resources regimes devote to keeping gasoline prices low. Using the measure in an empirical analysis of transitions toward more electorally democratic regimes confirmed a core insight of rentier theory, namely that as these benefits increase, the likelihood of a change toward a more democratic government decreases. Though the continued significance of oil income in the regression models validates Herb's (2017) argument about causal complexity in the resource curse, it does appear that citizens in non-democracies pay a high political price for their cheap gasoline.

Nevertheless, scholars should continue chasing this mechanism further back. For instance, what kinds of behaviors are discouraged by these sorts of benefits? Do citizens benefiting from these low prices demand less accountability (a hypothesis that could be evaluated with cross-national survey data)? The results should also encourage further exploration of subsidies as a programmatic distributive policy. Do greater subsidies generate specific electoral benefits for incumbents, such as increased turnout or vote share? Such effects may be particularly important in understanding the dynamics of hybrid or competitive authoritarian regimes, where democratization is forestalled not by the complete elimination of domestic opposition but rather by tilting the electoral playing field in the incumbents favor. What other sorts of spending programs are correlated with generous fuel subsidies? Regimes that subsidize fuel are also likely to spend more on other benefits. A better accounting of these processes will shed further light on the role of rentierism in understanding regime dynamics.

The new measure of fuel subsidies also opens up a number of additional avenues of inquiry. For instance, the cost and inefficiencies of fuel subsidies make them common targets for elimination under IMF lending programs. Do borrowing nations actually reduce or eliminate these subsidies? Alternatively, are fuel subsidies modified when the price of oil – and therefore the revenues petro-states receive and the underlying need to subsidize domestic consumption – declines? Existing work suggests that reform efforts, even in light of diminished oil revenues, are rarely implemented (Lockwood 2015), though this may be changing as the oil-producing world enters a more sustained period of low prices. Recent announcements in countries as diverse as Saudi Arabia and Nigeria suggest that current leadership intends to reduce fuel subsidies. The risk of incomplete implementation or rollback remains high, however, in part because of public outcry over increased prices (Sivaram and Harris 2016). Recent work by Kyle (2018) identifies an important political barrier to reform, namely that citizens may not trust the government to offset the pain of subsidy cutbacks with other welfare enhancing policies. More generally, fuel subsidies are one of the few concepts that sit squarely at the intersection of political, economic, and environmental consequences, giving the topic a unique combination of scholarly and policy relevance. For these reasons alone, future work should take up the challenge to understand both the causes and consequences of these costly subsidies.

Supplemental Information

Supplemental information for this paper includes an online appendix containing various robustness analyses and a detailed description of the fuel subsidy variable construction, as well as the data and code to replicate the findings. These materials are available at the author's website (https://sites.google.com/a/oakland.edu/mfails/research) and at the International Studies Quarterly data archive.

References

Ahmadov, Anar K. 2014. "Oil, Democracy, and Context A Meta-Analysis." *Comparative Political Studies* 47(9): 1238-1267.

Andersen, Jørgen, and Michael Ross. 2014. "The Big Oil Change: A Closer Look at the Haber–Menaldo Analysis." *Comparative Political Studies* 47(7): 993-1021.

Bates, Robert H., and Da-Hsiang Donald Lien. 1985. "A Note on Taxation, Development, and Representative Government." *Politics & Society* 14(1): 53-70.

Baum, Christopher, and Mark Schaffer. 2012. "Ivreg2h: Stata module to perform instrumental variable estimation using heteroskedasticity-based instruments." http://ideas.repec.org/c/boc/bocode/s457555.html

Bell, Andrew, and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3(1): 133-53.

Bermeo, Sarah Blodgett. 2016. "Aid is Not Oil: Donor Utility, Heterogenous Aid, and the Aid-Democratization Relationship." *International Organization* 70(1): 1-32.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph Siverson, and James Morrow. 2005. *The Logic of Political Survival.* Cambridge: The MIT Press.

Carter, David B., and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18(3): 271-292.

Chaudhry, Kiren Aziz. 1997. *The Price of Wealth: Economies and Institutions in the Middle East*. Ithaca, NY: Cornell University Press.

Cheon, Andrew, Johannes Urpelainen, and Maureen Lackner. 2013. "Why do Governments Subsidize Gasoline Consumption? An Empirical Analysis of Global Gasoline prices, 2002–2009." *Energy Policy* 56: 382-390.

Cheon, Andrew, Maureen Lackner, and Johannes Urpelainen. 2015. "Instruments of Political Control National Oil Companies, Oil Prices, and Petroleum Subsidies." *Comparative Political Studies* 48(3): 370-402.

Clements, Benedict, David Coady, Stefania Fabrizio, Sanjeev Gupta, Trevor Alleyne, and Carlos Sdralevich. 2013. *Energy Subsidy Reform: Lessons and Implications*. Washington DC: The International Monetary Fund.

Coady, David, Robert Gillingham, Rolando Ossowski, John Piotrowski, Shamsuddin Tareq, and Justin Tyson. 2010. "Petroleum Product Subsidies: Costly, Inequitable, and Rising." *IMF Staff Position Note 10/05.*

Crystal, Jill. 1995. *Oil and Politics in the Gulf: Rulers and Merchants in Kuwait and Qatar* (Vol. 24). Cambridge: Cambridge University Press.

del Granado, Francisco Javier Arze, David Coady, and Robert Gillingham. 2012. "The Unequal Benefits of Fuel Subsidies: A Review of Evidence for Developing Countries." *World Development* 40(11): 2234-48.

De La O., Ana. 2013. "Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico." *American Journal of Political Science* 57(1): 1-14.

Deutsche Gesellschaft für International Zusammenarbeit ("GIZ"), various years. "International Fuel Prices." Available from https://www.giz.de/expertise/html/4282.html

Dietrich, Simone, and Joseph Wright. 2015. "Foreign Aid Allocation Tactics and Democratic Change in Africa." *Journal of Politics* 77(1): 216-234.

Diaz-Rioseco, Diego. 2016. "Blessing and Curse: Oil and Subnational Politics in the Argentine Provinces." *Comparative Political Studies*, 49(14): 1930-64.

Dunning, Thad. 2008. *Crude Democracy: Natural Resource Wealth and Political Regimes.* Cambridge: Cambridge University Press.

Girod, Desha M., Megan A. Stewart, and Meir R. Walters. 2016. "Mass Protests and the Resource Curse: The Politics of Demobilization in Rentier Autocracies." *Conflict Management and Peace Science*. DOI: 0738894216651826.

Goldberg, Ellis, Erik Wibbels, and Eric Mvukiyehe. 2008. "Lessons from Strange Cases: Democracy, Development, and the Resource Curse in the US States." *Comparative Political Studies* 41(4-5): 477-514.

Gross, Justin H. 2015. "Testing What Matters (If You Must Test at All): A Context-Driven Approach to Substantive and Statistical Significance." *American Journal of Political Science* 59(3): 775-88.

Haber, Stephen and Victor Menaldo. 2011. "Do Natural Resources Fuel Authoritarianism? A Reappraisal of the Resource Curse." *American Political Science Review* 105(1): 1-26.

Herb, Michael. 2017. "Ontology and Methodology in the Study of the Resource Curse." *LSE Kuwait Program Paper Series 43.* Available at http://eprints.lse.ac.uk/82871/1/MichealHerb_%2043_Ontology_2017.pdf

International Monetary Fund. 2008. "Islamic Republic of Iran: Selected Issues." IMF Country Report No. 08/285.

Jones Luong, Pauline, and Erika Weinthal. 2010. *Oil is Not a Curse: Ownership Structure and Institutions in Soviet Successor States*. Cambridge: Cambridge University Press.

Karl, Terry Lynn. 1997. *The Paradox of Plenty: Oil Booms and Petro-states.* Los Angeles: University of California Press.

Kim, Sung Eun, and Johannes Urpelainen. 2015. "Democracy, Autocracy and the Urban Bias: Evidence from Petroleum Subsidies." *Political Studies* 64(3): 552-72.

Kojima, Masami, and Dough Koplow. 2015. "Fossil Fuel Subsidies: Approaches and Valuation." *World Bank Policy Research Working Paper 7720.*

Kyle, Jordan. 2018. "Local Corruption and Popular Support for Fuel Subsidy Reform in Indonesia." *Comparative Political Studies*. DOI: 10.1177/001041401878755

Labonne, Julien. 2013. "The Local Electoral Impacts of Conditional Cash Transfers: Evidence from a Field Experiment." *Journal of Development Economics* 104: 73-88.

Lall, Ranjit. 2017. "The Missing Dimension of the Political Resource Curse Debate." *Comparative Political Studies* 50(10): 1291-324.

Lewbel, Arthur. 2012. "Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models." *Journal of Business & Economic Statistics*, 30(1): 67-80.

Lindberg, Staffan I. 2015. "Ordinal Versions of V-Dem's Indices: For Classification, Description, Sequencing Analysis and Other Purposes." *V-Dem Working Paper* 2015:20. Available at http://dx.doi.org/10.2139/ssrn.2704786

Liou, Yu-Ming, and Paul Musgrave. 2016. "Oil, Autocratic Survival, and the Gendered Resource Curse: When Inefficient Policy Is Politically Expedient." *International Studies Quarterly* 60(3): 440-56.

Liou, Yu-Ming, and Paul Musgrave. 2014. "Refining the Oil Curse Country-Level Evidence From Exogenous Variations in Resource Income." *Comparative Political Studies* 47(11): 1584-610.

Lockwood, Matthew. 2015. "Fossil Fuel Subsidy Reform, Rent Management and Political Fragmentation in Developing Countries." *New Political Economy* 20(4): 475-94.

Lucas, Viola, and Thomas Richter. 2016. "State Hydrocarbon Rents, Authoritarian Survival and the Onset of Democracy: Evidence from a new dataset." *Research & Politics* 3(3): 1-9.

Luciani, Giacomo. 1987. "Allocation versus Production States: A Theoretical Framework." In *The Rentier State*, edited by Hazem Beblawi and Giacomo Luciani, 65-85. New York: Croom Helm.

Mahdavi, Paasha. 2015. "Explaining the Oil Advantage: Effects of Natural Resource Wealth on Incumbent Reelection in Iran." *World Politics* 67(02): 226-67.

Mahdavi, Paasha. 2016. "No Taxation, No Representation? Oil-to-Cash Transfers and the Dynamics of Government Responsiveness." Unpublished manuscript.

Mazaheri, Nimah. 2017. "Oil, Dissent, and Distribution." *World Development* 99: 186-202.

McGuirk, Eoin F. 2013. "The Illusory Leader: Natural Resources, Taxation and Accountability." *Public Choice* 154(3-4): 285-313.

Menaldo, Victor. 2016. *The Institutions Curse: Natural Resources, Politics, and Development*. Cambridge: Cambridge University Press.

Morrison, Kevin M. 2015. *Nontaxation and Representation*. Cambridge: Cambridge University Press.

Paler, Laura. 2013. "Keeping the Public Purse: An Experiment in Windfalls, Taxes, and the Incentives to Restrain Government." *American Political Science Review* 107(4): 706-25.

Ramsay, Kristopher. 2011. "Revisiting the Resource Curse: Natural Disasters, the Price of Oil, and Democracy." *International Organization* 65(3): 507-29.

Rentschler, Jun, and Morgan Bazilian. 2016. "Reforming Fossil Fuel Subsidies: Drivers, Barriers and the State of Progress." *Climate Policy* 17(7): 891-914.

Ross, Michael L. 2001. "Does Oil Hinder Democracy? *World Politics* 53(3): 325-61.

Ross, Michael L. 2004. "Does Taxation Lead to Representation?" *British Journal of Political Science,* 34(2): 229-49.

Ross, Michael L. 2012. *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton: Princeton University Press.

Ross, Michael L. 2015. "What Have We Learned About the Resource Curse?" *Annual Review of Political Science* 18: 239-259.

Schumpeter, Joseph. 1918/1954. "The Crisis of the Tax State." *International Economic Papers* 4: 5-39.

Sivaram, Varun, and Jennifer M. Harris. 2016. "Sustaining Fuel Subsidy Reform." Council on Foreign Relations Discussion Paper. Last modified 18 October 2016, accessed May 22, 2018.

Smith, Benjamin. 2017. "Resource Wealth as Rent Leverage: Rethinking the Oil-Stability Nexus." *Conflict Management and Peace Science* 34(6): 597-617.

Teorell, Jan, Michael Coppedge, Svend-Erik Skaaning, and Staffan I. Lindberg. 2016. "Measuring Electoral Democracy with V-Dem data: Introducing a new Polyarchy Index." *V-Dem Working Paper 2016:25.* Available at http://dx.doi.org/10.2139/ssrn.2740935

Ulfelder, Jay. 2007. "Natural Resource Wealth and the Survival of Autocracy." *Comparative Political Studies* 40(8): 995-1018.

Vadlamannati, Krishna C., and Indra de Soysa. 2016. "Do Rulers of Natural Resource Wealth Welcome Transparency? An Empirical Analysis of Institutional Change, 1980–2012." *International Studies Quarterly* 60(3): 457-74.

Vandewalle, Dirk J. 1998. *Libya Since Independence: Oil and State Building*. London: IB Tauris.

Victor, David. 2009. *The Politics of Fossil-Fuel Subsidies*. Geneva, Switzerland: Global Subsidies Initiative.

Wiens, David, Paul Poast, and William Roberts Clark. 2014. "The Political Resource Curse: An Empirical Re-evaluation." *Political Research Quarterly* 67(4): 783-94.

Wilson, Matthew C., and Joseph Wright. 2017. "Autocratic Legislatures and Expropriation Risk." *British Journal of Political Science,* 47(1): 1-17.

Wright, Joseph, Erica Frantz, and Barbara Geddes. 2015. "Oil and Autocratic Regime Survival." *British Journal of Political Science* 45(2): 287-306.

Young, Alwyn. 2017. "Consistency without Inference: Instrumental Variables in Practical Application." Working Paper.

Zucco, Cesar. 2013. "When Payouts Pay Off: Conditional Cash Transfers and Voting Behavior in Brazil 2002–10." *American Journal of Political Science* 57(4): 810-22.

FUEL SUBSIDIES LIMIT DEMOCRATIZATION: EVIDENCE FROM A GLOBAL SAMPLE, 1990-2014

SUPPLEMENTARY FILE

Table 1A: Main Results Robustness

Regression follows specification in Table 1:	(1) Model 1, but with Model 2 sample	(2) Model 2, but excluding oil income	(3) Model 4, but excluding oil income
Oil income per capita (log)	-0.08***	-	-
	(0.03)		
Fuel subsidies per capita (log)	-	-0.12**	-
		(0.05)	
Fuel subsidies per capita (log) - cm	-	-	-0.17***
			(0.06)
Fuel subsidies per capita (log) - deviation from cm	-	-	0.01
			(0.06)
Mean of DV	-	-	4.15
			(2.87)
Income per capita (log)	0.08	0.02	0.05
	(0.09)	(0.08)	(0.08)
Economic growth	-0.02**	-0.02***	-0.02**
	(0.01)	(0.01)	(0.01)
Polyarchy Index	-1.77***	-1.65***	-1.65***
	(0.35)	(0.37)	(0.42)
Sum of past transitions	0.38***	0.37***	0.28***
	(0.05)	(0.06)	(0.07)
Observations (clusters)	2214 (120)	2381 (125)	2381 (125)

Notes: Probit regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. 'cm' = country mean. *** p<0.01, ** p<0.05, * p<0.1

Table 2A: Linear Probability Model of Increases in Polyarchy Index, 1991-2014, following Table 1 specification

	(4) Standard Probit	(5) Between-Within
Oil income per capita (log)	-0.006**	-
	(0.003)	
Fuel subsidies per capita (log)	-0.009**	-
	(0.005)	
Oil income per capita (log) - cm	-	-0.006*
		(0.003)
Oil income per capita (log) - deviation from cm	-	0.000
		(0.005)
Fuel subsidies per capita (log) - cm	-	-0.016***
		(0.006)
Fuel subsidies per capita (log) - deviation from cm	-	-0.001
		(0.005)
Mean of DV	-	0.628*
		(0.336)
Income per capita (log)	0.011	0.018*
	(0.010)	(0.009)
Economic growth	-0.003**	-0.003**
	(0.001)	(0.001)
Polyarchy Index	-0.197***	-0.200***
	(0.034)	(0.032)
Sum of past transitions	0.048***	0.035***
	(0.006)	(0.008)
Observations (clusters)	2214 (120)	2333 (121)

Notes: Linear probability model regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. 'cm' = country mean. *** p<0.01, ** p<0.05, * p<0.1

Table 3A: Probit results using 3 Point Change in Polity2 as Dependent Variable, 1991-2014

	(6) Standard Probit	(7) Between-Within
Oil income per capita (log)	-0.006	-
	(0.035)	
Fuel subsidies per capita (log)	-0.111*	-
	(0.061)	
Oil income per capita (log) - cm	-	-0.018
		(0.038)
Oil income per capita (log) - deviation from cm	-	-0.011
		(0.074)
Fuel subsidies per capita (log) - cm	-	-0.155*
		(0.081)
Fuel subsidies per capita (log) - deviation from cm	-	0.044
		(0.084)
Mean of DV	-	1.823
		(1.928)
Income per capita (log)	-0.198	-0.100
	(0.121)	(0.115)
Economic growth	-0.034***	-0.036***
	(0.009)	(0.009)
Polity2	-0.157***	-0.160***
	(0.034)	(0.031)
Sum of past transitions	0.430***	0.387***
	(0.081)	(0.083)
Observations (clusters)	1932 (114)	2289 (116)

Notes: Probit regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polity score* increased by at least 3 points from the prior year. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4A: Robustness to using "Price Gap" measure of subsidies

	(8) Standard Probit	(9) Between-Within	(10) Between-Within, Lewbel Internal Instruments for *Fuel Subsidies per capita - country mean*
Oil income per capita (log)	-0.082***	-	-
	(0.031)		
Price gap per gallon	-0.005	-	-
	(0.005)		
Oil income per capita (log) - cm	-	-0.060*	-0.005
		(0.031)	(0.003)
Oil income per capita (log) - deviation from cm	-	-0.034	-0.002
		(0.068)	(0.005)
Price gap per gallon - cm	-	-0.021**	-0.002**
		(0.009)	(0.001)
Price gap per gallon - deviation from cm	-	0.009*	0.001
		(0.005)	(0.001)
Mean of DV	-	4.113	0.593*
		(2.797)	(0.354)
Income per capita (log)	0.110	0.152	0.015
	(0.095)	(0.093)	(0.009)
Economic growth	-0.020**	-0.024***	-0.003*
	(0.009)	(0.009)	(0.001)
Polyarchy Index	-1.823***	-1.850***	-0.200***
	(0.367)	(0.423)	(0.032)
Sum of past transitions	0.385***	0.298***	0.035***
	(0.055)	(0.071)	(0.008)
Observations (clusters)	2221 (120)	2333 (122)	2333 (122)

Notes: Models 1 and 2 report probit regression coefficients with standard errors clustered on countries in parentheses. Model 3 reports linear probability model regression coefficients, following Lewbel (2012) instrumental variable procedure. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. Price gap measure truncated at 0 for negative subsidy (i.e. taxed) values. 'cm' = country mean. *** p<0.01, ** p<0.05, * p<0.1

Table 5A: Tests of Heteroskedasticity in Table 2 IV estimates

	Pagan-Hall test statistic	P-value
Table 2, model 2	258.3	<0.001
Table 2, model 4	284.2	<0.001

Test statistic evaluates H0 that the errors are homoskedastic.
Test statistic calculated by using "Lewbel" internal instruments as first-stage predictors in IV procedure.

Table 6A: Controlling for Total Social Expenditures, Oil Price, and Foreign Reserves

	(11) Standard Probit	(12) Between-Within	(13) Between-Within
Oil income per capita (log)	-0.08	-	-
	(0.07)		
Fuel subsidies per capita (log)	-0.21*	-	-
	(0.12)		
Oil income per capita (log) - cm	-	0.00	-
		(0.09)	
Oil income per capita (log) - deviation from cm	-	-0.24	-
		(0.18)	
Fuel subsidies per capita (log) - cm	-	-0.31*	-0.13*
		(0.17)	(0.07)
Fuel subsidies per capita (log) - deviation from cm	-	-0.08	0.04
		(0.11)	(0.07)
Social expenditures as share of total expenditures	0.00***	0.00***	-
	(0.00)	(0.00)	
Mean of DV	-	-5.78	4.96*
		(8.72)	(2.75)
Oil price	-	-	-0.02*
			(0.01)
Foreign reserves in months of imports	-	-	-0.01
			(0.01)
Observations (clusters)	610 (68)	607 (68)	1,785 (110)

Notes: Models report probit regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes a measure of income per capita, economic growth, past transitions, prior year's *polyarchy index,* year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. 'cm' = country mean. *** p<0.01, ** p<0.05, * p<0.1

Table 7A: Exclude observations with negative yearly Polyarchy Index Change (i.e. “backsliding” observations)

	(14) Between-Within Excluding Backsliding Obs.	(15) Between-Within Exclude Backsliding Obs. IV Estimates using Lewbel (2012) Instruments
Oil income per capita - cm	-0.06*	-0.00
	(0.03)	(0.00)
Oil income per capita - deviation from cm	-0.04	-0.00
	(0.06)	(0.00)
Fuel subsidies per capita (log) - cm	-0.13**	-0.01**
	(0.06)	(0.01)
Fuel subsidies per capita (log) - deviation from cm	0.00	0.00
	(0.06)	(0.00)
Mean of DV	5.08*	0.71**
	(2.67)	(0.35)
Income per capita (log)	0.12	0.02
	(0.09)	(0.01)
Economic growth	-0.02**	-0.00**
	(0.01)	(0.00)
Polyarchy Index	-1.73***	-0.19***
	(0.39)	(0.03)
Sum of past transitions	0.28***	0.03***
	(0.07)	(0.01)
Observations (clusters)	2259 (121)	2259 (121)

Notes: Probit regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. 'cm' = country mean, calculated for the included sample years only. *** p<0.01, ** p<0.05, * p<0.1

Table 8A: Descriptive Statistics

Variable	Obs	Mean	Std.Dev	Min	Max
Polyarchy Index (Ordinal) increase in democracy	2,333	0.06	0.24	0	1
Oil income per capita (log) - country mean	2,333	2.14	2.54	0	10.15
Oil income per capita (log) - deviation from country mean	2,332	0.37	0.99	-6.58	5.61
Fuel subsidies per capita (log) - country mean	2,333	0.95	1.23	0	5.94
Fuel subsidies per capita (log) - deviation from country mean	2,333	0.02	0.85	-3.80	3.39
Country mean of Polyarchy transition variable	2,333	0.06	0.04	0	0.2
Income per capita (log)	2,333	8.37	1.04	5.86	11.79
Economic Growth	2,333	2.32	5.45	-62.21	37.13
Lagged Polyarchy Index (Ordinal)	2,333	0.43	0.26	0	1
Sum of past Polyarchy transitions	2,333	2.23	1.76	0	9
Region	2,333	4.43	3.35	1	11
Spatial lag Polyarchy Index	2,299	0.05	0.11	0	1
Polity 3-pt transition toward democracy alternative DV	2,248	0.04	0.19	0	1
Age	2,333	16.86	14.77	1	54
Age^2	2,333	502.56	718.34	1	2916
Age^3	2,333	18715.56	33656.15	1	157464
Price Gap	2,333	7.31	13.39	0	94.7
Price Gap - country mean	2,333	7.15	9.97	0	51.67
Price Gap - deviation from country mean	2,333	0.15	8.71	-25.88	80.38

PART B – MEASURING FUEL SUBSIDIES PER CAPITA

The key variable of interest to this project is a measure, in per capita terms, of the amount of money that a government spends to subsidize the domestic consumption of gasoline. As noted in the main text, this subsidy takes the form of either direct intervention to reduce the price consumers pay at the pump compared to market rates, or in cases where the ease of oil extraction and refinement would lead to below world market prices without this direct intervention, lost revenue from not selling gasoline to citizens at market rates.

Such a measure is not currently available for a large number of cases and multiple years. The closest available measure comes from a recent IMF report (Clements et al., 2013) which provides data measuring the percent of GDP countries devote to various energy sources, including petroleum. While these measures provide some ability to calculate per-capita energy subsidies, they are only available for a single year. A simple cross-sectional measure based on these IMF data are a useful check on the measurement strategy deployed here, but are ill-suited to empirical analysis of panel data.

To overcome these obstacles, I create my own measure of this concept. This appendix describes the steps taken to create this variable, including a discussion of how missing data were handled, and evaluates the extent to which this measures corresponds to other potential measurement strategies.

PART I: OVERVIEW

I follow existing conventions (e.g. Coady et al. 2010) and utilize the price-gap approach to the measurement of subsidies. This approach estimates the size of subsides by comparing the domestic price of a good to some benchmark price the good would fetch in a competitive market (Kojima and Koplow 2015). When the domestic price is lower than the benchmark price, a subsidy is assumed to exist. Or course, artificially lower prices encourage excess consumption of the particular good in question. Because I am interested in the total fiscal burden that such subsidies impose on regimes, however, I multiply this measure of the per-liter subsidy by the total liters of fuel consumed within a country per year, and finally divide by population. The result is a measure of the fiscal burden of fuel subsidies in per-capita terms.

The resulting variable *Fuel Subsidies Per Capita* requires separate data components. I discuss them individually in the paragraphs that follow.

Benchmark Fuel Price

For internationally traded commodities, such as oil and its various refined products like the gasoline analyzed here, the benchmark price is the international price for a good, with appropriate adjustments for transportation and distribution costs (Coady et al., 2010). Estimating a benchmark price for gasoline is, from a theoretical standpoint, a relatively straightforward affair. Given the high degree of global integration in the petroleum market, and the generally low transportation costs, most scholars studying the oil and gasoline markets view gasoline prices converging toward a single world price, so long as governments do not intervene to either lower the costs paid by consumer (a subsidy) or to increase the price in an effort to raise additional tax revenue to offset the negative externalities of gasoline consumption.

In actuality, of course, governments do regularly intervene to lower (or raise) the price of gasoline such that global prices vary wildly. Scholars must therefore look elsewhere to measure the benchmark price of fuel. I follow the approach outlined by the German Federal Ministry for Economic Cooperation and Development (Deutsche Gesellschaft für International Zusammenarbeit, denoted as 'GIZ' hereafter) and use the price per liter in the United States as the benchmark price. US prices represent a "revenue neutral" international benchmark, meaning that prices cover transportation and distribution costs along with industry margins, though do not include sufficient taxes to offset the environmental or health costs associated with fossil fuel usage (GIZ, various years).

US costs per liter are calculated using data reported by the US Energy Information Administration (EIA). Specifically, I collected data on the average United States cost for a gallon of gasoline in November of each year 1990-2014, and converted this to a price per liter using the EIA conversion guide. I used prices in November of each year to match the GIZ's reporting of fuel prices in other markets, discussed in more detail below.

Data from GIZ are the most widely available source of international fuel price data, constitute the underlying data for existing IMF (e.g. Coady et al., 2010) and World Bank data on global fuel prices, and report data for US fuel price for a number of years. Nevertheless, GIZ surveys are not done annually, so relying on the GIZ source for establishing the benchmark price introduces considerable missing data into the variable construction process.[14] Using data from the US EIA helps mitigate this problem.

This price serves as the benchmark fuel price for all countries except those that are net oil importers. For oil importers, I follow the IMF's (Clements et al., 2013) approach and add 20 cents per liter to cover the additional transportation and distribution costs. I use Ross's (2013) data on net oil exports to distinguish countries that are importers from exporters.

In sum, this price represents the price we should expect to observe if governments did not intervene to lower (or in many cases raise) the price of gasoline relative to a revenue-neutral benchmark.

<u>Domestic Fuel Price</u>

The second component of the price-gap approach to subsidy measurement is a measure of actual fuel prices in a given country in a given year. If the observed price falls below the benchmark price, the difference constitutes the extent of the subsidy.

I use data collected from the IMF (Coady et al., 2010) and supplemented by GIZ (various years) to measure domestic fuel prices across countries for a variety of years. These prices are measured in November of each year and prices are reported in constant US cents per liter. Country coverage is fairly broad, though temporal coverage is less so. The data from the IMF cover fuel prices for the years 2002-2009. Using additional reporting from GIZ, we are also able to obtain measures of domestic fuel prices for the additional years of 1995, 1998, 2000, 2010, 2012, and 2014. Some countries also have prices reported for 1991 and 1993, though country coverage for these years is considerably sparser.

[14] I describe how this affects the rest of the variable construction in the following section.

For all country-years in which fuel prices are reported, calculating the per-liter subsidy is as simple as subtracting the domestic price from the benchmark price. When this difference is positive (meaning that the domestic price is below the benchmark price), the difference is recorded as the per-liter subsidy. When this difference is negative (meaning that the domestic price is higher than the benchmark price), this indicates that gasoline consumption is taxed at levels exceeding the revenue-neutral cost, and therefore no subsidy is present. For purposes of this project, the per-liter subsidy level is set at 0 rather than a negative number. While the price-gap measures are calculated in cents per liter, the empirical analyses using this variable divides this measure by 100, so that the measure becomes denominated in dollars, not cents.

As is common in comparative politics research, however, missing data is a problem when it comes to time-series cross-sectional data. Even in the best case scenario (meaning that a country has non-missing data on fuel prices for all years covered by these two sources), fuel prices are still missing for 7 years out of 20. Of course, in other instances, the proportion of missing data are higher. In order to maximize the temporal coverage of the variable, I take steps to mitigate the extent of missing data.

There are a number of approaches to dealing with missing data, ranging from relatively simple interpolation strategies to multiple imputation strategies.[15] The choice of strategy should depend on a careful examination of the pattern of – and to the extent possible, the explanation of – missing data. In this case, data for fuel prices are missing largely because the agency that reports fuel prices only surveyed countries in certain years. Unlike the "missing at random" pattern that pervades other sources of time series cross sectional data in employed in the field, the vast majority of missing observations can be explained by the timing of the GIZ fuel price surveys, which are produced on an every-other-year cycle starting in 1998.

Given this pattern, I deal with missing data by a combination of simple linear interpolation, the use of moving averages, and relatively straightforward logic. I begin by identifying the countries that are – based on the reported fuel price data – long-term subsidizing countries versus those that are not. The decision rule I followed was based first and foremost on whether or not a country's observed fuel price was more than twenty-percent less than the US-based benchmark price in a given year. If the observed price fell below this threshold, I considered the country-year to be a subsidizer. Countries in which more years were designated as subsidizers than not are considered long-term subsidizers.[16]

The designation of a country as a long-term subsidizer or not is important for my approach to dealing with missing data. Specifically, I deal with the missing data problem differently depending on whether or not a country is a long-term subsidizer. For those that are long-term subsidizers, I fill in missing values by using the moving average of observed gas prices in the year immediately before and after the year in which fuel price data is missing. For instance, if 1994 prices are missing for Venezuela, I assign 1994 the value of the average of the 1993 and 1995 reported prices. If there are no non-missing values immediately adjacent, the value stays missing. For instance, if 1993, 1994, and

[15] I discuss multiple imputation and report robustness analyses in section III.
[16] I use the twenty-percent below US benchmark price threshold because we are interested in distinguishing long-term fuel subsidizers from those whose domestic prices may fall below the US benchmark price by very small amounts or only occasionally.

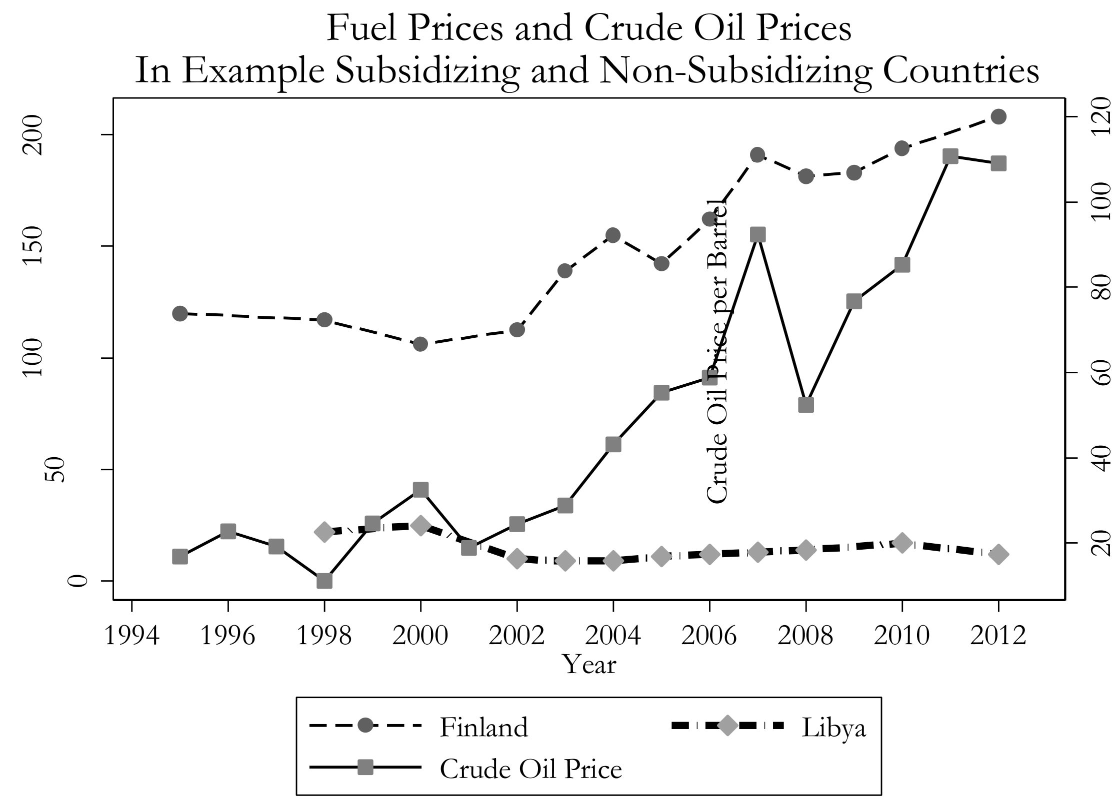
1995 were all missing, than 1994 stays missing. In this way, I maximize the likelihood that the moving average approach covers years that are missing because the GIZ did not survey fuel prices in that year, and not for any other random reason. The logic behind this approach is that countries that are long-term subsidizers of fuel tend to have relatively stable fuel prices that are immune to fluctuations in the underlying price of oil – this price stability, after all, is what subsidies are designed to accomplish. If a country has a history of subsidizing fuel costs, then averaging over missing years is a reasonable way of estimating fuel costs in years that were not covered by the GIZ surveys.

For countries that are not long-term fuel subsidizers, I fill in missing data by using a panel-specific linear interpolation of fuel prices based on the observed price of crude oil in missing years. The logic behind this approach is that for countries that do not intervene to lower the costs of fuel, the main driver of fuel prices is the underlying price of oil – when oil prices increase, fuel prices should increase, and vice versa. When this interpolation process is completed on a country-by-country basis, the process is able to account for country-specific trends in added taxation, such as distinguishing high-tax from relatively low-tax countries when it comes to domestic gasoline consumption. I use data from the EIA to measure the spot price of Brent crude oil in Europe. I use oil prices in November of each year to be consistent with the reporting of fuel price data from the IMF (2013) and GIZ.

Though each of these approaches to missing data are rooted in logic, I can also assess their appropriateness empirically. For instance, while the logic of fuel subsidization suggests that prices should remain relatively constant, this assertion can be tested empirically with the non-missing data reported from the IMF and GIZ. Using the classification scheme reported above, I calculated the average standard deviation of fuel prices for countries considered long-term subsidizers and for countries that are not classified as such. If subsidizers tended to have more stable prices than those countries that do not subsidize fuel, we would expect the average standard deviation of prices in the former category to be considerably lower than in the later. This is exactly the pattern we observe. For long-term subsidizers, the average standard deviation of non-missing fuel prices is 29.9. For countries that are not long-term subsidizers, this figure is 45.7. This result is consistent with the expected behavior of subsidizers versus non-subsidizers, namely that the former should experience considerably less price volatility than the latter. These simple statistics also suggest that the approach to dealing with missing data is also reasonable.

These same patterns can be seen visually in Figure B1. This graph plots the observed (that is, non-missing) fuel prices for a prototypical long-term subsidizer and non-subsidizer alike. The representative countries are Libya and Finland respectively, though the choice makes little difference to the underlying conclusions from such an exercise. Along with the observed fuel prices (in US cents per liter) for these countries, Figure B1 also plots the oil spot price, which is used as the basis for the linear interpolation for missing fuel price values in non-subsidizing countries, though should have little relationship to the price of gasoline in subsidizing countries.

Figure 1B



This is exactly the pattern we observe. In long-term subsidizer Libya, the non-missing fuel prices are very stable and lack any meaningful correlation with the underlying price of oil. As noted above, this should not be surprising – one of the underlying stated purposes of extensive fuel subsidies is in fact to minimize price volatility. On the other hand, fuel prices in non-subsidizing countries such as Finland have a close correlation with the underlying price of oil.

This exercise is not meant to offer a definitive account of variations in fuel subsidies across countries or a full model of how the price of oil and fuel are related across countries. Instead, both of these exercises provide empirical support for the approach I have employed to deal with missing data in terms of observed fuel prices globally.

Total Fuel Consumption

The third data component is a measure of total fuel consumption in a country per year. I use data from the EIA to create this measure.[17] The EIA provides annual measures of the total consumption of gasoline, reported in thousands of barrels per day. I use these measures to first transform the data in annual consumption (by multiplying the daily measure by 365, and adjusting for leap years). This

[17] Available from http://www.eia.gov/beta/international/data

creates a measure of total gasoline consumption annually, in thousands of barrels. I then convert these figures into liters, so the end result is a measure of estimated annual gasoline consumption in liters.

Population

The final data component is a measure of total population. I use the World Bank's World Development Indicators for this measure.

Combining Into "Fuel Subsidies Per Capita" Variable

Once these four data components are collected, I turn to construction of my key variable *Fuel Subsidies Per Capita*. To calculate this variable, I first measure the per-liter fuel subsidy by subtracting the domestic price from the benchmark price adjusted for importer status. For all country-years where the domestic price is lower than the benchmark, I record the differences as the per-liter subsidy. For all country-years in which the domestic price exceeds the benchmark price, I record the per-liter subsidy as 0.

I then multiply the per-liter subsidy by the measure of total consumption of gasoline in liters annually, and divide by population so as to weight the variable in per-capita terms.

The resulting variable ranges from \$0 to \$722 (Bahrain in 2011). As the variable is highly skewed, empirical analysis using this variable employs a log transformation.

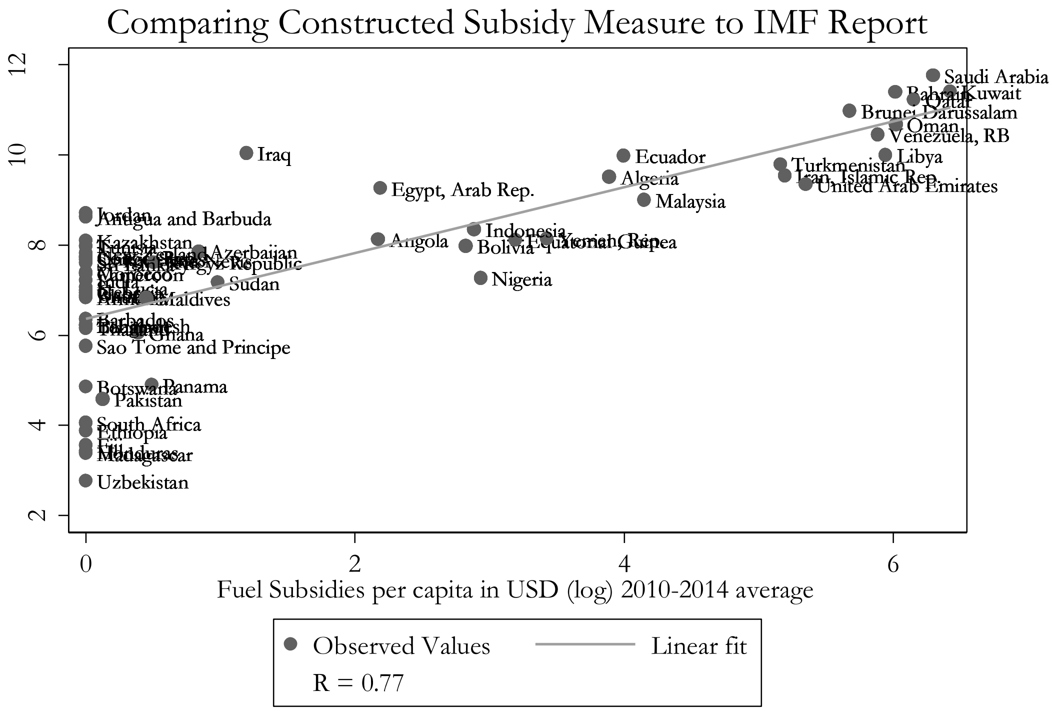
PART II: COMPARISON WITH ALTERNATIVE MEASUREMENT STRATEGIES

As noted at the outset, there are no other existing sources of data measuring fuel subsidies for a wide array of countries and years. The closest alternative source is a report from the IMF (Clements et al., 2013) which provided a single year (2011) measure of the pre-tax subsidies for various energy sources, including petroleum, electricity, natural gas, and coal as a share of GDP. As the subsidy estimates were reported separately for each sector, it is possible to use the reported share of GDP allocated to pre-tax subsidies for petroleum to construct a similar indicator as the one outlined in detail above. Though this measure includes petroleum subsidies more broadly defined than the focus on domestic gasoline consumption I've described above, the measures should at least be broadly comparable. If they are closely related, then we can have greater confidence in the measurement strategy employed here.

To compare the two measures, I take the IMF's reported share of GDP devoted to pre-tax subsidies for petroleum products and multiply this by the IMF's reported GDP estimate in 2011, the same year in which the subsidy measures are reported. This converts the IMF's reported "share of GDP" figures into actual dollar figures. I then divide by the population in 2011 so that the total spending is reported in per-capita terms. As is the case for my own variable construction, the resulting variable is highly skewed so I employ a log transformation for analysis purposes.

Figure 2B below plots the variable constructed from the single-year IMF estimate against my own variable construction, averaged between 2010 and 2014 to maximize comparability. As can be clearly seen, the variables are closely related to one another ($r = 0.77$). The main difference lies on the left side of the graph, where the IMF reports some meaningful petroleum subsidies per capita whereas my own variable construction based on gasoline prices is considerably smaller, and in some cases is recorded as zero. Though the IMF's methodology is not entirely clear, these differences are likely attributed to the IMF's reliance on a broader set of petroleum subsidies, including other forms of fuel such as diesel, than my narrower definition. Nevertheless, the results of this comparison are comforting, since it shows that – despite using a very different approach to measure the size of fuel subsidies – my variable construction is closely related to the IMF's single year snapshot.

Figure 2B



PART III: THE IMPACT OF MISSING DATA ON EMPIRICAL FINDINGS

Finally, this appendix examines the extent to which missing data may affect the empirical findings presented in the main text. Table 1B describes each variable employed in the main and supplemental analysis in terms of the share of non-missing values. This calculation is based on the maximum possible sample size as defined by the following three criteria: country-years between 1990 and 2014, scores on the *Polyarchy Index (ordinal*) of 0.75 or lower, and country-years designated as fully autonomous according to the VDem dataset. The first of these criterion limits the analysis to the years in which it is possible to measure the size of fuel subsidies given data availability. The second limits the sample to only those country-years in which an increase in the *polyarchy index* was possible. Finally, the third criterion removes from the sample countries that may have experienced externally imposed regime changes. In total, this results in 2,900 country-years.

Table 1B: Missing Data Description Variable	Valid Observations	Share Non-Missing
Polyarchy Index (Ordinal) increase in democracy	2,900	100.0
Oil income per capita (log) - country mean	2,713	93.6
Oil income per capita (log) - deviation from country mean	2,662	91.8
Fuel subsidies per capita (log) - country mean (interpolated as described in Supplemental File)	2,723	93.9
Fuel subsidies per capita (log) - deviation from country mean (interpolated as described in Supplemental File)	2,497	86.1
Country mean of Polyarchy transition variable	2,900	100.0
Income per capita (log)	2,650	91.4
Economic Growth	2,665	91.9
Lagged Polyarchy Index (Ordinal)	2,900	100.0
Sum of past Polyarchy transitions	2,900	100.0
Region	2,783	96.0
Polity 3-pt transition toward democracy alternative DV	2,579	88.9
Age	2,900	100.0
Age^2	2,900	100.0
Age^3	2,900	100.0
Price Gap (interpolated as described in Supplemental File)	2,516	86.8
Price Gap - country mean (interpolated as described in Supplemental File)	2,723	93.9
Price Gap - deviation from country mean (interpolated as described in Supplemental File)	2,516	86.8

Note: Table reports information for all variables included throughout the manuscript. Share non-missing based on maximum possible sample size, defined as all country-years from 1990-2014 with a non-missing *polyarchy index (ordinal)* score less than or equal to 0.75 and coded as "fully autonomous" in the Vdem dataset. See text for further details.

There are two main inferential concerns that arise from this pattern of missing data. First, do the patterns of missing data throughout the entire dataset bias the reported estimates? That is, would having complete data for each independent variable – and the resulting increase in sample size – change the estimates from those reported in the main text?

I address this question by employing multiple imputation methods.[18] The multiple imputation datasets use the known values of variables to create five separate imputed datasets. Once created, I use the built-in multiple imputation commands in Stata to re-estimate the baseline model (the so-called Between-Within model reported in Table 1, column 4). The results of the multiple-imputation estimates are reported in Table 2B on the following page. The sample size reduction reflects the lagged variables. In terms of substantive interpretation, the results are identical to those reported throughout the manuscript: the between-effect of higher fuel subsidies continues to exert a large and statistically significant negative impact on the likelihood of a transition toward democracy. The rest of the control variables demonstrate similar patterns of results when compared to the estimates reported in the main paper.

[18] Specifically, the Amelia II program by Honaker, King, and Blackwell (2012).

Table 2B: Multiple Imputation Estimates of Increases in Polyarchy Index, 1991-2014

	(16) Replicates Table 1, Model 4
Oil income per capita - country mean	-0.05*
	(0.03)
Oil income per capita - deviation from country mean	-0.02
	(0.06)
Fuel subsidies per capita (log) - country mean	-0.11**
	(0.05)
Fuel subsidies per capita (log) - deviation from country mean	0.05
	(0.07)
Country mean of DV	2.72
	(2.38)
Income per capita (log)	0.08
	(0.07)
Economic growth	-0.02**
	(0.01)
Polyarchy Index (ordinal)	-1.46***
	(0.27)
Sum of past transitions	0.24***
	(0.06)
Observations / Clusters	2744 (132)

Notes: Column 1 reports probit regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes a year time trend variable, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. *** p<0.01, ** p<0.05, * p<0.1

The second question involves the role of missing data in the calculation of the key variable *fuel subsidies per capita*. As noted above, due to the biennial nature of the GIZ's reporting, there is a much larger degree of missingness in the data measuring domestic fuel prices, which is a key component of the overall *fuel subsidy* variable. Part I of this supplemental file described in detail how I filled in these missing values using a moving average of observed prices for gasoline subsidizing countries or a country-specific linear interpolation based on the time-varying price of crude oil for non-subsidizing countries. As Part I demonstrated, these decisions are rooted in a careful logic and have considerable empirical evidence suggesting that they are valid approaches. Table 1B above shows that this interpolation process for domestic fuel prices results in a final measure of *fuel subsidies per capita* that has roughly the same share of non-missing values as the other variables used throughout the analysis. Nevertheless, I use two different approaches to assess just how consequential the interpolation process is for the results presented in the main analyses.

First, I have re-estimated the main model using a constrained temporal sample that minimizes the potential impact of the interpolation process. As Part I of this supplemental file notes, Coady et al. (2010) report fuel price data for the years 2002-2009 for a sample of 186 countries. Though not every country-year has a corresponding value of domestic gasoline prices, restricting the analysis to these years minimizes the extent that the values of *fuel subsidies per capita* employed in the regression depend on the interpolated values of domestic gasoline prices. The GIZ data source also reports data for the years 2001 and 2010, such that a slightly longer sample (2001 – 2010) can be constructed while continuing to minimize the impact of the interpolation process.

On the next page, Table 3B replicates the main model (Table 1, Model 4) using both of these limited time samples. The results of these re-analyses are identical to those presented in the manuscript. This provides some additional assurance that the main results reported are not dependent on the interpolation process used to construct one part of the *fuel subsidies per capita* variable.

Second, I have used a multiple imputation process to directly deal with the missing values in the measure of domestic gasoline prices. Specifically, I created 5 draws of a dataset containing the following variables: the non-missing domestic gasoline prices, a measure of oil production, population size, and the price of crude oil price. I then averaged these 5 MI estimates of domestic gasoline prices and used these values in an alternative construction of the main *fuel subsides per capita* variable. The net result of this process is an alternative measurement of the *fuel subsidies* variable that uses multiple imputation methods to deal with missing data in its construction as opposed to the moving average and interpolation process described above and employed in the main analyses.

Table 4B replicates the main model with this alternative version of the key variable. The results are substantively identical to those reported above, and thus reinforce the robustness of the main findings and the country-specific interpolation process used in the paper.

Table 3B: Robustness of Main Results to Years with non-missing Fuel Price Data

	(17) Between-Within Years: 2000-2010	(18) Between-Within Years: 2002-2009
Oil income per capita - cm	-0.06	-0.10*
	(0.04)	(0.06)
Oil income per capita - deviation from cm	-0.05	-0.04
	(0.10)	(0.12)
Fuel subsidies per capita (log) - cm	-0.21**	-0.19*
	(0.09)	(0.11)
Fuel subsidies per capita (log) - deviation from cm	-0.02	0.05
	(0.09)	(0.11)
Mean of DV	9.69***	7.95*
	(3.46)	(4.18)
Income per capita (log)	0.14	0.18
	(0.13)	(0.16)
Economic growth	-0.02	-0.02
	(0.01)	(0.02)
Polyarchy Index	-1.86***	-2.29***
	(0.55)	(0.72)
Sum of past transitions	0.11	0.16*
	(0.08)	(0.10)
Observations (clusters)	1,175 (114)	861 (113)

Notes: Probit regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. 'cm' = country mean, calculated for the included sample years only.
*** p<0.01, ** p<0.05, * p<0.1

Table 4B: Robustness of Main Results to Using Multiple Imputation in Construction of Domestic Gas Prices

	(19) Fuel Subsidy Variable created with MI estimates of domestic gas prices
Oil income per capita - cm	-0.08**
	(0.03)
Oil income per capita - deviation from cm	-0.08
	(0.07)
Fuel subsidies per capita (log) - cm	-0.21*
	(0.13)
Fuel subsidies per capita (log) - deviation from cm	-0.00
	(0.04)
Mean of DV	4.41
	(2.81)
Income per capita (log)	0.18*
	(0.10)
Economic growth	-0.01
	(0.01)
Polyarchy Index	-1.90***
	(0.39)
Sum of past transitions	0.29***
	(0.07)
Observations (clusters)	2,495 (123)

Notes: Probit regression coefficients with standard errors clustered on countries in parentheses. Dependent variable equals 1 if the *polyarchy index (ordinal)* score increased by at least one step (0.25) from the prior year. Each model also includes year fixed effects, regional dummy variables, time polynomials measuring years since last DV change (age, age^2, and age^3) and a constant, but these are not reported to save space. 'cm' = country mean, calculated for the included sample years only.
*** p<0.01, ** p<0.05, * p<0.1

References (Appendix B Material Only)

Clements, Benedict, David Coady, Stefania Fabrizio, Sanjeev Gupta, Trevor Alleyne, and Carlos Sdralevich. 2013. *Energy Subsidy Reform: Lessons and Implications*. Washington DC: The International Monetary Fund.

Coady, David, Robert Gillingham, Rolando Ossowski, John Piotrowski, Shamsuddin Tareq, and Justin Tyson. 2010. "Petroleum Product Subsidies: Costly, Inequitable, and Rising." *IMF Staff Position Note 10/05.*

Deutsche Gesellschaft für International Zusammenarbeit ("GIZ"), various years. "International Fuel Prices." Available from https://www.giz.de/expertise/html/4282.html

Honaker, James, and Gary King. (2010). "What to Do about Missing Values in Time-Series Cross-Section Data." *American Journal of Political Science* 54 (no. 2): 561-581.

Kojima, Masami, and Dough Koplow. 2015. "Fossil Fuel Subsidies: Approaches and Valuation." *World Bank Policy Research Working Paper 7720.*