Does Oil Hinder Transparency?*

Hamid Mohtadi, Michael L. Ross, Stefan Ruediger and Uchechukwu Jarrett†

August 2017

Abstract

Evidence suggests that petroleum wealth is associated with less transparency. But the reasons remain unclear. We propose an explanation supported by extensive evidence, using a simple model: Officials trade-off greater transparency to improve tax compliance against less transparency to increase gains from corruption. Oil windfalls diminish tax revenue needs, causing officials to optimize on less transparency and citizens to optimize on less compliance, at equilibrium. Using giant oil discoveries dataset and oil price data, we combine dynamic and static instruments for oil, with geography instruments for per capita income, to find robust support for our explanation and the model's deep structure across three distinct estimation techniques for 130+ countries for 1980-2010.

Keywords: resource curse, transparency, taxation, public goods, oil

JEL codes: H1, P26, O13, O17

*Financial support was provided by the Economic Research Forum (ERF). Prior versions were presented at the Universities of Wisconsin and Minnesota, Pontifica Universidad Catolica de Chile, Santiago, Chile, the Economic Research Forum, Cairo, Egypt, the University of Southern California, Los Angeles, CA, and the 2016 Midwest International Economic Development Conference, Minneapolis, MN. We thank Scott Adams, Niloy Bose, Xi Chen, Scott Drewianka, Ibrahim Elbadawi, Jason Kerwin, Jeffrey Nugent, Hashem Pesaran, Kamiar Mohaddes, Hadi Salehi-Esfehani, Raimundo Soto and three ERF internal reviewers of the project for their helpful comments.

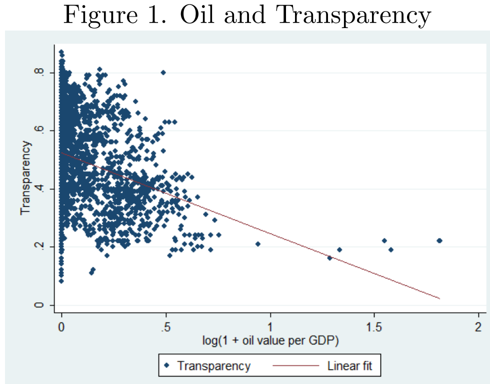
†Hamid Mohtadi is the corresponding author. He is Professor of Economics, University of Wisconsin in Milwaukee and Visiting Professor of Economics, University of Minnesota. Michael Ross is Professor of Political Science, UCLA, and Visiting Professor of Political Science, University of Oxford, UK. Stefan Ruediger is Associate Professor of Economics, Arizona State University. Uchechukwu Jarrett is Assistant Professor of Economics, University of Nebraska.

Does Oil Hinder Transparency?

1 Introduction

Natural resource wealth has been linked to a wide range of adverse economic and political outcomes,[1] particularly when institutional quality is low.[2] But can resource wealth itself influence the quality of the institutions in the first place? We address this question by focusing on government transparency. A number of studies suggest that petroleum wealth is associated with reduced transparency (see below and also Figure 1), meaning fewer public disclosures about government policies, institutions, and activities[3,4]. It is not well-established, however, why resource wealth might reduce transparency, and the conditions under which this relationship is observed.

Figure 1. Oil and Transparency



This paper seeks to bring greater clarity to this topic using the framwork of a simple game, accompanied by extensive empirical evidence, that examines the link between oil revenues and government transparency via taxation. In it public officials trade-off greater transparency to improve tax compliance against less transparency to increase gains from corruption. Oil windfalls diminish tax revenue needs, causing officials to optimize on less transparency and citizens to optimize on less compliance, at equilibrium.

Our model is simple, but our explanation is new and robust to variations across three different estimation techniques using several rich datasets. We exploit plausibly exogenous cross country variations on new giant oil field discoveries (e.g. Arezki, Ramey, Sheng 2017) complied by Horn (2014) along with exogenous temporal variations from oil price data to instrument for oil revenues, using an instrumental variables method that allows us to combine static and dynamic instruments. We also use geographic variables (latitude and world regions) to instrument for income per capita. We utilize two high quality datasets, one on taxation that is meticulously complied by the International Center for Tax and Development (ICTD, 2014) and one

[1] For summaries of this voluminous literature, see Ross (1999), Stevens and Dietsche (2008), Wick and Bulte (2009), Frankel (2010) and Ross (2015).

[2] Studies indicating that the impact of natural resources on income depends on the quality of institutions are found in Lane and Tornell (1996), Tornell and Lane (1999), Mehlum, Moene and Torvik (2006), Robinson, Torvik and Verdier (2006), and Boschini, Pettersson and Roine (2007).

[3] A prevailing definition of transparency used by the International Monetary Fund (2012) is: "the clarity, reliability, frequency, timeliness, and relevance of public reporting and the openness to the public of the government's fiscal policy-making process."

[4] Several international initiatives are based on the belief that this outcome is not inevitable. They include the Extractive Industries Transparency Initiative, the IMF's Fiscal Transparency Code, the UNDP's Strategy for Supporting Sustainable and Equitable Management of the Extractive Industries, and the mandatory disclosure laws adopted by the US and EU for extractive industry firms operating abroad.

on oil values (Ross, 2013) both of which remedy shortcomings of other alternatives. These datasets and their quality are detailed in the Data section of the paper. (For robustness we also use other data). We test the deep structure of the model and its implications for the period 1980-2010 and for 130-150 countries by employing three distinct estimation techniques; Instrumental Variables (IV) estimation to address the various endogeneity issues using the above instruments, Structural Equation Model (SEM) to capture the interrelationships among the variables as implied by the model, and Cross-Sectionally Augmented Autoregressive Distributive Lag (CS-ARDL) to address the dynamic dimensions of the relationships and sort out the long-term from the short-term effects. All three methods strongly support the model and its implications.

This paper is related to several strands of the literature. First, although this is the first paper that examins and finds a connection between oil wealth, government transparency and citizen tax compliance, its key channel, i.e., the link between tax compliance and government transparency, is inspired by the taxation-representation arguments in which the government's need for direct taxation induced greater political accountability both in the nineteenth century Europe (North 1990, North and Weingast 1989, Hoffman and Norberg 1994) and in contemporary societies (Ross 2004, Brautigam et al. 2008, McGuirk 2010). Viewed in this perspective, the contribution of the paper is how the presence of oil revenues may have compromised this historical social contract.

This paper is also related to the existing evidence which suggests a negative association between oil or natural resource wealth, with both transparency (Ross 2011; Williams 2011; Mohtadi, Ross and Ruediger 2014)[5] and also taxation (C´ardenas et al., 2011; Jensen, 2011; Crivelli and Gupta 2014).[6] Egorov, Guriev, and Sonin (2009) also find that resource windfalls reduce freedom of information. However, while their model focuses on the threat of revolt, we offer a different explanation with a more direct relevance to policy. Our measure of transparency also covers a much longer period (1980-2010); our results are robust to variations in estimation techniques; and our empirics are distinguished by their test of the structure of the model (as opposed to reduced forms) and their use of several novel instruments.

The relation of our paper to the state capacity literature is more nuanced. While the state's ability to collect taxes can be an indicator of its capacity as in Besley and Persson (2009, 2010), citizens tax compliance in our paper may respond either to better enforcement or to better government transparency (or both). Although we do allow for both channels in the model, the empirical results tend to support the transparency channel as the key strategic variable.

The notion that oil revenues reduce government accountability and by extension citizen input has a potential policy implication. For example, Devarajan et. al. (2011) develop a model in which resource revenues, if transferred directly to citizens, and then taxed to finance public expenditures, would increase efficiency. The argument is that public spending, when based on taxation rather than resource revenue is more likely to encourage citizen scrutiny and thus efficiency.

Since the lack of transparency is likely to promote corruption, this paper is also related to the literature on government corruption and tax evasion. For example, Tanzi and Zee (2000) link citizens' tax evasion to

[5] Williams (2011) finds a correlation between resource wealth and reduced transparency in a smaller panel of countries over a briefer span of time. Ross (2011) reports a cross-national correlation between oil wealth and lower scores on the Open Budget Index, but only among authoritarian states. Mohtadi, Ross and Ruediger (2014) estimate the relationship between petroleum wealth and transparency, finding that oil is associated with less transparency when it constitutes a significant share of national income (representing a country's dependence on oil), but not when it merely contributes to a higher average wealth of the country (an indicator of oil abundance). Consistent with this finding, the present paper also uses oil as a fraction of national income both in its model and in its empirical formulation.

[6] As anecdotal evidence of the latter, Juan Pardinas, the General Director of the Mexican Institute for Competitiveness stated in 2013 that, "We collect few taxes because we have oil. . . That allows us to pay less in taxes..." (New York Times, September 9, 2013).

their perception of government corruption. Mohtadi and Polasky (2016) develop a model in which corrupt governments, facing citizen tax non-compliance, choose trade protection in order to rely on trade taxes and tariffs, rather than on income tax, to finance government budget as well as corrupt activities.

The empirical findings support the structure of the model in which oil wealth reduces the government's optimal transparency level and tax payers respond by optimally reducing their tax compliance. In an appendix, we also examine an alternative mechanism based on tax enforcement, using an available proxy.[7] What we find is that although tax compliance also responds to more strict enforcement besides more transparency, it is *transparency* that turns out to be the key strategic policy variable of choice by governments.

In the remainder of this paper, section 2 presents the model; section 3 presents the precise linkages between the model and the empirical methodology by mapping theoretical propositions to empirically testable hypotheses; section 4 discusses the data and choice of the instruments; section 5 discusses the estimation strategies, econometric model, data, instrument choices and the results; section 6 offers concluding remarks.

2 The Model

To capture the key ideas, a simple stylized model of the economy is developed, with government and citizens as key players in a game in which the government moves first and citizens follow (described later in more detail). The economy produces a single final good with labor (l), capital (k) and a non-excludable public input provided by the government (e.g., infrastructure) G:

$$Y = A(k^{\alpha} l^{1-\alpha}) G^{\beta} \quad \alpha, \beta \in (0,1) \tag{1}$$

where, A is the usual scale factor representing the level of technology. Demand for labor and capital are given by their marginal products, derived from maximization of profits, $\pi_Y = A(k^{\alpha} l^{1-\alpha}) G^{\beta} - wl - rk$, where Y is the numeraire good with price of 1. Labor is fully employed with measure 1, so that $l^* = 1$ where l^* is demand for labor. We assume that capital markets are open so that return to capital, r, is exogenous, given by the international capital markets: $r = \bar{r}$. Profit maximization then yields aggregate output Y and labor demand, given G. G is exogenous to the agents but subject to balanced budget equation (below). Thus total output, labor, and capital income are,

$$Y = \Omega G^{\frac{\beta}{1-\alpha}} \tag{2}$$

$$Y_l = w^* l = w^* = (1-\alpha)\Omega G^{\frac{\beta}{1-\alpha}} \tag{3}$$

$$Y_k = \bar{r} k^* = \alpha \Omega G^{\frac{\beta}{1-\alpha}} \tag{4}$$

where $\Omega \equiv A^{\frac{1}{1-\alpha}} (\frac{\alpha}{\bar{r}})^{\frac{\alpha}{1-\alpha}}$ which is a function of the constants of the economy.

[7] This result is subject to the limitation that the best available proxy is rule of law. While this rule of law includes enforcement at large, it does not specifically focus on the enforcement of tax laws.

2.1 Public Goods and Government Budget

Public goods G are financed by both natural resource revenue R and income tax revenue X. (We reserve the notation T for Transparency introduced later.) Income is taxed at the rate t, but is subject to citizen under-reporting, the extent of which is to be determined. For simplicity, but no loss of generality, we assume that capital and labor incomes are taxed at the same rate and are subject to the same level of under-reporting. This is represented by a fraction, $1-\mu$, where μ is the *reported* fraction of total income. This simplification allows us to aggregate both types of income so that total reported income, Y_r, is proportional to aggregate income:

$$Y_r = \mu Y \tag{5}$$

Resource revenue is also expressed as a fraction ρ of total income, Y as:

$$R = \rho Y \tag{6}$$

We allow for "leakage" of some revenue to the outside of the transparent system. This is represented by a fraction b for total government revenue and represents the degree of extractive (corrupt) behavior by government officials. With this background, government spending on public goods is given by,

$$G = (1-b)(R+X) = (1-b)(\rho + t\mu)Y \tag{7}$$

where $X = t\mu Y$ is aggregate tax revenue. This simple result is really an accounting identity. Still, it is significant and will be put to an empirical test. For this reason, we formalize the result here as the first hypothesis of the paper:

Proposition 1. For a given level of public spending to GDP, more oil revenue per GDP means less income tax revenue per GDP.

This proposition, and the future empirical test of its validity, represent what amounts to a reduced form relationship. That relationship will then form the basis for the model's structural relationships that will be developed shortly below and empirically tested.

Substituting for Y from equation 2 into 7 yields, $G = (1-b)(\rho + t\mu)\Omega G^{\frac{\beta}{1-\alpha}}$. Solving for G, we find:

$$G = [(1-b)(\rho + t\mu)\Omega]^{\frac{1-\alpha}{1-\alpha-\beta}} \tag{8}$$

2.2 Tax Evasion/Compliance Technology

Citizens decide what fraction of income to underreport. Although our main focus in this paper is on transparency, it is natural to expect that tax compliance should positively respond, not only to transparency, but also to tax enforcement. In addition, one might conjecture that both variables should also enter the government's decision making process. Thus we will formulate the model with this possibility in mind such that both variables are allowed in, and conduct our empirical test for both channels (to the extent that the data allows).[8]

First, we show how rational tax payers comply with, or evade, taxation in response to both transparency and tax enforcement policies of the government. Let C_u denote the cost/penalty of non-compliance, measured by under-reporting a fraction $1-\mu$ of income. We assume that, tax payers who are caught for underreporting (see below for the probability of being caught), face a convex cost for tax evasion $1-\mu$, i.e., they receive progressively higher penalty for higher rates of tax evasion. This assumption will be needed to produce an interior solution. For analytical tractability, we assume that a quadratic form reflects this convexity. We also assume that C_u is proportional to the level of transparency which we denote by T: higher T raises the cost of tax evasion because a more transparent government is likely less likely to be corrupt and divert tax revenues for personal use, so tax payers understand the higher cost of non-compliance in terms of the forgone efficiency of services they would otherwise receive. This dependence is shown by a function $\gamma(T)$ where $\gamma' > 0$ (see later for empirical test of this assumption).

$$C_u(1-\mu) = \frac{\gamma(T)}{2}(1-\mu)^2 \text{ with } \gamma^{'} > 0 \tag{9}$$

Let π_b^e denote the perceived probability of being caught and penalized for under-reporting. Without loss of generality we assume that, if caught, the individual is always penalized, thus ignoring the possibility of bribing the functionary to avoid penalty. While it is easy to incorporate this aspect, for example by considering μ to be *net* of the bribe rate, the paper's focus is on sovereign or grand corruption, rather than bribery and functionary corruption as in Polinsky and Shavell (2001). The distinguishing feature is that high officials who divert resources for personal use are also the ones who make policy decisions and this is what we are interested in.

In light of the above, a risk averse taxpayer facing convex costs considers $E[C_u(1-\mu)]$, i.e., the expected penalty for tax evasion per unit of income, in her calculations[9]. This is given by:

$$E[C_u(1-\mu)] = \frac{\gamma(T)}{2}\pi_b^e(1-\mu)^2 \tag{10}$$

Given G, citizens choose their compliance ratio by balancing the benefits and the costs of non-compliance. They do so by maximizing their expected incomes from labor and capital. Since we have assumed, for simplicity, similar tax evasion behavior from both sources, this amounts to maximizing expected *total* income as follows:

$$\underset{\{\mu\}}{Max}[E(Y)|_{\bar{G}}] = \underset{\{\mu\}}{Max}[(1-\mu t) - E(C_u(1-\mu)].Y|_{\bar{G}} \tag{11}$$

Substituting from 10 into 11, citizens' best response to government transparency is given by,[10]

$$\mu^*(t,T,L) = 1 - \frac{t}{\gamma(T)\pi_b^e} \tag{12}$$

[8] What we will learn from the evidence is that while the level of tax compliance responds to both transparency and the enforcement of tax laws (or its empirical proxy), it is only transparency that turns out to be the "strategic variable" in the sense that the government optimizes over how much information to allow and citizens respond by deciding how much to report, in a game theory setting. However, we will structure the model so that it is neutral with respect to both variables (and thus falsifiable) in this respect.

[9] This is because, $E[C_u(1-\mu)] = \frac{\gamma}{2}\pi_b^e(1-\mu)^2 > C_u[E(1-\mu)] = \frac{\gamma}{2}[\pi_b^e(1-\mu)]^2$

[10] First order condition is, $-t + \gamma(T)\pi_b^e(1-\mu) = 0$. Second order condition is, $-\gamma(T)\pi_b^e < 0$.

Proposition 2. Citizens choose their tax compliance level optimally

Citizens estimate π_b^e from experience and casual observation. In the aggregate, (a) expectations are realized so that $\pi_b^e = \pi_b$, and (b) the probability of being caught increases with greater enforcement of the tax laws, denoted by L, so that $\pi_b^e = \pi_b = f(L)$ with $f^{'} > 0$ and $f < 1$. Thus, optimum compliance becomes[11]:

$$\mu^*(t, T, L) = 1 - \frac{t}{\gamma(T)f(L)} \text{ with } \gamma' > 0,\ f^{'} > 0 \tag{13}$$

Note that γ' and $f^{'} > 0$ means that tax compliance increases with better transparency and tax enforcement:

$$\frac{\partial \mu^*}{\partial T} > 0, \frac{\partial \mu^*}{\partial L} > 0, \frac{\partial \mu^*}{\partial t} < 0 \tag{14}$$

Proposition 3a. Optimal tax compliance increases in transparency T.
Proposition 3b. Optimal tax compliance increases in tax enforcement L.

In addition, if $\gamma^{"} < 0$ and $f^{"} < 0$, then tax compliance is concave in T and L[12]. In turns out that this concavity is sufficient for the government (as the first mover) to optimize in T and possibly L. This is discussed in the next section. Both the increasing and the concavity assumptions will be empirically tested later.[13]

2.3 Government

Government moves first by anticipating citizens' response in setting the level of transparency, T, and possibly enforcement of tax laws, L. However, we assume that G is determined by the sociopolitical process and *drives* the need for taxation via balanced budget equation 8, i.e., politics driving spending and spending driving taxes. This assumption is in line with the political economy literature where either the interest groups lobby the government for G (e.g., Mohtadi and Roe, 1998, 2003) or voters choose the types of government associated with different levels of G (e.g., Persson, et al., 2007) but not the optimal spending/taxation literature (e.g., Barro, 1990) where a welfare maximizing government chooses the efficient level of G. As long as citizens' voting/lobbying behavior is independent of their tax evasion behavior, this approach is valid so that endogenizing the political process is unrelated to the task at hand and can be avoided. In the empirical section, employing an instrumental variable (IV) approach, we will examine whether G is endogenous. For example anticipating higher tax revenues may drive spending decisions, or oil revenues may drive the size of G (Ross, 2012) as oil producers attempt to "buy" constituent loyalty. In addition, democracies and non-democracies may have distinct political economic structures which may have systematic

[11] In the equation below, since μ^* is a fraction, f and γ have lower bounds, i.e., $\gamma(T = 0) = \underline{\gamma} > 0$ and $f(L = 0) = \underline{f} > 0$

[12] We show this in the case of T : $\mu^*_{TT} = (t/f).(\gamma^{"}\gamma^{-2} - 2\gamma^2\gamma^{-3}) < 0$, if $\gamma^{"} < 0$.

[13] We do not have systematic data on t to test the sign of $\frac{\partial \mu^*}{\partial t}$. But the literature suggests that compliance does decline with the tax rate (e.g., Friedman, et. al. 2000). As we will see, at equilibrium, t itself will depend on T, L, and G.

effects on government spending. We will control for this effect by including a polity variable. We will show that the main findings of this paper remain robust despite all these considerations.

Using citizens best response, i.e., equation 13, in equation 8, the level government spending is found to be:

$$G = (1-b)[\rho + t\mu^{*}(t,T,L)\Omega]^{\frac{1-\alpha}{1-\alpha-\beta}} \tag{15}$$

From this equation, it is seen that the tax rate t is the solution to this simple budget equation and therefore depends on G (assumed exogenous per above discussion)[14], given T and L.

$$t = h(G|T,L,b) \tag{16}$$

But both T and L are assumed to be optimally chosen by the government. The government in making its optimization decision will need to take account of the implicit dependence of t on T and L via the balance budget equation 15, as is shown in 16. Therefore, before the government's optimization decision is discussed, this relationship must be analyzed and understood. Focusing on T first, we make the reasonable assumption that extra-budgetary diversions b are easier to hide when there is some opacity, i.e., $\partial b/\partial T \leq 0$. This leads to the following lemma.

Lemma 1. Under reasonable upper bound on the tax rate, t, $\partial b/\partial T \leq 0$ implies that $\partial t/\partial T < 0$

Proof: See appendix 1.

In reality the fact that more democratic societies which are also more transparent may also experience higher tax rates is because such societies may also spend a higher fraction of their GDP on public spending (G/Y) which we hold exogenous here. In the empirical section this issue is addressed by controlling both for the level of democracy and the variations in G/Y.

The application of Lemma 1 to the case of enforcement of the tax laws is somewhat more complex. On one hand, corrupt officials might be interested in extracting more taxes from the citizen to generate more revenues in order to divert more rents in which case $\partial b/\partial L > 0$, on the other, better tax collection is a part of better rule of law institutions in which case there is likely to be less corruption so that $\partial b/\partial L < 0$. This leads to the second lemma:

Lemma 2. Under reasonable upper bound on the tax rate, t, $\partial b/\partial L \leq 0$ implies that $\partial t/\partial L < 0$ while $\partial b/\partial L > 0$ implies that $\partial t/\partial L \gtreqless 0$.

Proof: See appendix 2.

We will take advantage of both lemmas in determining the government's optimum transparency and tax enforcement policy. Let the government payoff depend on how much "it" diverts for personal use, b. This yields the following utility function:

[14] Holding L and T constant, it is easy to show that $t_G > 0$, consistent with the idea that goverment spending drives the tax rate. To see this, totally differentiate 15 to find,

$$t_G(\mu + t\mu_G) = \frac{1-\alpha-\beta}{\Omega(1-b)(1-\alpha)}(\rho + t\mu\Omega)^{-\frac{\beta}{1-\alpha-\beta}}$$

Since right side >0, $t_G > 0$ if $\mu + t\mu_G > 0$. The latter holds if there is an upper bound on the tax rate: $t < \gamma(T)f(L)/2$. This is the same condition as we will need for Lemmas 1 and 2 below.

$$U_G(T,L) = \bar{V}_{oG} + b(T,L).[\rho + t\mu^*(t,T,L)]Y =$$
$$\bar{V}_{oG} + b(T,L).[\rho + t\mu^*(t,T,L)]\Omega G^{\frac{\beta}{1-\alpha}} \text{ with } b_T \leq 0 \text{ and } b_L \lesseqqgtr 0 \quad (17)$$

The second equality in 17 comes from substituting for Y in terms of G from equation 2. The notion that G is determined exogenously (by sociopolitical processes) and drives the tax rate allows us to empirically identify compliance in one of our estimation methods where direct compliance data, when combined with other variables, yields too few data points. Nevertheless that exogeneity assumption will be put to an empirical test in a robustness section.

In equation 17, $\bar{V}_{oG}$ represents some base level government utility independent of L and T . In a democracy $\bar{V}_{oG}$ is determined by the median voters and b is small. Then variations in $U_G(T)$ come mostly from $\bar{V}_{oG}$. In the empirical sections, we control for the variations in $\bar{V}_{oG}$ by including a democracy variable.

The government seeks to optimize over T and L. First we focus on T. The idea that the government would optimize on transparency is found in monetary and public policy discussions based on both the benefits and the costs of transparency (e.g., Conrad and Heinemann, 2008), but to our knowledge it's application to the case where a corrupt government optimizes on transparency based on tax payer behavior is not seen in the literature. In this context, the transparency decision involves clear trade-offs; on one hand, $\partial b/\partial T < 0$ as already discussed; but on the other hand, the fact that citizen tax compliance increases with transparency, i.e. $\partial\mu^*/\partial T > 0$ (see equation 14), gives the government a larger income base from which to extract and divert resources. In this case we expect an interior solution for optimum transparency:

$$T^* = \arg\max_{\{T\}} U_G(T,.)|_G \quad (18)$$

The first order condition for T^* demonstrates the balance of costs and benefits to the government, from a transparency policy:

$$T^*|_G : b(t\mu_T^* + t\mu_t^* t_T) = -b_T(\rho + t\mu^*) - bt_T\mu^* \quad (19)$$

The left side of this equation shows the marginal payoff of greater transparency to the rent seeking government; via the expansion of the tax base from better compliance by proposition 3a ($t\mu_T^* > 0$), *and* the compliance response to lower tax rate via Lemma 1 ($t\mu_t^* t_T > 0$); the right hand side is the marginal cost of transparency via lower rent seeking ($-b_T > 0$) *and* the marginal loss due to the reduction in the tax rate ($t_T\mu^* < 0$). Note that even though G is exogenous, it still influences government decision via balanced budget equations 15 and 16 coming from the effect of transparency (T) on the tax rate (t).

Using the condition $b_T < 0$ and lemma 1, appendix 3 shows that the second order condition for optimal T^* holds if tax compliance is concave in T, that is if $\partial^2\mu/\partial T^2 < 0$. In short, governments limit transparency to their citizens. Their choice of this limit depends on how tax payers behave and, soon we will learn, also on the available natural resource revenue. The empirical section examines the validity of this result which we summarize as follows:

Proposition 4a. If $b_T < 0$, *lemma 1 (upper bound on tax rate) holds,* and *tax compliance increases with transparency* ($\partial\mu/\partial T > 0$–*proposition 3a) at a decreasing rate* ($\partial^2\mu/\partial T^2 < 0$), *then an optimal transparency policy* T^* *exists.*

Notice that μ is not being maximized with respect to T here. Proposition 4a only requires that $\mu(T)$ be concave in T for a value of T^* to exist (i.e. for the government as the Stackelberg leader to choose a T^*). This is because neither the tax payers nor the government ever maximize μ in T; the former maximize their expected utility in μ; the latter maximizes its utility in T. In this setting, concavity of $\mu(T)$ is all we need for the Stackelberg game to go through.

We now turn to the enforcement of tax laws, $L^* = \arg\max U_G(L,.)|_G$ as an alternative mechanism by which a government may improve tax compliance. The first order conditions for optimum enforcement, L^*, is identical to equation 19 with L replacing T. But now the sign of b_L plays a role. If $b_L < 0$, the result is identical to the above and an optimum level of enforcement of taxes will exist. But if $b_L > 0$ it is possible that $t_L > 0$ by Lemma 2. In such a case, no interior solution and thus no optimum enforcement will exist (see Appendix 4). This outcome arises because there will be no trade-off when stronger tax enforcement leads to more rent extraction by officials *and* also increased tax revenue coming from better compliance. We can summarize this result as follows:

Proposition 4b. If $b_L > 0$ so that *enforcing taxes exacerbates the government's rentier behavior, there may be no optimal tax enforcement. If* $b_L < 0$ so that *tax enforcement counters government's rentier behavior,* and *tax compliance increases with enforcement ($\partial\mu/\partial L > 0$–proposition 3b) at a decreasing rate ($\partial^2\mu/\partial L^2 < 0$), then an optimal tax enforcement policy L^* exists.*

In the empirical section, we will test for both propositions 4a and 4b and find support for the optimal transparency policy but not for the optimal tax enforcement policy.

2.4 Equilibrium

Citizens react to the strategic variables, T and possibly L, chosen by the government. But the tax rate, t, is also a variable entering into the government decision (see for example the use of Lemmas 1 and 2 in propositions 4a and 4b). This then means that *at equilibrium t* must enter the tax payer's post optimization value of μ^*. Focusing on T first, this means:

$$\frac{d\mu^*}{dT}|_{T=T^*(eq)} = \mu^*_{T|T=T^*(re)} + \mu^*_t t_T|_{gov.budget} > 0 \tag{20}$$

where $\mu^*_{T|T=T^*(re)}$ is the slope of citizens' optimal response at T^*, denoted by subscript "re". This is the direct effect and is positive. However, since an increase in transparency also reduces the tax rate (second term on right), it follows that an increase in tax payer compliance at equilibrium (denoted by subscript *"eq"* on the left) may be either due to citizen's greater participation in a more transparent economy or a reduction in the tax rate.

The second differential, via the transparency and the taxation paths, is:

$$\frac{d^2\mu^*}{dT^2} = \mu^*_{TT} + 2\mu^*_{Tt} < 0 if \mu^*_{TT} < 0 \tag{21}$$

Empirically, systematic observations on the tax rate t are not available. Thus, we can only estimate the total effects $d\mu/dT$ and $d^2\mu/dT^2$. A positive total differential and a negative total second differential are consistent with rational tax payer's optimizing compliance. Similar argument holds when tax enforcement, L^*, is the variable of interest.

2.5 Resource Rents and Transparency

We are interested here in how rents from natural resources impact the government's optimal transparency policy. To see this re-write the first order condition in equation 19 as $\Psi(T^*, \rho) = 0$ (suppressing the dependence on L). Collecting all the terms in 19 to the right, we see that $\Psi_\rho(T, \rho) < 0$. Similarly, the second order condition established in appendix 3 can be written as $\Psi_T(T, \rho)|_{T^*} < 0$. Totally differentiating the first order condition $\Psi(T^*, \rho) = 0$ in ρ we have,

$$\Psi_T(T, \rho)|_{T^*} \cdot \frac{dT}{d\rho} + \Psi_\rho(T, \rho)|_{T^*} = 0 \tag{22}$$

from which we obtain,

$$\frac{dT^*}{d\rho} = -\frac{\Psi_\rho(T, \rho)|_{T^*}}{\Psi_T(T, \rho)|_{T^*}} = -\frac{(-)}{(-)} < 0 \tag{23}$$

That is, conditional on its existence, optimum transparency is negatively associated with natural resource revenues.

Proposition 5a: More resource revenue as a share of national income means a lower level of transparency will be optimally chosen by the government.

Similar Result holds for tax enforcement:

Proposition 5b: Conditional on the existence of an optimal tax enforcement policy, more resource revenue as a share of national income means a lower level of tax enforcement will be optimally chosen by the government.

3 From Model to Hypothesis Testing

Before we examine the theory, we must be able to map the propositions of the model into empirically testable hypotheses. Beginning with proposition 1 and noting that this was a simple proposition regarding the government's budget accounting, we have the following hypothesis:

H1: For a given level of public spending, more resource revenue means less income tax revenue (all relative to GDP)

It may be noted that H1 is a "reduced form" hypothesis because its underlying proposition describes a simple reduced form relationship. Other propositions connect different aspects of the model with each other and can be described as "structural". For example, propositions 4a and 4b link the tax payers' optimal compliance behavior, described in propositions 3a and 3b, to the government's optimal transparency or tax enforcement policy as a necessary condition. They also relate the curvature of this compliance behavior to the existence of an optimal government policy. Therefore, combining the two sets of propositions in pairs we will have the following pair of hypotheses that, if confirmed, not only test for tax payer behavior, but also for the *existence* of an optimizing government with regard to T and L:

H2a: Tax compliance increases concavely in transparency (implying optimal transparency policy exists).
H2b: Tax compliance increases concavely in enforcement (implying optimal enforcement policy exists).

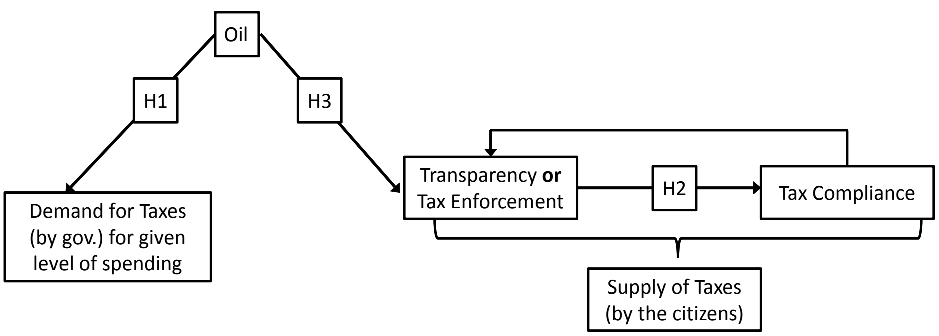
Given the possibility that $b_L > 0$ or $b_L < 0$ from the earlier theoretical discussion, H2b requires some elaboration. If concavity is rejected this indicates that there is no optimality in enforcement. Theoretically, this implies corrupt governments might prefer to enforce tax rules more vigorously to get more taxes (i.e., $b_L > 0$), as we described in proposition 4b and its preceding paragraph. In this case, tax enforcement behaves differently from transparency: instead of a trade-off between corruption and tax enforcement, one would reinforce the other, hence the reason why tax enforcement optimality would be rejected. The empirical section will examine these hypotheses.[15]

The final hypothesis is related to resource revenue. The optimum level of transparency and tax enforcement discussed are both for a *given* level of resource revenue. An increase in this revenue source acts to reduce the optimum from propositions 5a and 5b. Mapping these propositions we have:

H3a: In relation to H2a, Transparency falls with resource revenue as a share of national income.
H3b: In relation to H2b, Tax enforcement falls with resource revenue as a share of national income.

Figure 2 below presents a conceptualization of these hypotheses and the relationship among them. For example, one can view hypothesis H1 as signifying the "demand side" for tax revenues (by the government) while hypotheses H2 and H3 as signifying the "supply side" (by the citizens).

Figure 2. Hypotheses and Model Structure



4 Data

We employ several datasets. For transparency we use the Release of Information (RI) index, developed by Williams (2009) and updated on his website to 2010, which counts the frequency and extent of annual data released by governments and published in World Bank's World Development Indicators (WDI) and the IMF's International Financial Statistics. Its main advantage is its annual availability from 1970 to 2010, covering every country listed in WDI. However, the common time period between this dataset and a second dataset on taxation and government revenue (below), limits us to the 1980-2010 period. Recently Hollyer et al. (2014) have developed a similar dataset based on a Bayesian methodology. However, their transparency measure contains 1500 fewer observations than Williams. Moreover, the correlation between the two datasets is very high (82 percent). Thus, we opted for the dataset by Williams. While simply counting the amount of information is not an accurate measure of the quality of information released by governments, Williams (2011) shows that the RI index is highly correlated with World Bank's Statistical Capacity index and thus

[15] In the absence of direct data on tax enforcement, testing hypothesis H2b is limited by the use of a proxy variable.

also casts light on the quality of information released by governments. This dataset allows us to use panel estimation techniques and hence examine within-country effects over time. The index ranges from -10 to +10. which we normalize to (0,1) to maintain ordinal ranking when we have to square it.

For tax compliance while there exists some direct tax compliance data from the Fraser-Chain Linked index (Fraser Institute Annual Report 2010), which we *will* use for our robustness checks, this data is only available for a limited period, i.e., from 2004 to 2010. On the other hand, high quality data is available on government taxes and revenues from the International Center for Tax and Development (ICTD, 2014) from 1980 to 2010: A key challenge to cross-country research on the role of revenue and taxation has been the weakness of available data. The new dataset released by ICTD addresses this problem and significantly improves on existing data in quality, standardization and coverage.[16] This is therefore our preferred dataset to use for its quality and length of coverage. Of the many tax revenue variables available from ICTD, we choose revenue from *personal income taxes* because this variable matches the theory most closely. This variable is the actual variable of interest when testing the direct effect of oil on taxation, i.e., hypothesis H1. We can also use this variable as a proxy for personal income tax compliance given several controls coming from the model.[17] In addition, we also check the robustness of our results with the Fraser-Chain Linked compliance data which is available, where possible. Figure A5-1 of Appendix 5 shows a remarkably strong relation between tax revenue and the Fraser-Chain tax compliance data for 2004-2010 when the latter data which is available.[18]

The fourth dataset is on oil wealth. Data on resource rents including oil are available from the World Development Indicators (WDI). But there are several problems associated with using this data. First, while the measure of rents nets out the cost of production of a resource and therefore in theory seems a suitable measure, in practice the measure of the cost is a rough estimate taken at a single point in time and is heavily dependent on accounting assumptions; it is hence both uncertain and does not capture actual variations in extraction costs over time, which can be very large. Second, using the WDI 'rents' measures create a serious endogeneity problem: The measure relies on countries' self-reporting of their resource data to the World Bank and since our measure of transparency is based on the frequency and extent of reporting of data by countries (see next section), any missing observations may be an indication of low transparency rather than being randomly distributed. Third, there are many missing observations, which could bias our estimate of oil's impact. Fourth, due to the missing observations the scope of the coverage of WDI's data is somewhat limited, especially when it is combined with other variables.

To circumvent these problems, we will use an alternative measure, namely "oil and gas value" which is available from Ross (2013) which is simply the quantity of oil and gas extracted in a given year multiplied

[16] Wilson, Cobham and Goodall (2014) describe the new ICTD data as follows: "The dataset meticulously combines data from several major international databases, as well as drawing on data compiled from all available International Monetary Fund (IMF) Article IV reports. It achieves marked improvements in data coverage and accuracy, including a standardized approach to revenue from natural resources."

[17] we must account for variations in the tax rate, t, in the $t\mu^*$ expression where μ^* is the optimal compliance rate. No systematic data exists for the tax rate across countries and over time. We try to address this challenge as best we know how: We "back out" tax compliance from the tax revenue data by noting from equation 16 that, $t = h(G|L, T)$. This means that controlling for variations in G, L and T should also control for variations in t, at least theoretically. Then optimal compliance $\mu^*(\bar{t}, L, T)$ can then be backed out from $\bar{t}\mu^*(\bar{t}, L, T)$. (Note that equations 15 or 16 are in terms of G, while the discussion here is in terms of G/Y. These are equivalent since the difference is absorbed in the value of Ω in 15.)

[18] Friedman, et. al. (2000) have shown that tax avoidance is associated with a greater share of the unofficial economy. But thier data is for 1995 only. Using their estimate for 1- share of unofficial economy for their 69 countries in 1995 (which is their available data) to convert tax avoidance to tax compliance, a simple same IV model was estimated for their tax compliance variable and for ours, when our data is tailored to conform to theirs. We then examined the correlation between their predicted values of compliance and ours from these IV regressions and found the correlations ranged anywhere from a low value of 78% to a high value 87% depending on the choice of controls and regressors. Thus, there is some confidence that the measure above captures actual tax compliance.

by the per-unit world price. Ross cleaned and compiled data from the U.S. Geological Survey, the U.S. Energy Information Administration's International Energy Statistics, the World Bank, and the BP Statistical Review. This dataset covers 172 countries, of which 95 have produced oil, from 1932–2011. Due to its approximate independence from the pure WDI measure, this measure allows us to circumvent, to a large extent, the above endogeneity issue and yet it is highly correlated with the WDI measure. Furthermore, our measure is able to avoid the thorny challenge of estimating extraction costs. Consequently, it is able to cover nearly all countries.

Most of the other variables are from WDI, but democracy data is from Polity IV. The data coverage is 1980-2010 for approximately 150 countries. The latter varies slightly depending on the variables used. Appendix 5 provides information on variables, data sources, and descriptive statistics.

4.1 Data on Instruments

Despite its many merits discussed above, oil value still represents revenue and much has been written about its potential endogeneity to institutions. We cover this issue in our discussion of the instrumental variable (IV) method. For this purpose, we use two instruments for oil value; *oil prices* (in 2000 constant prices) and *oil reserves* from the giant oil field discoveries data by Horn (2014). The appropriateness of using oil reserves is extensively discussed in Cassidy (2016). This type of data is also used in a recent paper by Arezki, Ramey and Sheng (2017). Although we utilize these instrument in our IV regressions, we also use them as "inputs" for oil values in our SEM regressions. Besides instruments for oil, we also employ data on geography variables (latitude and regions of the world) as instruments for per capita income, following Alexeev and Conrad (2009). We do this in order to address endogeneity concerns such as the possibility of oil boosting per capita incomes (Alexeev and Conrad 2009; Smith 2015; Arezki, Ramey and Sheng 2017). These instrument choices and the related literature are discussed more extensively in the discussion of the IV regressions.

5 Estimation Strategies, Models and Results

To examine different aspects of the model and the theory, we use three different estimation techniques. Each techniques has a strength and a weakness. Each weakness is then addressed by another technique. The hope is that a fuller and more robust examination of the theory emerges. Specifically, structural equation modeling (SEM) by its *systems* estimation capability captures key interrelationships within the model, but it is not sensitive to common endogeneity concerns that prevail in this literature. To address these concerns, instrumental variables (IV) techniques are used. Here we discuss the various endogeneity concerns and show how the choice of our instruments addresses these concerns. Finally, the question arises whether the relation between oil, taxation and transparency is something that holds primarily across countries, or over time and, if the latter, whether they are short-lived or are longer-sustained processes. Given the large T in the sample and the distinction between short-run and long-run, such considerations cannot be fully addressed by pure panel techniques. Here, we rely on a most recent variant of the well known autoregressive distributive lag model that also allows for cross-sectional dependence in the error structure, known as CD-ARDL (e.g., Chudik, et. al. 2015). In what follows, each estimation strategy is discussed in greater detail, followed by relevant econometric model, estimation results and, where appropriate, robustness checks.

5.1 Estimating Model 1: Structural Equations Model (SEM)

The SEM methodology is employed here to estimate the combined relationship that connects the oil-transparency path as described by hypothesis H3a with the transparency-tax compliance path as described by hypothesis H2a. (The hypotheses relating to tax enforcement will be tested separately). In addition, a separate SEM is also used to estimate the reduced form hypothesis H1 linking oil revenues directly to taxation. The usefulness of SEM for a case such as ours has been known for some time. For example, SEM allows for the estimation of multiple simultaneous equations containing several endogenous variables as well as structural relationships between variables (Bollen and Long 1993). This allows us to properly test the validity of our hypotheses. With this approach, we can simultaneously test measurement and structural relationships and account for potential measurement errors, making it superior to regular first generation techniques such as SURE (Chin 1998). For a more recent exposition of SEM see Maskus (2007) and Bowen and Guo (2012).

The need for SEM to test for structural relationships is obvious. But using SEM to test the simple reduced form relationship of hypothesis H1 requires some explanation. This is because we will allow for oil revenue to depend on oil prices and oil reserves *prior* to estimating the relationship between oil revenue and taxation. As a result, we would have a two-equation system to estimate rather than a single equation. The equations are shown below. While strictly speaking, oil prices and reserves are not "instruments" in the framework of the SEM regressions, the instrumental variable (IV) approach which we present later does utilize both variables as proper instruments. This is in keeping with the literature that we cover extensively in the section on IV. For SEM regressions we can think of oil prices and reserves as simply some exogenous inputs.

5.1.1 SEM 1: Estimating Direct Oil-Taxation Hypothesis (H1)

Based on this discussion, we estimate the following two equations in a simple SEM model:

$$(oil\ value/Y)_{it} = a_0 + b_0(oil\ price)_t + c_0(oil\ reserves)_i + u_{0i} + v_{0t} + \varepsilon_{0it} \tag{24}$$

$$(tax\ rev/Y)_{it} = a_1 + b_1(oil\ value/Y)_{it} + c_1(G/Y)_{it} + \tilde{\mathbf{d}}_1'.\tilde{\mathbf{X}}_{it} + u_{1i} + v_{1t} + \varepsilon_{1it} \tag{25}$$

The residuals, u_{0i} and v_{0t} are country and time fixed effects and ε_{0it} is the regression error that includes unobserved heterogeneities. Equation 25 is a simple accounting relationship that was represented by equation 7 and reflects the government's budget constraint. In that equation, $t\mu$ was the government's need ("*demand*") for tax revenues, driven by government spending G/Y. In 25 $\vec{X}_{it}$ is the column vector of other controls, $\tilde{\mathbf{X}}_{it}' = (GDP/cap_{it},\ Polity2_{it},\ trade\ taxes_{it})$ discussed shortly below, and $\tilde{\mathbf{d}}_1'$ is the row vector of the corresponding coefficients. (u_{1i}, v_{1t} and ε_{1it} are defined similarly to 24). We expect $b_1 < 0$ and $c_1 > 0$. Equation 25 assumes that G/Y is exogenous. This assumption will be tested in the IV estimation subsection.

5.1.2 SEM 2: Estimating Oil-Transparency & Transparency-Compliance Hypothesis (H2 and H3)

The system of SEM equations associated with this estimation is as follows:

$$(oil\ value/Y)_{it} = a_0 + b_0(oil\ price)_t + c_0(oil\ reserves)_i + u_{0i} + v_{0t} + \varepsilon_{0it}$$

$$(trans)^*_{it} = a_2 + b_2(oil\ value/Y)_{it} + \tilde{\mathbf{c}}_2'.\tilde{\mathbf{X}}_{it} + u_{2i} + v_{2t} + \varepsilon_{2it} \quad (26)$$

$$(tax\ compliance)_{it} = a_3 + b_3(trans)^*_{it} + c_3(trans^*)^2_{it} + d_3(G/Y)_{it} + \tilde{\mathbf{f}}_3'.\tilde{\mathbf{X}}_{it} + u_{3i} + v_{3t} + \varepsilon_{3it} \quad (27)$$

The first equation (not numbered) is a repeat of 24 because a_o and b_o need to be re-estimated simultaneously with the other two equations. Equation 26 tests hypothesis H3a and equation 27 tests hypothesis H2a. The order reversal in the testing of the two hypotheses follows from the game structure in that tax payers (followers) respond to government's (leader) actions. From the hypotheses, we expect that $b_2 < 0$ and $b_3 > 0$ and $c_3 < 0$. We may note that the right hand side of both equations 27 and 26 are assumed exogenous. Departures from this assumption are examined later under IV estimation.

5.1.3 Control Variables

The inclusion of GDP/cap controls for the possibility that a positive resource shock (an increase in resource revenues per capita) would lead to an increase in GDP causing taxes/GDP ratio to fall. Alexeev and Conrad (2009) raise a concern that is relevant to the transparency equation (26) in which a boost in GDP/cap due to resource revenues places oil producers in the class of richer countries with better (in our case more transparent) institutions so that by comparison they look less transparent, a result that is less due to an oil curse than a boost in the income of oil producers. In the IV estimation section we will use Alexeev and Conrad's specific instruments in addition to other instruments to address this issue.

The inclusion of $Polity2$, which measures democratic freedoms, controls for variations in G/Y and in the dependent variable, $trans$, related to the degree of democracy. For example, governments that are more responsive to their citizens may spend differently (or more) on public goods than otherwise, all else constant. More democratic regimes are also likely to be more transparent. The inclusion of $Polity2$ also controls for "the Norway factor", i.e., it controls for countries that have high quality institutions to begin with. (The further possibility that polity2 may be endogenous due to reverse causality with transparency is considered later in the IV estimation.)

The inclusion of *trade taxes* addresses the potential influence that a weak income tax environment can lead to increased resource extraction by the government. Since trade taxes are easier to impose and collect in many developing economies with weak income tax environments, trade taxes would be expected to also be higher in such an environment. Thus, one should observe a correlation between trade taxes and oil revenues so that the inclusion or the exclusion of trade taxes would have an effect on the size of the oil coefficient in the various regressions above.

5.1.4 SEM Estimation Results

Table 1 reports the results. The left panel estimates the system of SEM1 consisting of the regression of oil value on oil price and oil reserves (equation 24) followed by the regression of taxation on oil value (equation 25). Results show that both oil price and oil reserves are strongly and positively associated with oil value and that the oil value is significantly and negatively associated with tax revenues (hypothesis H1). Additionally, as we expected, government spending has a highly significant and positive coefficient, consistent with the notion that G drives the demand for taxation.

In sum, the underlying structural mechanism is examined by the second and the third panels testing respectively, the effect of oil on transparency (equation 26) and transparency on tax compliance (equation 27) but both simultaneously estimated. Results strongly confirm that oil is adversely associated with transparency and that transparency affects tax compliance positively. The latter effect is also found to be concave (negative quadratic coefficient). This is consistent with the citizens responding to transparency and the government optimizing on transparency, confirming hypotheses H2 and H3.

Tables 1 goes here

5.1.5 Robustness

Two robustness tests are included. First, introducing rule of law as a proxy of tax enforcement (see Appendix 6) does not deter from the robustness of the transparency-tax compliance results: As shown in Appendix 6, we see that although stricter rule of law does improve tax compliance in some of the empirical specifications, tax compliance still responds to transparency as it did in table 1, even in the presence of the rule of law, that is, concavity still holds in the transparency-tax compliance relationship. By contrast no concavity is observed in the relationship between rule of law and tax compliance. This suggests that of the two factors, transparency and rule of law, it is only transparency that remains the key strategic variable in the interaction between the citizens and the government within the framework of the game.

Second, we re-examine table 1, using direct tax compliance data available for 2004-2010 from the Fraser-Chain linked index. Results are reported in table 2. The first three columns are the same as table 1 since they reflect the reduced form regression which is about tax revenue, not compliance. The rest of the results reinforce the main aspects of table 1: the negative role of oil on transparency and the concavity in the effect of transparency on tax compliance.[19]

Tables 2 goes here

A word about the effect of international trade taxes: In both tables 1 and 2, international trade taxes have a negative and highly significant effect on transparency, suggesting that more closed economies may be less transparent. But as for the notion that a weak tax environment may drive oil extraction and therefore its effect on transparency, the coefficient of the oil variable in the transparency equation does not change by much from column 5 to column 6. On the other hand, in table 2 that same coefficient falls in size and loses all significance. Since neither the model nor the variables change in columns 4-6 between the two tables, the difference must be either due to the more recent and shorter timing of table 2 (2004-2010) versus table 1 (1980-2010) or from the smaller number of observations in table 2 compared to table 1.

In sum, despite the differences in the timing and the measure of tax compliance between the two tables, it is remarkable that the key concavity results remain robust.

5.2 Estimation Model 2: Instrumental Variables (IV)

Although the SEM estimation can better reflect the structural features of the model, it is not as sensitive to endogeneity issues. There are several different endogeneity concerns that arise depending on the equation under investigation. We will discuss each separately.

[19] Here, we had to drop the oil reserves as a factor in oil values in column 4-6 due to the small number of obervations.

5.2.1 Endogeneity in the Direct Oil-Taxation Equation

In estimating equation (25), several endogeneity issues arise that are due to possible reverse causalities. First, although public spending was considered exogenous to, and a driver of, the taxation decision in the model, the empirics must allow for the possibility that taxation may also drive public spending. For example, expectation of higher tax revenues might influence public officials to spend more. Second, the taxation-representation argument discussed in the introduction implies that the need for taxation by the government may actually drive more open political systems so that taxation may also influence the Polity2 variable. Third, the need for more tax revenues may drive the decision to extract more oil so that taxation could also drive oil revenues. These possibilities will be considered presently in the IV estimation. The instrument choices are discussed shortly below.

5.2.2 Endogeneity & Instrument Choice in Oil-Transparency Equation

In estimating the transparency equation several potential endogeneity issues and one critical collinearity issue arise. First there is a potential endogeneity of oil revenue with respect to transparency. Countries that are more transparent may provide a better environment for foreign investments in oil (and other natural resources), creating a reverse causality from transparency to oil revenues. For example, Bohn and Deacon (2000) find that democratic institutions and political stability positively affect investments in oil exploration. Cust and Harding (2014) show that when oil is potentially located on a national border, 95 percent more exploration occurs in the country with relatively better institutions. One approach to address this issue has been to use oil price fluctuations as an identification strategy (Bruckner et. al. 2012; Dube and Vargas 2013; Andersen et. al. 2014; Carreri and Dube; 2015; Caselli and Tesei 2016). The problem with this approach is the lack of cross country variations. Specifically, the use of price shock ignores revenue variability which depends on quantity of oil extracted in each country. Price does offer the possibility of exogeneity (except perhaps Saudi Arabia) but does not offer cross country variations reflective of the countries' differences in endowments and cannot be used without some reference to the available or extracted oil. Cassidy (2016) uses geologically determined oil basins to generate cross country variation in oil wealth. We also use this variable. But this instrument lacks a time dimension. To address both issues, we use both instruments each dealing with a different problem, one the cross country variation and the other the time variation.

Second, there is a possibility of reverse causality in the relation between polity and transparency: one might ask whether it is democratic institutions that influence transparency, or that greater transparency induces more democratic institutions?

Third, in the model public spending, G, is exogenous to the government's decision of how transparent to be. But Cassidy (2016) argues that resource windfall reduces the marginal utility of public goods to policy makers, causing them to invest less in fiscal capacity, though he does not provide evidence. In addition, there may be an omitted variable correlated both with G and transparency: In the fiscal capacity model of Besley and Persson (2009, 2010, 2011) when institutions are relatively more "cohesive" a "common-interest" state is more likely to emerge. Such a state probably spends more on public goods, G. If greater cohesion also implies greater transparency because such a state is less threatened by disclosures due to greater public trust, then the unobserved variable "cohesion" is correlated with both G and transparency.

Fourth, so far we have used per capita income as a control in the transparency (and other) regressions. But oil revenues boost income as is shown by Arezki, Ramey and Sheng (2017) using giant oil field discoveries and also by Smith (2015). Moreover, Alexeev and Conrad (2009) show that the negative cross-sectional association between oil and the quality of institutions disappears after controlling for (instrumented) GDP. We follow Alexeev and Conrad's (2009) lead, using their geography instruments (absolute latitude and regions of the world) for GDP per capita when it enters as a control. This would make it possible to address both the overestimation of the effect of oil on institutions (Alexeev and Conrad's concern), coming from the association between GDP per capita and oil revenues, and the potential endogeneity of GDP vis-a-vis transparency, since greater transparency may boost incomes through improvements in domestic and foreign investments in the oil sector as was discussed above.

5.2.3 IV Estimation Strategy

The estimation challenge lies in the fact that the geography instruments as well as oil reserves instrument are static[20] but the oil price instrument is dynamic. In general, estimation is difficult in such cases in that differencing methods used in panel estimation remove the static instruments. However, a panel-based IV estimation method, known as ec2sls, exists that allows for combining static and dynamic instruments. While the commonly used g2sls uses only the demeaned variables as instruments in the first stage, thus eliminating the static instruments, the ec2sls uses both the demeaned variables and their mean again as instruments in the first stage. The demeaned static instruments are still eliminated, but their means survive (Baltagi, 2008). Baltagi and Li (1992) and Baltagi and Liu (2009) discuss the superiority of these estimators over g2sls estimators. The Stock-Yogo weak instrument test, which is used in the case of multiple instruments, turns out to work very well here (Stock and Yogo 2005).[21]

Below, we first report the results for the direct effect of oil on taxation. This is followed by the results for the oil–transparency effect (path 1) and then for the transparency–tax compliance effect (path 2). Each result is followed by its own robustness test.

5.2.4 Results: The Direct Effect of Oil on Taxation

First we examine the *direct* (reduced form) effect of oil on taxation revenue, just discussed. This is analogous to equation 25 in the SEM estimation technique except that the variables are allowed to be endogenous. The instruments used for this purpose are the oil prices (in 2000 dollars), and oil reserves from giant oil field discoveries dataset for oil revenue, as well as several geography variables, i.e., latitude and region dummies (Europe, East Asia, Latin America, Middle East) for per capita income. Table 3 indicates the second-stage estimation results. Different columns are based on which variable(s) or combination of variables are considered endogenous. All columns indicate negative and significant coefficients of the oil variable, at least at 10 percent by mostly at 5 percent, regardless of the choice of the endogenous variables. However, notice that both the size and the significance of the effect increases with the number of endogenous variables that are instrumented for. The F statistics for weak instrument tests, along with their critical values, are reported at the bottom of each column. The F values are from Stock and Yogo (2005) based on (a) the maximal possible 2SLS bias (at 5% level of significance) of the estimated coefficients of the instruments in the first stage estimation relative to the OLS estimates by at most 20%, 10% or 5%, and (b) the number of instruments

[20] The latter is not entirely static: oil reserves data is sporadically available in different times for different counties over the sample period.

[21] Other overidentification and underidentification tests perform differently. But in our case the F tests are more restrictive and therefore preferable. For example in estimating the transparency equations, the Anderson (1984) underidentification test almost always implied no underidentification, while the F test's results were more restrictive. As for overidentification, the Sargan (1958) test whose null is that instruments are *not* overidentified, was rejected suggesting overidentification. But Chao et al (2012) show that the Sargan test is overly sensitive to the number of instruments. They provide a new statistic, corrected for rejection frequency, which increases with the number of instruments.

and endogenous variables (see Stock and Yogo's table 1 for details). Except for column 1 which admits a bias of as much as 20% and column 4, for all other columns the instrument choices strongly reject the weak test.

Table 3 goes here

Robustness: Using lagged oil instead, and repeating the regressions, table 4 finds negative and significant coefficient of oil and strong F test relative to their critical values indicating the rejection of the weak instrument test.

Table 4 goes here

5.2.5 Results: Oil–transparency Effect

With the evidence that has been established so far, pointing to the negative effect of oil revenues on taxation, the next task in the IV framework is to re-examine our explanation that this effect arises because oil abundance represses governments' desire for transparency which in turn represses tax compliance.

Estimating the oil-transparency path, similar to equation 26, table 5 shows negative and significant effects mostly at 5% level regardless of the choice endogenous variables or any combination of them. The coefficient is generally stable but varies somewhat from -0.17 to -0.31. The large F values (generally far above their critical F values) indicate rejection of the weak instrument hypothesis.

Table 5 goes here

Robustness: Table 6 examines what happens if we replace current oil with lagged oil. To make the robustness analysis comparable, we keep the same choice of endogenous variables in each column of table 6 as in the corresponding column of table 5. Results continue to confirm the significant and negative coefficient of oil on transparency. Taken together, tables 5 and 6 give us indication that oil adversely influences transparency across countries and across time.

Table 6 goes here

5.2.6 Results: Transparency–Tax Compliance Effect

This part examines the hypothesis that tax payers' compliance behavior is rational and an optimal reaction to government transparency. It is the IV counterpart of equation 27 in SEM regressions. Since this effect is conditional on the government itself choosing the transparency level optimally (first mover), it is the *predicted* values of transparency (both linearly and quadratically) that are used here. For this purpose the values from table 5 are chosen although those from table 6 could be equally used. Results are reported in table 7, where column uses the predicted values of trans_hat(i) and trans_hat sq(i) (i=1...10) from the ith column in table 5. For example, trans_hat(3) and trans_hat sq(3) are predicted from column 3 of table 5. Finally, table 7 allows for GDP to be endogenous (the next table checks the robustness of the results by also allowing government spending to be endogenous).

Results from all ten columns indicate highly significant (at 1% and 5%) coefficients of transparency consisting of a positive linear and negative quadratic term. This concavity result is consistent with the prediction of the model regarding the tax payers' tax compliance/avoidance behavior. The F tests reject the strictest version of the weak instrument hypothesis per Stock and Yogo (2005) (i.e., with maximal allowable bias of at most 5% relative to OLS) mosly at 5% significance level with two columns rejecting it at the next allowable bias of at most 10%)

Table 7 goes here

Robustness: Table 8 re-estimates the transparency-compliance hypothesis by allowing government consumption to also enter as an endogenous regressor, besides income. In all cases the signs of the linear and quadratic coefficients confirm the concavity results. In five of the cases the results are significant or highly significant. In the remaining cases the quadratic terms are often highly significant but the linear term is not. Moreover, the absolute size of the coefficients are smaller than in table 7. Thus, while general concavity of tax revenues relative to transparency holds, it seems that the model's original assumption of *exogenous* government spending G produces sharper results than allowing G to be endogenous. But table 8 also rejects the weak instrument test as table 7 did. So, which result is to be considered more valid, endogenous or exogenous G? Some guidance can be gained from the much larger value of F statistic in table 7 compared to table 8, suggesting that for the instruments at hand an exogenous G is a more plausible assumption, at least for the type of regressions that test the transparency-tax compliance hypothesis.

Table 8 goes here

5.3 Estimation Model 3: Cross-Sectionally Augmented Autoregressive Distributive Lag (CS-ARDL)

Neither SEM nor IV methods can test for dynamics or distinguish between short-run and long-run. To examine these aspects, we turn to a dynamic estimation technique that adopts some of the more recent innovations in dynamic cross-sectional models with large N and large T. These innovations include lagged dependent variables, common correlated effects arising from cross-sectional dependencies and unobserved common factors, and slope heterogeneities (Pesaran 2006, Chudik and Pesaran, 2015). This involves augmenting a traditional autoregressive distributive lag (ARDL) model with cross-section averages to filter out the effects of the unobserved common factors, hence the acronym, CS-ARDL. This allows for correct estimation of the long-run effects.

An overview of the general ARDL approach is presented here. Consider the following regression model:

$$y_{it} = \sum_{j=1}^{p} \lambda_{ij} y_{it-j} + \sum_{j=0}^{q} \delta'_{ij} \mathbf{X}_{it-j} + \mu_i + u_{it} \tag{28}$$

where y_{it} is the dependent variable, $\mathbf{X}_{it-j}$ is the $k \times 1$ vector of explanatory variables, μ_i captures the individual country fixed effects and u_{it} is the error term. This equation can be reparametrized to obtain,

$$\triangle y_{it} = \phi_i y_{it-1} + \beta'_i \mathbf{X}_{it} + \sum_{j=1}^{p-1} \lambda^*_{ij} \triangle y_{it-j} + \sum_{j=0}^{q-1} \delta^{*\prime}_{ij} \triangle \mathbf{X}_{it-j} + \mu_i + u_{it} \tag{29}$$

where,

$$\phi_i = -\left(1 - \sum_{j=1}^{p} \lambda_{ij}\right), \boldsymbol{\beta}_i = \sum_{j=0}^{q} \delta_{ij}, \lambda^*_{ij} = -\sum_{m=j+1}^{p} \lambda_{im}, \, and \delta^*_{ij} = -\sum_{m=j+1}^{q} \delta_{im} \tag{30}$$

For a long-run relationship between the dependent and independent variables to exist, the condition $\phi_i < 0$ must be satisfied, where $\frac{\boldsymbol{\beta}_i}{\phi_i}$ and δ^*_{ij} are the long and short-run coefficients respectively.

The lag length of the model is chosen to be long enough to ensure that the error term is exogenous and serially uncorrelated, but not so long as to impose excessive data requirements, especially where time series observations are limited (Cavalcanti, et al 2015). Given the intent to see if the results can corroborate those from the earlier SEM specification, we allow for one lag to ensure the most use of the limited observations. Thus, to recreate the pattern in the SEM approach above, we estimate the following set of equations:

$$\begin{aligned} \triangle\, tax\_rev/Y_{it} = {} & \phi_{1i} tax\_rev/Y_{it-1} - \beta_{01i} - \beta_{11i} oil\_value_{it} - \beta_{21i} GDP\_PC_{it} \\ & - \beta_{31i} gov\_cons_{it} + \delta_{11i} \Delta oil\_value_{it} - \delta_{21i} \Delta GDP\_PC_{it} - \delta_{31i} \Delta gov\_cons_{it} + u_{1it} \end{aligned} \tag{31}$$

$$\begin{aligned} \triangle\, trans_{it} = {} & \phi_{2i} \left(trans_{it-1} - \beta_{02i} - \beta_{12i} oil\_value_{it} - \beta_{22i} GDP\_PC_{it} \right) \\ & + \delta_{12i} \Delta oil\_value_{it} - \delta_{22i} \Delta GDP\_PC_{it} + u_{2it} \end{aligned} \tag{32}$$

$$\begin{aligned} \triangle\, tax\_rev/Y_{it} = {} & \phi_{3i} tax\_rev/Y_{it-1} - \beta_{03i} - \beta_{13i} trans_{it} - \beta_{23i} trans^2_{it} \\ & - \beta_{33i} GDP\_PC_{it} - \beta_{43i} gov\_cons_{it} + \delta_{13i} \Delta oil\_value_{it} \\ & - \delta_{23i} \Delta GDP\_PC_{it} - \delta_{33i} \Delta gov\_cons_{it} + u_{3it} \end{aligned} \tag{33}$$

where, ϕ_{ji} are the error correction terms of the jth equation; β_{kji} are the long run coefficients of the kth variable in the jth equation, with $\beta_{0ji} = \frac{\mu_i}{\phi_{ji}}$ representing the country fixed effects; δ_{kji} are the short run coefficients of the kth variable in the jth equation; tax_rev/Y_{it} is the tax revenue for country i at time t; oil_value_{it} is the natural log of 1+oil value as a fraction of GDP; GDP_PC_{it} is the real GDP per capita; gov_cons_{it} is government consumption; $trans_{it}$ is the transparency measure for country i at time t.

Equation 31 captures the direct impact of oil revenues on tax revenues, while equations 32 and 33 show the impact of oil on tax compliance via its effect on transparency, per hypotheses.

5.3.1 ARDL Estimation Results

Results from this model are reported in table 9, with the long-run effects indicated by the top three panels and the short-run effects by the bottom three.

Table 9 goes here

In table 9, the first three columns show the direct (reduced-form) effects of oil on tax revenues. As indicated, all three long-run coefficients are negative and highly significant, consistent with our previous estimates, further supporting hypothesis H1. While the short-run effects are generally positive, the negative and significant coefficient of the error correction term points to convergence to the long run results.

The second three columns indicate the negative effect of oil wealth on transparency in the long run, consistent with our previous estimates, further supporting hypothesis H3a. Again, while the short-run results are in the reverse (positive) direction, we see evidence of convergence to the long-run.

The last three columns present the long and short-run effects of transparency on tax compliance as measured by tax revenues. Here while the coefficient of the linear and the quadratic terms have the right signs in columns 7 and 8 and are highly significant, indicating the usual concave structure, the quadratic term is not statistically significant in column 9 where we include Polity2. Thus, the long-run results are sensitive to the inclusion of Polity2. To further explore this issue, we conduct a robustness check below, using actual tax compliance data.

5.3.2 Robustness

We now re-estimate the last three columns of table 9 with actual tax compliance data from Fraser-Chain Linked Index rather than the ICTD tax revenue data. Results, reported in table 10, show that the long-run concavity in the last column is now robust to the inclusion of Polity2, while Polity2 itself is no longer significant–unlike table 9. We also note the reversal in the sign of government spending between the two tables. Despite these differences the remarkable consistency of the key result, i.e., the concavity pattern of the transparency variable, is noteworthy. This suggests that whether tax compliance is "backed out" from tax revenues (as we described earlier) or is directly measured, and whether the period of coverage is shorter (2004-2010) in the case of tax Fraser compliance data or longer (1980-2010) in the case of tax revenue data, transparency plays a key role in policy as well in tax compliance behavior.

When we combine the above outcome with our other results, i.e., the negative effects of oil wealth on taxation and on transparency, which also passed the robustness tests and held up throughout our investigation, we see remarkable support for the main hypothesis of this paper and our simple underlying expository model; statically and dynamically, across dramatically diverse economies, vastly different spans of time, and radically different estimation methods.

Table 10 goes here

6 Conclusion

We develop a simple analytic game in which an extractive (rentier) government moves first, optimizing over the level of transparency since, on one hand, transparency is costly to a rentier state by making it harder to divert resources and, on the other hand, it is needed to elicit citizen tax compliance. Citizens, knowing this, comply only partly with taxation since they cannot tell where their taxes go. Besides transparency as an inducement we also examine the role of tax enforcement as a force. Testing the model with extensive data and several distinct estimation techniques, we find support quite consistent with the structural and analytical mechanism suggested by the model.

Economic historians have suggested that the need for public finance in early modern Europe eventually forced governments to become more accountable to their citizens (North 1990, North and Weingast 1989, Hoffman and Norberg 1994). Seen in this light, our paper extends this insight in reverse in that windfall revenues from oil (and possibly other natural resources) diminish the incentive of governments to make their policies and institutions more transparent.

The relation between resource wealth and transparency is particularly salient in oil-rich Middle East and North Africa (MENA) region. But rigorous empirical analysis of this question is constrained by the paucity of taxation data for this region. Further, by removing the MENA group from the sample and re-running the regressions to test the MENA effect by its absence, one runs into another difficulty; namely removing the key source of variability, i.e., oil revenues. However, by simply comparing mean values, we find that the oil-rich MENA countries have exceptionally low taxes and low transparency, while the non-oil MENA countries have higher tax and transparency levels similar to the rest of the developing world. From this we infer that low transparency in the MENA region is largely a function of the region's concentration of oil wealth, rather than cultural or historical traits.[22]

If more oil revenues mean less accountability, less taxation and less citizen tax compliance, will the recent collapse in oil prices, if it becomes permanent, lead to a change in reverse by ushering a new era of greater accountability by governments, reliance on taxation, and more tax compliance coming from greater citizen participation? The answer would be affirmative, if the empirical findings were symmetric in either direction. But we cannot know this until more time has passed and more data has been gathered.

What can be said about the policy relevance of these findings? Besides an exogenous price collapse, in a world where the government does act in public interest (for instance if new government comes to power that is less burdened by rentier behavior), the policy relevance of our findings is tied to several proposals (Sala-i-Martin and Subramanian 2003, Devarajan et al.2011) in which the transfer of oil revenues directly to the public is a way of avoiding the resource curse. These proposals rely on the principle that households can better spend the money than an oil-rich government with weak accountability and suggest that transferring oil revenues to citizens and then taxing them to finance public goods may improve the quality of public expenditures. Our paper provides broad support of the proposition in the sense that higher levels of taxation are associated with more transparency. But the challenge is that a rentier government is likely to resist such proposals since adopting them may be self-defeating. Citing the Roman poet, Juvenal, the late Nobel laureate Leo Hurwicz, famously asked (2008) "But Who Will Guard the Guardians?"

7 References

Alexeev, M., and R. Conrad (2009). "The Elusive Curse of Oil." The Review of Economics and Statistics 91(3):586-598.

Andersen, J. J., N. Johannesen, D. D. Lassen, and E. Paltseva (2014). "Petro Rents, Political Institutions, and Hidden Wealth: Evidence from Offshore Bank Accounts." Working paper.

Anderson, T. W. (1984). Introduction to Multivariate Statistical Analysis. 2nd ed. New York: Wiley.

Arezki, R., V. A. Ramey, and L. Sheng (2017). "News Shocks in Open Economies: Evidence from Giant Oil Discoveries." Quarterly Journal of Economics. 32 (1): 103-155.

Baltagi, B. (2008). Econometric analysis of panel data. John Wiley & Sons.

Baltagi, B. and Q. Li (1992). "A Note on the Estimation of Simultaneous Equations with Error Components." Econometric Theory (81):113-119.

Baltagi, B. and L. Liu (2009). "A note on the application of EC2SLS and EC3SLS estimators in panel data models." Statistics & Probability Letters 79(20):2189–2192.

Barro, R. (1990). "Government Spending in A Simple Model of Endogenous Growth." Journal of Political Economy. 98(5) (part 2):S103 S125.

Besley, T. and T. Persson (2009). "The Origins of State Capacity: Property Rights, Taxation, and Politics." American Economic Review 99(4):1218–1244.

Besley, T. and T. Persson (2010). "State Capacity, Conflict, and Development." Econometrica 78(1):1–34.

[22] The analysis is available from the authors but not included here for brevity.

Besley, T. and T. Persson (2011). Pillars of Prosperity: The Political Economics of Development Clusters, Princeton, NJ: Princeton University Press.

Bohn, H. and R. T. Deacon (2000). "Ownership Risk, Investment, and the Use of Natural Resources." American Economic Review 90(3):526–549.

Bollen, K. A. and Long J.S. (1993). Testing Structural Equation Models. Sage, Newbury Park, CA.

Boschini, A. D., J. Pettersson, and J. Roine (2007). "Resource Curse or Not: A Question of Appropriability." Scandinavian Journal of Economics, 109(3):593–617.

Bowen, N.K. and S. Guo (2012). Structural Equation Modeling, Oxford, UK: Oxford University Press.

Brautigam, D., O.H. Fjeldstad, and M. Moore (eds.) (2008). Taxation and State-Building in Developing Countries: Capacity and Consent. New York: Cambridge University Press.

Brückner, M., A. Ciccone, and A. Tesei (2012). "Oil Price Shocks, Income, and Democracy." Review of Economics and Statistics, 94(2):389–399.

C´ardenas, M., S. Ramirez, and D. Tuzemen (2011). "Commodity Dependence and Fiscal Capacity." Working paper, Brookings Institution.

Carreri, M. and O. Dube (2015). "Do Natural Resources Influence Who Comes to Power, and How?" Working paper.

Caselli, F. and A. Tesei (2016). "Resource Windfalls, Political Regimes, and Political Stability." Review of Economics and Statistics 98(3):573-590.

Cassidy, T. (2016). "The Long-Run Effects of Oil Wealth on Democracy and Fiscal Capacity." Working Paper, University of Michigan.

Cavalcanti, T., K. Mohaddes and M. Raissi (2015). "Commodity Price Volatility and the sources of Growth." Journal of Applied Econometrics. 30(6): 857-873.

Chao, J., J. Hausman, W. Newey, N. Swanson & T. Woutersen (2012). "Testing Overidentifying Restrictions with Many Instruments and Heteroskedasticity." Working Paper, MIT

Chin, W. (1988) "Commentary: Issues and opinions on Structural equation modeling" MIS Quarterly. 22(1):7–16 (http://www.jstor.org/stable/249674)

Chudik, A. and M. H. Pesaran (2015). Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors. Journal of Econometrics 188:393-420.

Cornand, Camille and Frank Heinemann (2008). "Optimal Degree of Public Information Dissemination." The Economic Journal. 118(528):718–742.

Crivelli, E. and S. Gupta (2014). "Resource Blessing, Revenue Curse? Domestic Revenue Effort in Resource-Rich Countries." Working paper, IMF.

Devarajan, S., Hélène Ehrhart, Tuan Minh Le, and Gaël Raballand (2011). "Direct Redistribution, Taxation, and Accountability in Oil-Rich Economies: A Proposal." CGD Working Paper 281. Washington, D.C.: Center for Global Development. http://www.cgdev.org/content/publications/detail/1425822

Dube, O. and J. F. Vargas (2013). "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." Review of Economic Studies, 80:1384–1421.

Egorov, G., S. Guriev and K. Sonin (2009). "Why Resource-Poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." American Political Science Review 103(4):645.

Frankel, Jeffrey A. (2010). "The Natural Resource Curse: A survey." NBER Working Paper 15836.

Fraser Institute Annual Report (2010).

Friedman, E., S. Johnson, D. Kaufmann and P. Zoido-Lobaton (2000). "Dodging the Grabbing Hand: the Determinants of Unofficial Activity in 69 Countries." Journal of Public Economics 76:459–493.

Hoffman, P. T., and K. Norberg (eds) (1994). Fiscal Crises, Liberty, and Representative Government, 1450-1789. Stanford: Stanford University Press.

Hollyer, J., P. Rosendorff and J. Vreeland (2014). "Measuring Transparency." Political Analysis 22:413–434.

Horn, M. (2014) "Giant Oil and Gas Fields of the World,"

http://www.datapages.com/AssociatedWebsites/GISOpenFiles/HornGiantFields.aspx.

Hurwicz, L. (2008). "But Who Will Guard the Guardians?" The American Economic Review 98:577-585.

ICTD (2014) International Centre for Tax and Development Government Revenue Dataset, Brighton: International Centre for Tax and Development

International Monetary Fund (2012). "Fiscal Transparency, Accountability, and Risk." Fiscal Affairs and Statistics Department.

Jensen, A. (2011). "State-Building in Resource-Rich Economies." Atlantic Economic Journal 39:171–193.

Lane, P. R. and A. Tornell (1996). "Power, Growth, and the Voracity Effect." Journal of Economic Growth 1:213–241.

Maskus, K. (2007). "Structural Equation Modeling" in Encyclopedia of Industrial and Organizational Psychology. S.G. Rogelberg (editor). pp. 774-776. Thousand Oaks, CA: Sage Publications

McGuirk, E. (2010). "The Illusory Leader: Natural Resources, Taxation, and Accountability." Dublin: Trinity College.

Mehlum, H., K. Moene and R. Torvik (2006). "Institutions and the Resource Curse." The Economic Journal 116:1–20.

Mohtadi H. and T. Roe (1998). "Growth, Lobbying and Public Goods." European Journal of Political Economy 14(4):453-573.

Mohtadi, H. and T. Roe (2003). "Democracy, Rent Seeking, Public Goods and Economic Growth." Journal of Public Economics 87(3-4):445-466.

Mohtadi, H., S. Polasky (2016). "To Choose or not to Choose Free Trade? Role of Corruption and Tax Evasion." Working Paper, University of Wisconsin and University of Minnesota.

Mohtadi, H., M. Ross and S. Ruediger (2014). "Do Natural Resources Inhibit Transparency?" Economic Research Forum working paper 906; Working Paper, University of Wisconsin and UCLA.

North, D. (1990). Institutions, Institutional Change, and Economic Performance (New York: Cambridge University Press), p. 49.

North, D. and B. Weingast (1989). "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England" Journal of Economic History 49(4):803-832.

New York Times (2013). "President of Mexico Proposes Tax Overhaul" September 9. Panagariya, A. Miracles and Debacles: Do Free-Trade Skeptics Have a Case? Paper presented at the Allied Social Science Association Meetings, San Diego Calif. Jan. 2004.

Persson, T., G. Roland and G. Tabellini (2007). "Electoral Rules and Government Spending in Parliamentary Democracies" Quarterly Journal of Political Science. 2(2):155-188.

Pesaran, M. H. (2006). "Estimation and Inference in Large Heterogeneous Panels with Multifactor Error Structure" Econometrica 74:967-1012.

Polinsky, A. Mitchell and Steven Shavell (2001) "Corruption and Optimal Law Enforcement." Journal of Public Economics. 81:1–24.

Polity IV PROJECT (2013). Monty G. Marshall Center for Systemic Peace and Societal-Systems Research Inc by Gurr, T. and K. Jaggers.

Robinson, J. A., R. Torvik, and T. Verdier (2006). "Political Foundations of the Resource Curse." Journal of Development Economics 79:447–468.

Ross, M. (1999). "The Political Economy of the Resource Curse." World Politics 51(2):297-322.

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

Ross, M. (2011). "Mineral Wealth and Budget Transparency" Working paper, UCLA, Department of Political Science

Ross, M. (2012). "The Oil Curse: How Petroleum Wealth Shapes the Development of Nations?" Princeton: Princeton Univ. Press

Ross, M. (2013). "Oil & Gas Data, 1932-2011" available at:

http://thedata.harvard.edu/dvn/dv/mlross

Ross, M. (2015). "What have we learned about the resource curse?" Annual Review of Political Science:239-259.

Tsui, K. (2010). "Resource Curse, Political Entry, and Deadweight Costs." Economics & Politics 22(3):471-497.

Sala-i-Martin, X. & A. Subramanian (2003). "Addressing the Natural Resource Curse: An Illustration from Nigeria" NBER Working Paper 9804.

Sargan, J.D. (1958). "The estimation of economic relationships using instrumental variables" Econometrica 26:393-415.

Smith, B. (2015). "The Resource Curse Exorcised: Evidence from a Panel of Countries." Journal of Development Economics 116:57–73.

Stevens, P. and E. Dietsche (2008). "Resource curse: An analysis of Causes, Experiences and Possible Ways Forward." Energy Policy 36(1):56-65.

Stock, J. and M. Yogo (2005). "Testing for Weak Instruments in Linear IV Regression." Ch. 5 in J.H. Stock and D.W.K.Andrews (eds), Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg, Cambridge University Press. (Originally published 2002 as NBER Technical Working Paper No. 284).

Tanzi, V. and H. Zee (2000). "Interview: Developing countries face range of challenges in devising effective tax strategy." IMF Survey 29(13): 209 and 217-221.

Tornell, A. and P. R. Lane (1999). "The Voracity Effect." The American Economic Review, 89(1):22–46.

Wick, K. and E. Bulte (2009). "The curse of natural resources. Annual Review of Resource Economics." 1(1):139-156.

Wilson, P., A. Cobham and A. Goodall (2014). "The ICTD Government Revenue Dataset." ICTD Working paper 19

Williams, A. (2009). "On the Release of Information by Governments: Causes and Consequences." Journal of Development Economics 89(1):124-138.

Williams, A. (2011). "Shining a Light on the Resource Curse: An Empirical Analysis of the Relationship Between Natural Resources, Transparency, and Economic Growth" World Development 39(4):490-505.

Table 1: SEM Test of Model's Direct & Structural Relations with Tax Revenue Data (1980-2010)

	SEM1 System: Oil and Taxes			SEM2 System: Oil, Transparency and Tax Revenue					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable	Oil Value			Oil Value					
Oil Price	0.011***	0.011***	0.011***	0.011***	0.010***	0.011***			
	(0.003)	(0.003)	(0.003)	(0.0029)	(0.0031)	(0.0031)			
Oil Reserves	0.01***	0.009***	0.009***	0.0095***	0.0090***	0.0089***			
	(0.002)	(0.002)	(0.002)	(0.0021)	(0.0023)	(0.0022)			
Constant	-1.349*	-1.166	-1.347*	-1.355*	-1.171	-1.354*			
	(0.721)	(0.788)	(0.811)	(0.725)	(0.791)	(0.816)			
	Direct effect of oil value on tax revenue			Indirect effect of oil value on taxes through Transparency					
				Path 1: From Oil to Transparency			Path 2: From Transparency to tax Revenue		
Dependent Variable	Tax Revenue			Transparency			Tax Revenue		
Transparency							11.84***	8.619***	9.803***
							(3.907)	(2.937)	(2.890)
Transparency squared							-9.990**	-10.15***	-11.19***
							(4.421)	(3.191)	(3.332)
Oil Value	-7.41***	-5.86***	-5.7***	-0.248***	-0.148***	-0.163****			
	(1.195)	(0.593)	(0.551)	(0.0054)	(0.0249)	(0.0304)			
GDP PC	1.733***	1.603***	1.585***	0.0539**	0.0416***	0.0387***	1.692***	1.649***	1.606***
	(0.007)	(0.092)	(0.095)	(0.0021)	(0.00389)	(0.00394)	(0.0112)	(0.149)	(0.150)
Gov. Cons	0.137***	0.135***	0.137***	-024***	-.0027***	-.0025***	0.140***	0.130***	0.135***
	(0.007)	(0.007)	(0.007)	(0.0002)	(0.00003)	(0.00005)	(0.0043)	(0.0031)	(0.0034)
Polity2		0.0736*	0.0829**		0.0064**	0.0057*		0.118**	0.127**
		(0.0405	(0.0397		(0.0030)	(0.0030)		(0.0562)	(0.0551)
Inter. Taxes			-0.02***			-.0037***			-0.039***
			(0.004)			(0.0001)			(0.0077)
Constant	85.02***	102.7***	105.8***	-.269***	-4.681***	-4.380**	86.92***	92.27***	98.90***
	(4.261)	(12.78)	(13.48)	(0.785)	(1.785)	(1.894)	(19.29)	(23.21)	(23.39)
Countries	145	138	134	146	140	122	146	140	122
Observations	1538	1461	1407	1528	1451	1397	1528	1451	1397

For variable measures see Appendix 5. *,**,***: significant at 10%, 5% and 1%. Standard errors in parentheses. All estimates use structural equation modelling (SEM). Errors correct for clustering around MENA countries. First 3 columns show oil's direct effect on tax revenue; next 6 show oil's indirect effect through transparency: columns 4-6 show oil's effect on transparency; columns 7-9 show transparency's effect on tax revenue as a proxy for tax compliance. Column-pairs 4 & 7, 5 & 8 and 6 & 9 are estimated simultaneously with SEM but are split for reporting efficiency.

Table 2: Robustness Test of Table 1 with Tax Compliance Data (2004-2010)

	SEM1 System: Oil and Taxes			SEM2 System: Oil, Transparency and Tax Compliance					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable	Oil Value			Oil Value					
Oil Price	0.011***	0.011***	0.011***	0.0215***	0.0232***	0.00732			
	(0.003)	(0.003)	(0.003)	(0.00414)	(0.00411)	(0.0118)			
Oil Reserves	0.01***	0.009***	0.009***						
	(0.002)	(0.002)	(0.002)						
Constant	-1.349*	-1.166	-1.347*	4.128***	4.769***	7.723***			
	(0.721)	(0.788)	(0.811)	(1.547)	(1.377)	(2.964)			
	Direct effect of oil value on tax revenue			Indirect effect of oil value on tax compliance through Transparency					
				Path 1: From Oil to Transparency			Path 2: From Transparency to tax Compliance		
Dependent Variable	Tax Revenue			Transparency			Tax Compliance		
Transparency							13.96**	14.88***	10.06**
							(6.218)	(5.589)	(4.975)
Transparency squared							-14.30***	-15.13***	-12.03***
							(4.089)	(4.213)	(4.068)
Oil Value	-7.41***	-5.86***	-5.7***	-0.319***	-0.182**	-0.138			
	(1.195)	(0.593)	(0.551)	(0.0848)	(0.0885)	(0.0945)			
GDP PC	1.733***	1.603***	1.585***	0.0398***	0.0342***	0.0295***	0.509***	0.472**	0.483***
	(0.007)	(0.092)	(0.095)	(0.0103)	(0.0068)	(0.0057)	(0.157)	(0.225)	(0.182)
Gov. Cons	0.137***	0.135***	0.137***	-.000315	-0.0015	-0.0008	0.0407***	0.0388***	0.0247***
	(0.007)	(0.007)	(0.007)	(0.0016)	(0.0019)	(0.0028)	(0.0111)	(0.0022)	(0.0009)
Polity2		0.0736*	0.0829**		0.0062***	0.0059*		0.0231	0.0139
		(0.0405	(0.0397		(0.0020)	(0.0031)		(0.0780)	(0.0696)
Inter. Taxes			-0.02***			-.00421**			0.0162
			(0.004)			(0.0017)			(0.0106)
Constant	85.02***	102.7***	105.8***	-.269***	16.10***	13.49***	-132.9***	-135.2***	-212.0***
	(4.261)	(12.78)	(13.48)	(0.785)	(2.656)	(2.614)	(0.615)	(1.531)	(35.30)
Countries	145	138	134	146	140	122	146	140	122
Observations	1538	1461	1407	892	866	656	892	866	656

For variable measures see Appendix 5. *,**,***: significant at 10%, 5% and 1%. Standard errors in parentheses. All estimates use structural equation modelling (SEM). Errors correct for clustering around MENA countries. First 3 columns show oil's direct effect on tax revenue; next 6 show oil's indirect effect through transparency: columns 4-6 show oil's effect on transparency; columns 7-9 show transparency's effect on *actual* tax compliance. Column-pairs 4 & 7, 5 & 8 and 6 & 9 are estimated simultaneously with SEM but are split for reporting efficiency.

Table 3. Direct effect of oil on taxation (instrumental variables regressions with various endogeneity assumptions)

	Taxes						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Oil Value	-1.303*	-1.217*	-2.274*	-5.561*	-6.991**	-6.749**	-9.908**
	(0.433)	(0.313)	(0.544)	(1.753)	(1.261)	(1.229)	(1.610)
GDP PC	1.061***	0.845***	1.789***	0.735**	1.778***	1.665***	1.850*
	(0.0241)	(0.0279)	(0.0849)	(0.0846)	(0.0493)	(0.0204)	(0.481)
Gov. Cons	-0.0452*	-0.0419	0.0546***	0.0299	0.131*	0.117*	0.158***
	(0.0127)	(0.0176)	(0.0013)	(0.0213)	(0.0346)	(0.0312)	(0.0023)
polity2	0.0489***	0.0583***	0.0437***	0.0561***	0.0370***	0.0413***	-0.0668
	(0.0011)	(0.00026)	(0.0011)	(0.0011)	(0.0023)	(0.0003)	(0.202)
Inter. taxes	0.0472**	-0.356***	0.0450*	-0.368**	0.194**	0.0375	0.0523
	(0.0096)	(0.0118)	(0.0132)	(0.0439)	(0.0229)	(0.0156)	(0.0450)
Constant	-4.432***	-2.068***	-11.92***	-2.169*	-13.15***	-11.75***	-13.30**
	(0.0895)	(0.134)	(0.567)	(0.595)	(0.813)	(0.470)	(3.084)
Endogenous Variables	Gov.Cons	Gov. Cons Inter. taxes	GDP PC Oil Value	Gov. Cons Oil Value Inter. taxes	Gov. Cons Oil Value GDP PC Inter. taxes	Gov. Cons Oil Value GDP PC	Gov. Cons GDP PC Oil Value polity2
F	$9.05^{\$}$	$15.10^{\$\$}$	$17.49^{\$\$\$}$	$10.06^{\$\$}$	$30.74^{\$\$\$}$	$30.14^{\$\$\$}$	$37.86^{\$\$\$}$
Stock-Yogo critical F&	6.73	9.92	16.88	8.50	<13.95	<13.95	<13.95
Countries	134	134	134	134	134	134	134
N	2394	2394	2394	2394	2394	2394	2394

For variable measures see Appendix 5. *,**,***: significant at 10, 5 and 1 percent. Standard errors are in parentheses. &: The critical values of Stock-Yogo (2005) F test are based on the maximal possible TSLS bias (at 5% level of significance) of the coefficients of the instruments in the first stage estimation relative to OLS estimates at most by 20% ($\$$), 10% ($\$\$$) or 5% ($\$\$\$$). The critical values are also based on the number of instruments and the number of included endogenous variables. Instruemnts are oil reserves, oil prices, latitudes and world regions. All regressions were carried out using instrument variable estimation for panel data (xtivreg command on stata), errors are corrected for clustering around the MENA country and year fixed effects are introduced in all columns.

Table 4. Robustness check of Table 3: Direct Effect of *Lagged* Oil on Taxation (iv regressions)

	Taxes								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.Oil Value	-0.463***	-7.698**	-8.244**	-6.971**	-6.731**	-0.424***	-6.752**	-6.697**	-7.331*
	(0.0246)	(1.420)	(1.126)	(0.871)	(0.797)	(0.0094)	(1.267)	(1.152)	(2.429)
GDP PC	0.867***	1.064***	0.748**	1.785***	1.650***	0.971***	1.119***	0.886***	2.009*
	(0.0045)	(0.0610)	(0.0879)	(0.0674)	(0.0417)	(0.0089)	(0.0612)	(0.0884)	(0.537)
Gov. Cons	-0.0243**	0.0446	0.0456	0.140**	0.118**	-0.0607***	0.0109	0.00882	0.155**
	(0.0038)	(0.0245)	(0.0217)	(0.0218)	(0.0189)	(0.0054)	(0.0311)	(0.0272)	(0.0157)
polity2	0.0611***	0.0492***	0.0585***	0.0390***	0.0445***	0.0588***	0.0495***	0.0557***	-0.0335
	(0.0011)	(0.0012)	(0.0028)	(0.0008)	(0.0009)	(0.00002)	(0.0007)	(0.0027)	(0.213)
Inter. taxes	-0.361***	0.0385*	-0.390**	0.230**	0.0378				
	(0.0248)	(0.0112)	(0.0456)	(0.0239)	(0.0136)				
L.Inter. taxes						-0.339**	0.0492*	-0.340**	0.281*
						(0.0591)	(0.0156)	(0.0676)	(0.0878)
Constant	-2.610***	-5.705***	-2.423*	-13.48***	-11.70***	-2.731***	-5.620***	-2.926**	-15.32*
	(0.153)	(0.297)	(0.617)	(0.783)	(0.478)	(0.273)	(0.176)	(0.548)	(3.724)
Endogenous Variables	Gov. Cons Inter. taxes	Gov. Cons L.Oil Value	Gov. Cons L.Oil Value Inter. taxes	Gov. Cons L.Oil Value GDP PC Inter. taxes	Gov. Cons L.Oil Value GDP PC	Gov. Cons L.Inter. taxes	Gov. Cons L.Oil Value	Gov. Cons L.Oil Value L.Inter. taxes	Gov. Cons GDP PC L.Oil Value polity2 L.Inter. taxes
F	31.60$^{\$\$\$}$	29.39$^{\$\$\$}$	53.56$^{\$\$\$}$	63.99$^{\$\$\$}$	71.24$^{\$\$\$}$	39.36$^{\$\$\$}$	28.40$^{\$\$\$}$	33.78$^{\$\$\$}$	9.107$^{\$\$}$
Stock-Yogo critical F[&]	16.88	16.88	13.95	<13.95	13.95	16.88	16.88	13.95	<8.50
Countries	134	134	134	134	134	134	134	134	134
N	2339	2339	2339	2339	2339	2286	2286	2286	2286

Additional endogeneity tests are carried out over those in table 3 for robustness check. For variable measures see Appendix 5. *,**,***: significant at 10, 5 and 1 percent. Standard errors are in parentheses. &: The critical values of Stock-Yogo (2005) F test are based on the maximal possible TSLS bias (at 5% level of significance) of the coefficients of the instruments in the first stage estimation relative to OLS estimates at most by 20% (\$), 10% (\$\$) or 5% (\$\$\$). The critical values are also based on the number of instruments and the number of included endogenous variables. Instruemnts are oil reserves, oil prices, latitudes and world regions. All regressions were carried out using instrument variable estimation for panel data (xtivreg command on stata), errors are corrected for clustering around the MENA country and year fixed effects are introduced in all columns. "L". signifies the first lag of the variable in question.

Table 5. Effect of Oil on Transparency (instrumental variables regressions with various endogeneity assumptions)

	Transparency									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Oil Value	-0.270**	-0.314**	-0.128**	-0.253**	-0.139**	-0.282**	-0.130**	-0.176**	-0.174**	-0.192*
	(0.0577)	(0.0722)	(0.0285)	(0.0461)	(0.0274)	(0.0635)	(0.0207)	(0.0373)	(0.0330)	(0.0470)
GDP PC	0.0532***	0.0605***	0.0482***	0.0541**	0.0613***	0.0557**	0.0503**	0.0549**	0.0535***	0.0669***
	(0.0020)	(0.0031)	(0.00120)	(0.0070)	(0.0036)	(0.0064)	(0.0073)	(0.0058)	(0.0053)	(0.0001)
Gov. Cons	-0.00122*	-0.00128	-0.000940*	-0.00110*	-0.000820	-0.00115*	-0.000826*	-0.00409*	-0.00429*	-0.00481
	(0.0003)	(0.0004)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0012)	(0.0010)	(0.0019)
polity2	0.00607**	0.00342***	0.00821**	0.00432**	0.00351***	0.00350***	0.00610**	0.00373***	0.00546***	0.00356***
	(0.0014)	(0.00007)	(0.0012)	(0.0005)	(0.00007)	(0.00005)	(0.0007)	(0.00002)	(0.0002)	(0.00001)
Inter. taxes	-0.00161	-0.00110	-0.00158	-0.00132	-0.000450	-0.00123	-0.00127	-0.0000148	-0.000227	0.000804
	(0.0016)	(0.0021)	(0.0017)	(0.0023)	(0.0021)	(0.0022)	(0.0023)	(0.0017)	(0.0018)	(0.0015)
Constant	0.141***	0.0981	0.158***	0.139	0.0701	0.132	0.148	0.180	0.187	0.0962
	(0.0077)	(0.0419)	(0.0034)	(0.0674)	(0.0349)	(0.0668)	(0.0657)	(0.0721)	(0.0663)	(0.0332)
Endogenous Variables	GDP PC Oil Value polity2	GDP PC Oil Value	GDP PC polity2	Oil Value polity2	GDP PC	Oil Value	polity2	Gov. Cons	Gov. Cons polity2	Gov. Cons GDP PC
F	21.91$$$	18.92$$$	20.09$$$	30.06$$$	25.81$$$	19.75$$	39.24$$$	22.24$$$	27.65$$$	16.76$$
Stock-Yogo critical F&	13.95	16.88	16.88	16.88	19.86	11.22	19.86	19.28	16.88	9.92
Countries	149	149	149	149	149	149	149	149	149	149
Observations	3111	3111	3111	3111	3111	3111	3111	3111	3111	3111

For variable measures see Appendix 5. Standard errors in parentheses. Errors corrected for MENA clustering and year-fixed effects. *,**,***: significant at 10, 5 and 1 percent; &: critical values of Stock-Yogo (2005) F test based on maximal possible TSLS bias (at 5% significance) of the coefficients of instruments in first stage estimates, relative to OLS estimates at most by 20% ($), 10% ($$) or 5% ($$$). The critical values are based on the number of instruments and included endogenous variables.

Table 6. Robustness check of Table 5: Effect of Lagged Oil on Transparency (instrumental variables regressions)

	Transparency									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
L.Oil Value	-0.271*	-0.331**	-0.124*	-0.224**	-0.140**	-0.244*	-0.133**	-0.149**	-0.139**	-0.159*
	(0.0666)	(0.0734)	(0.0295)	(0.0485)	(0.0237)	(0.0806)	(0.0169)	(0.0343)	(0.0316)	(0.0386)
GDP PC	0.0517***	0.0593***	0.0467***	0.0569**	0.0604***	0.0575**	0.0534**	0.0541**	0.0521***	0.0680***
	(0.0034)	(0.0029)	(0.0035)	(0.0079)	(0.0031)	(0.0065)	(0.0083)	(0.0062)	(0.0047)	(0.0013)
Gov. Cons	-0.0005	-0.0004	-0.0005	-0.0004	-0.0004	-0.0004	-0.0004	-0.0045*	-0.0052*	-0.0056
	(0.0003)	(0.0004)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0014)	(0.0013)	(0.0024)
polity2	0.0062*	0.0034***	0.0084**	0.0036*	0.0035***	0.0035***	0.0052*	0.0041***	0.0062***	0.0037***
	(0.0019)	(0.00008)	(0.0016)	(0.0011)	(0.00002)	(0.00005)	(0.0013)	(0.0001)	(0.00004)	(0.00004)
Inter. taxes	-0.0017	-0.0013	-0.0016	-0.0010	-0.0005	-0.0011	-0.0010	-0.0001	-0.0001	0.0013
	(0.0018)	(0.0022)	(0.0018)	(0.0026)	(0.0022)	(0.0025)	(0.0026)	(0.0015)	(0.0015)	(0.0012)
Constant	0.140***	0.0937	0.161***	0.104	0.0691	0.100	0.119	0.190	0.208*	0.0963*
	(0.0011)	(0.0400)	(0.0062)	(0.0700)	(0.0311)	(0.0665)	(0.0703)	(0.0768)	(0.0664)	(0.0306)
Endogenous Variables	GDP PC L.Oil Value polity2	GDP PC L.Oil Value	GDP PC polity2	L.Oil Value polity2	GDP PC	L.Oil Value	polity2	Gov. Cons GDP PC polity2	Gov. Cons GDP PC	Gov. Cons polity2
F	$16.55^{\$\$\$}$	$20.33^{\$\$\$}$	$17.73^{\$\$\$}$	$21.28^{\$\$\$}$	$34.82^{\$\$\$}$	$9.169^{\$}$	$62.26^{\$\$\$}$	$19.02^{\$\$\$}$	$19.41^{\$\$\$}$	$16.87^{\$\$(1)}$
Stock-Yogo critical F$^{\&}$	13.95	13.95	16.88	16.88	19.86	6.73	19.86	13.95	16.88	9.92
Countries	149	149	149	149	149	149	149	149	149	149
Observations	3043	3043	3043	3043	3043	3043	3043	3043	3043	3043

(1): The F value for the last column is exceedingly close to the 5% critical value of 16.88.

For variable measures see Appendix 5. Standard errors in parentheses. Errors corrected for MENA clustering and year-fixed effects. *,**,***: significant at 10, 5 and 1 percent; &: critical values of Stock-Yogo (2005) F test based on maximal possible TSLS bias (at 5% significance) of the coefficients of instruments in first stage estimates, relative to OLS estimates at most by 20% ($), 10% ($$) or 5% ($$$). The critical values are based on the number of instruments and included endogenous variables.

Table 7. IV Estimates of the Effect of Transparency (predicted from Table 5) on Taxes: Case of Endogenous GDP per capita

	Taxes									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GDP PC	2.191***	2.768***	1.952***	2.458***	2.686***	2.672***	2.066***	2.554***	2.180***	2.857***
	(0.0714)	(0.0956)	(0.0853)	(0.006)	(0.149)	(0.0565)	(0.0548)	(0.0932)	(0.005)	(0.210)
Gov. Cons	0.0606***	0.0565***	0.0647***	0.0598***	0.0619***	0.0580***	0.064***	0.0468**	0.0667***	0.0315*
	(0.0012)	(0.0028)	(0.0013)	(0.002)	(0.0029)	(0.002)	(0.0017)	(0.006)	(0.003)	(0.0102)
Inter. taxes	0.0437*	0.0396*	0.0435*	0.0414*	0.0475**	0.0391*	0.0440**	0.0452**	0.0408**	0.0556**
	(0.0112)	(0.0106)	(0.0105)	(0.01)	(0.0097)	(0.0100)	(0.009)	(0.009)	(0.008)	(0.0118)
ri_hat_1	11.00**									
	(2.228)									
ri_hat_1_sq	-13.97**									
	(2.625)									
ri_hat_2		7.198**								
		(1.599)								
ri_hat_2_sq		-18.73***								
		(0.430)								
ri_hat_3			16.03**							
			(2.854)							
ri_hat_3_sq			-15.61**							
			(3.282)							
ri_hat_4				12.30**						
				(2.129)						
ri_hat_4_sq				-19.55**						
				(1.997)						
ri_hat_5					12.85**					
					(2.961)					
ri_hat_5_sq					-24.11***					
					(1.485)					
ri_hat_6						9.896**				
						(1.850)				
ri_hat_6_sq						-20.59***				
						(1.030)				
ri_hat_7							18.55**			
							(2.825)			
ri_hat_7_sq							-19.99**			
							(3.156)			
ri_hat_8								18.47**		
								(2.421)		
ri_hat_8_sq								-28.17***		
								(1.434)		
ri_hat_9									20.04**	
									(2.197)	
ri_hat_9_sq									-23.47***	
									(2.165)	
ri_hat_10										11.44**
										(2.536)
ri_hat_10_sq										-24.20***
										(0.603)
_cons	-16.72***	-17.57***	-17.18***	-17.78***	-18.49***	-17.77***	-18.10***	-19.01***	-18.71***	-18.46***
	(1.085)	(0.243)	(1.303)	(0.776)	(0.188)	(0.481)	(1.119)	(0.424)	(0.751)	(0.0870)
Endogenous Variables	GDP PC	GDP PC	GDP PC	GDP PC	GDP PC	GDP PC	GDP PC	GDP PC	GDP PC	GDP PC
F	28.34$$$	20.26$$$	22.60$$$	33.41$$$	18.83$$	28.60$$$	40.15$$$	58.19$$$	83.19$$$	20.36$$$
Stock-Yogo critical F&	19.86	19.86	19.86	19.86	11.29	19.86	19.86	19.86	19.86	19.86
Countries	134	134	134	134	134	134	134	134	134	134
N	2394	2394	2394	2394	2394	2394	2394	2394	2394	2394

ri stands for transparency. Predicted values of ri_hat_i (i=1...10) are the transparency estimates from table 5 for where "i" stands for each of the ith columns of table 5. ri_hat_i_sq is the square of ri estimates. SE's in (.). For variable measures see Appendix 5. Errors corrected for MENA clustering & year-fixed effects. *,**,***: significant at 10%, 5% and 1%. ; &: critical Stock-Yogo (2005) F test based on maximal possible TSLS bias (at 5%) of instruments in 1st stage coefficient estimates, relative to OLS, by at most 20% ($), 10% ($$) or 5% ($$$). Critical values are based on # of instruments & included endogenous variables.

Table 8. IV Estimates of the Effect of transparency (predicted from Table 5) on taxes: Case of endogenous Gov Spending and GDP per cap.

	Taxes									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GDP PC	1.941***	2.595***	1.776***	1.977**	1.179	2.424***	1.727**	0.927	1.420	0.340
	(0.0914)	(0.164)	(0.0615)	(0.239)	(0.684)	(0.165)	(0.197)	(0.970)	(0.534)	(1.211)
Gov. Cons	0.09**	0.075**	0.09**	0.093**	0.109*	0.078**	0.093**	0.230	0.181	0.276
	(0.0115)	(0.0107)	(0.0152)	(0.0156)	(0.0322)	(0.0103)	(0.0183)	(0.113)	(0.0884)	(0.119)
Inter. taxes	0.0395*	0.0374*	0.0394*	0.0362*	0.0234	0.0367*	0.0362	0.0213	0.0248	0.00196
	(0.0119)	(0.0119)	(0.0131)	(0.0111)	(0.0206)	(0.0109)	(0.0133)	(0.0237)	(0.0207)	(0.0355)
ri_hat_1	8.356*									
	(2.115)									
ri_hat_1_sq	-8.770***									
	(0.629)									
ri_hat_2		6.768*								
		(1.649)								
ri_hat_2_sq		-16.44***								
		(0.255)								
ri_hat_3			9.464*							
			(2.754)							
ri_hat_3_sq			-7.637**							
			(1.511)							
ri_hat_4				8.743*						
				(2.321)						
ri_hat_4_sq				-10.14***						
				(1.015)						
ri_hat_5					4.377					
					(7.437)					
ri_hat_5_sq					3.018					
					(1.470)					
ri_hat_6						8.973**				
						(1.935)				
ri_hat_6_sq						-16.66***				
						(0.264)				
ri_hat_7							8.835			
							(3.179)			
ri_hat_7_sq							-6.814***			
							(0.270)			
ri_hat_8								39.04		
								(15.26)		
ri_hat_8_sq								-25.19***		
								(1.277)		
ri_hat_9									31.43	
									(11.06)	
ri_hat_9_sq									-24.79**	
									(3.713)	
ri_hat_10										38.93
										(15.42)
ri_hat_10_sq										-19.92***
										(0.395)
Constant	-15.46***	-17.03***	-15.09***	-15.56***	-11.24***	-16.90***	-14.71***	-21.53**	-20.51**	-19.14***
	(0.532)	(0.0631)	(0.886)	(0.0936)	(0.273)	(0.186)	(0.504)	(2.233)	(2.238)	(0.718)
Endogenous Variables	Gov. Cons GDP PC	Gov. Cons GDP PC	Gov. Cons GDP PC	Gov. Cons GDP PC	Gov. Cons GDP PC	Gov. Cons GDP PC	Gov. Cons GDP PC	Gov. Cons GDP PC	Gov. Cons GDP PC	Gov. Cons GDP PC
F	15.60$^{\$\$}$	16.84$^{\$\$(1)}$	11.81$^{\$\$}$	14.19$^{\$\$}$	0.346	21.49$^{\$\$\$}$	7.725$^{\$}$	6.547$^{\$}$	8.073$^{\$}$	6.375$^{\$}$
Stock-Yogo critical F$^{\&}$	9.92	9.92	9.92	9.92	6.16	19.86	6.16	6.16	6.16	6.16
Countries	134	134	134	134	134	134	134	134	134	134
N	2394	2394	2394	2394	2394	2394	2394	2394	2394	2394

(1): The F value for the second column is exceedingly close to the 5% critical value of 16.88.

ri stands for transparency. Predicted values of ri_hat_i (i=1...10) are the transparency estimates from table 5 for where "i" stands for each of the ith columns of table 5. ri_hat_i_sq is the square of ri estimates. For variable measures see Appendix 5. SE's in (.). Errors corrected for MENA clustering & year-fixed effects. *,**,***: significant at 10%, 5% and 1%. &: critical Stock-Yogo (2005) F test based on maximal possible TSLS bias (at 5%) of instruments in 1st stage coefficient estimates, relative to OLS, by at most 20% (\$), 10% (\$\$) or 5% (\$\$\$). Critical values are based on the number of instruments and included endogenous variables.

Table 9. Dynamic Effects: CS-ARDL Estimates of the Reduced Form & Structural Equations

	Direct effect of oil on tax revenue			Indirect effect of oil on tax revenue through Transparency					
				Path 1: From Oil to Transparency			Path 2: From Transparency to tax revenue		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Long Run Effects								
Oil Value	-1.204***	-3.035***	-2.758***	-0.216***	-0.174***	-0.179***			
	(0.355)	(0.495)	(0.106)	(0.00634)	(0.00639)	(0.00366)			
GDP PC	0.978***	0.620***	0.597***	0.0356***	0.0264***	0.0382***	0.520***	0.308***	0.552***
	(0.0362)	(0.0288)	(0.0450)	(0.00838)	(0.00394)	(0.00441)	(0.0266)	(0.0388)	(0.0806)
Gov. Cons		0.111***	0.0961***		0.000646	0.000428		0.114***	0.103***
		(0.000113)	(0.000324)		(0.000549)	(0.000594)		(0.000515)	(0.0000740)
polity2			0.0640***			0.00439***			0.0620***
			(0.000842)			(0.000584)			(0.00467)
Transparency							6.442***	3.367***	0.655***
							(0.388)	(0.132)	(0.188)
Transparency squared							-2.944***	-0.747***	-0.163
							(0.339)	(0.215)	(0.240)
	Short Run Effects								
Error Correction	-0.162***	-0.171***	-0.194***	-0.214***	-0.229***	-0.235***	-0.174***	-0.179***	-0.195***
	(0.00197)	(0.00199)	(0.00523)	(0.000454)	(0.000822)	(0.00254)	(0.00215)	(0.00186)	(0.00502)
Oil Value	-0.552***	0.488***	0.479***	0.0296***	0.0338***	0.0309***			
	(0.0680)	(0.133)	(0.108)	(0.000169)	(0.0102)	(0.00812)			
GDP PC	0.0249	0.269***	0.247***	0.0227*	0.0324***	0.0342***	-0.121**	0.158***	0.179***
	(0.0567)	(0.0477)	(0.0347)	(0.0123)	(0.00887)	(0.00852)	(0.0535)	(0.0449)	(0.0475)
Gov. Cons		0.0214***	0.0209***		0.000144	0.000138		0.0162***	0.0168***
		(0.00131)	(0.00211)		(0.000165)	(0.000162)		(0.00205)	(0.00251)
polity2		0.0287	0.0380		-0.000181	-0.00187***			-0.00811***
		(0.0334)	(0.0337)		(0.00243)	(0.000278)			(0.000575)
Transparency			-0.00830***			-0.000524***	-1.519***	-1.411***	-0.800***
			(0.000406)			(0.0000603)	(0.227)	(0.265)	(0.258)
Transparency squared							1.116***	1.036***	0.695**
							(0.265)	(0.293)	(0.294)
Constant	24.32***	31.50***	52.55***	0.0796	-0.0337	1.475	2.677**	11.30***	-40.24*
	(0.729)	(3.632)	(19.80)	(0.154)	(0.0587)	(0.938)	(1.082)	(2.550)	(23.33)
Countries	144	137	133	163	154	148	155	137	133
Observations	2473	2312	2222	4337	4085	2997	2582	2297	2207

CS-ARDL refers to cross sectionally augmented autoregessive distributive lag estimates (see text for details). All regressions adopt dynamic fixed effects method (xtpmg command on stata with dfe option). For variable measures see Appendix 5. *,**,***: significant at 10, 5 and 1 percent. Standard errors are in parentheses. Errors are corrected for clustering around the MENA countries. Tax data are from ICTD. The first three columns capture the direct effect of oil on tax compliance, the next six columns capture the indirect effect through transparency, with columns 4 to 6 capturing the first path (effect of oil on transparency) and columns 7 through 9 capturing the second path (effect of transparency on tax revenue). All columns capture both the short and long run effects and all regressions are cross sectionally adjusted to account for cross sectional dependence such as time fixed effects or global shocks that impact all countries. To do this, cross sectional averages for all variables for each time period are introduced to the regressions. These averages are not reported.

Table 10. Transparency and Tax Compliance: Robustness Test of Table 9-columns 7-9, with Direct Tax Compliance Data

	robust check of col. 7 table 9	robust check of col. 8 table 9	robust check of col. 9 table 9
		Long Run Effects	
GDP PC	3.061***	2.933***	2.958***
	(0.599)	(0.581)	(0.624)
Gov. Cons		-0.0333**	-0.0339**
		(0.0138)	(0.0142)
polity2			-0.00170
			(0.00892)
Transparency	10.87***	5.410***	6.326**
	(2.996)	(1.802)	(2.465)
Transparency squared	-8.522***	-4.574***	-5.262**
	(2.674)	(1.750)	(2.232)
		Short Run Effects	
Error Correction	-0.405***	-0.413***	-0.414***
	(0.00590)	(0.00528)	(0.00481)
GDP PC	-0.673**	-0.806**	-0.813**
	(0.267)	(0.327)	(0.319)
Gov. Cons		0.0162***	0.0171***
		(0.00113)	(0.00166)
polity2			-0.00803
			(0.00607)
Transparency	-1.368	0.137	0.00364
	(1.208)	(1.210)	(1.318)
Transparency squared	0.911	-0.143	-0.0599
	(1.082)	(1.120)	(1.208)
Constant	-34.93	155.6***	-133.7***
	(187.8)	(18.99)	(42.24)
Countries	133	131	127
Observations	776	747	726

Estimates use tax compliance data from Fraser Chain-Linked Index for 2004-2010 instead of tax revenue data from ICTD that were used in Table 9. For variable measures see Appendix 5. *,**,***: significant at 10, 5 and 1 percent. Standard errors are in (.). All regressions adopt CS-ARDL dynamic fixed effect estimation methody (xtpmg command on stata with dfe option), errors correct for clustering around the MENA countries. First three columns show oil's direct effect on tax compliance, next six show oil's indirect effect through transparency, with columns 4-6 capturing the first path (oil's effect on transparency) and columns 7-9 capturing the second path (transparency's effect on tax compliance). Both short and long run effects are shown for all regressions. Regressions are cross sectionally adjusted to account for cross sectional dependence such as time fixed effects or global shocks that impact all countries. To do this, cross sectional averages for all variables for each time period are introduced to the regressions. These averages are not reported.

8 Appendices

8.1 Appendix 1: Proof of Lemma 1

Statement of Lemma 1: *Under reasonable upper bound on the tax rate, t,* $\partial b/\partial T \leq 0$ implies that $\partial t/\partial T < 0$

Rewrite the budget constraint, equation 7, for optimum compliance, μ^*:

$$(1-b)[\rho + t\mu^*(t, L, T)] = G/Y$$

Focusing on the effect of taxes t on transparency, holding enforcement constant, differentiation of the above equation in T yields:

$$(1-b)(t_T\mu^* + t\mu_T^* + t\mu_t^* t_T) - b_T(\rho + t\mu^*) = 0$$

Rearranging terms,

$$(1-b)(\mu^* + t\mu_t^*)t_T = -(1-b)t\mu_T^* + b_T(\rho + t\mu^*) = 0$$

Since $b_T \leq 0$ by assumption and with $\mu_T^* > 0$, it follows that $RHS < 0$. Using optimum μ^* from 13 in the expression $\mu^* + t\mu_t^*$ on the LHS of the above equation, we find $\mu^* + t\mu_t^* = 1 - 2t/[\gamma(T).f(L)]$. This shows that $\mu^* + t\mu_t^* > 0$ if $t < \gamma f/2$ which puts an upper bound on the tax rate. It follows that,

$$b_T \leq 0 \rightarrow t_T < 0$$

8.2 Appendix 2: Proof of Lemma 2

Statement of *Lemma 2. Under reasonable upper bound on the tax rate, t,* $\partial b/\partial L \leq 0$ implies $\partial t/\partial L < 0$ but $\partial b/\partial L > 0$ implies $\partial t/\partial L \gtreqless 0$.

Repeating the above procedure this time for L we have

$$(1-b)(\mu^* + t\mu_t^*)t_L = -(1-b)t\mu_L^* + b_L(\rho + t\mu^*) = 0$$

So, as in above case, with the same upper bound taxes as under lemma 1, and with $\mu_L^* > 0$, we see that $b_L \leq 0$ implies that the RHS < 0 so that $t_L < 0$ so that the behavior of tax enforcement is similar to transparency in this case. However, if $b_L > 0$, RHS can have any sign and hence so can t_L, i.e., $t_L \gtreqless 0$.

8.3 Appendix 3: Proof of Second Order Condition for Compliance

Differentiating the first order condition, equation 19, in T and rearranging the terms, we have:

$$Q \equiv b_T(t_T\mu^* + t\mu_T^* + t\mu_t^* t_T) + b_T(t_T\mu^* + t\mu_T^*) + b[2t_T\mu_T^* + t\mu_{TT}^* + t_T^2\mu_t^* + t\mu_{tT}^* t_T]$$

Use the first order condition equation 19 once again to substitute for the first two terms $b_T(t_T\mu^* + t\mu_T^*)$ in the above equation. We have:

$$Q = -2\frac{b_T^2}{b}.(\rho + t\mu^*) - t\mu_t^* t_T + tb_T\mu_t^* t_T + b[2t_T\mu_T^* + t\mu_{TT}^* + t_T^2\mu_t^* + t\mu_{tT}^* t_T]$$

In this equation the first term (with minus sign) is negative; the second term, $-t\mu_t^* t_T$ is < 0 as μ_t^* and t_T are both < 0. In addition, since $b_T < 0$ the third term $tb_T\mu_t^* t_T$ is also < 0. As for the last term in the brackets, we know from optimum μ^* that $\mu_T^* > 0$, $\mu_t^* < 0$ and $\mu_{tT}^* > 0$ (equation 13) Moreover, $t_T < 0$ by Lemma 1. Thus, it is sufficient that $\mu_{TT}^* < 0$ in order that $Q < 0$ and second order condition be satisfied[23]. Note, however, that $\mu_{TT}^* < 0$ is *not* necessary. For example, a more relaxed condition such as $\mu_{TT}^*/\mu_T^* < -t_T/t$ will also satisfy second order condition.

8.4 Appendix 4: Proof of non-optimality of law enforcement if $t_L > 0$

The first order condition for L, if it existed, would derive from differentiating the government's utility function, in L. This would yield:

$$\frac{\partial U_G}{\partial L} = b(\mu^* + t\mu_t^*)t_L + bt\mu_L^* + b_L(\rho + t\mu^*)$$

Recall from Lemma 2 that $\mu^* + t\mu_t^* > 0$ under reasonable upper bound on tax rate. Also recall from that lemma that the case of $t_L > 0$ can only arise from $b_L > 0$. It follows that $\mu_L^* > 0$, above derivative U_G will always be increasing in L i.e., $\partial U_G/\partial L > 0$, and there is no optimum enforcement. Such a government would always benefit from maximum enforcement.

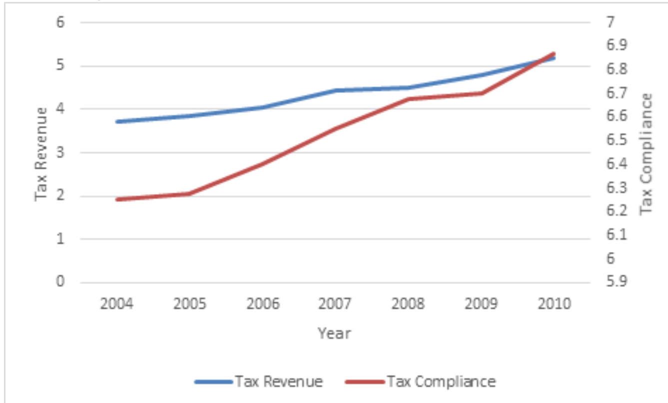
[23] For analytic simplicity, two simplifying linearity conditions are assumed: $\tau_{TT} = b_{TT} = 0$.

8.5 Appendix 5: Descriptive Statistics and Graph

Table A5-1. Descriptive Statistics, Variable Names and Variable Transformations

Variable Names as Shown in Tables	Actual Transformed Variable, if any, Used in Regressions	Summary Statistics						Variable Description	Source
		Obs.	Countries	Mean	Std. dev.	Min	Max		
Transparency (also labeled ri in tables 7 and 8)		5326	184	0.48	0.17	0.01	0.87	An Index of Transparency ranging from 0 to 1. (The index is linearly transformed without being compacted so that all values are positive)	Release of Information Index by Williams (2009 & 2011)
Tax Compliance		952	149	6.54	2.29	0	9.87	An index of tax compliance ranging from 0 to 10	Fraser Institute Annual Report (2010)
Tax Revenue	Taxes on Individuals per GDP (in percent)	2968	165	4.21	4.62	0	26.26	Gov. Tax revenue from Individuals. Include taxes on total income, capital gains and profit.	ICTD
Oil	Log of (1+Oil Value per GDP)	4743	167	0.06	0.13	0	1.82	Where oil value is the quantity of oil and gas extracted in a given year multiplied by the per-unit world price divided by GDP.	Ross (2013)
Oil Price	Log of Oil Price (const. 2000 US$)	5239	169	3.48	0.48	2.61	4.35	Where oil price is the Average annual per unit price of oil.	Ross (2013)
Oil Reserves	Log of Oil Reserves	1550	50	7.68	1.90	4.37	11.47	Where oil reserves is the estimated value of the fraction of recoverable crude oil from underground reservoirs with current petroleum extraction technologies.	Horn (2011) Giant Oil Discoveries dataset
GDP per capita	Log of GDP per capita (const. 2005 US$)	5480	199	8.01	1.64	3.91	11.98	GDP per capita in constant 2005 US$.	WDI
Gov. Cons	Gov't Consumption per GDP (in percent)	4976	186	16.65	8.19	1.38	164.70	Government consumption divided by GDP	WDI
Polity2		4669	165	1.68	7.24	-10	10	Updated Combined Polity Score; measuring on a scale from -10 to +10 the polity of a country	Polity2 entails a well-known "fix" to Polity, with "standardized scores, -66, -77, & -88, converted to conventional scores (in the range, -10 to +10): -66 (‘foreign interruption’) is treated as 'system missing;' -77 ('interregnum,' or anarchy), is converted to a 'neutral' (zero); -88 ('transition') is prorated across the span of transition" (Polity IV Project, 2012).
Inter. Taxes	Taxes international trade and Transactions per GDP (in %)	3880	182	3.10	3.65	-0.03	39.18	Taxes on international trade and transactions. Including import and export taxes	ICTD
Rule of Law		2396	203	-0.04	1.00	-2.67	2.00	An Index of rule of law ranging from -2.7 to 2.7. (When squared value is needed the index is first linearly transformed without being compacted so that all values are positive)	WDI

Figure A5-1 Tax Revenue and Tax Compliance



8.6 Appendix 6: Tax Enforcement and Rule of Law

We test whether tax enforcement is also a "strategic" variable (similar to transparency) in the sense that it enters citizens' compliance and government's transparency decisions within the framework of the game. No direct tax enforcement data is available to our knowledge. Thus we use the rule of law variable from the World Development Indicators (WDI) as a proxy. We recognize that rule of law is not the same as the *enforcement* of the law, in general, and of tax laws, in particular. Yet, it is reasonable to expect that having rule of law is necessary, if not sufficient, for its enforcement in the first place. Thus while the inclusion of rule of law may give some clue as to its enforcement, this is not a decisive clue. Yet, this is the best that can be done with the present data.

Including rule of law from WDI into the SEM equations reduced the number of observations drastically, by about one half. Given this measurement dilemma, we opted for a separate 5-year panel regression from 1996 to 2010 that smooths over missing observations and still gives a sense of time as well as cross-section.

For tax enforcement (or its proxy rule of law) to be a strategic variable between tax payers and the government, it must be that tax compliance μ is increasing but concave in L (hypothesis H2b). The corresponding empirical model to estimate is very similar to equation 27 but with rule of law either replacing or augmenting transparency:

$$(tax\ compliance)_{it} = a_4 + b_4(law\ and\ order)_{it} + c_4(law\ and\ order)^2_{it} + \\ d_4(trans)^*_{it} + e_4(trans^*)^2_{it} + f_4(G/Y)_{it} + \tilde{\mathbf{g}}'_4.\tilde{\mathbf{X}}_{it} + u_{4i} + v_{4t} + \varepsilon_{4it} \quad (34)$$

Compliance Measure and Estimation Results Results are reported in tables A6-1 testing the quadratic dependence of tax compliance on both L and T. Of the 9 columns, three include both rule of law *and* transparency, three include only rule of law, three include only transparency. Without exception, the transparency coefficients confirm the concavity of the results of table 1, but rule of law exhibits a significant convexity. Thus, it fails the optimality test as would be required by hypothesis 2b. This means that unlike transparency, rule of law does not appear to be a strategic policy tool of the government, to affect tax compliance.

Table A6-1. Testing optimum tax payer response to transparency & rule of law (panel regressions)

	Tax Revenue								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Transparency	9.491***	9.791***	11.45***	8.397***	9.196***	10.79***			
	(0.442)	(0.729)	(0.664)	(0.486)	(0.998)	(0.969)			
Transparency squared	-7.527***	-8.342***	-9.659***	-7.043***	-8.366***	-9.776***			
	(0.635)	(0.887)	(0.944)	(0.557)	(1.022)	(1.107)			
Norm. Rule of law	-5.530***	-3.807**	-5.684***				-2.405**	-1.884	-3.576**
	(1.116)	(1.845)	(1.673)				(1.024)	(1.810)	(1.555)
Norm. Rule of law sq.	8.901***	7.529***	10.10***				6.089***	6.136***	8.593***
	(0.750)	(1.114)	(0.772)				(0.630)	(1.031)	(0.589)
GDP PC	0.611***	0.724***	0.721***	1.048***	1.146***	1.240***	0.692***	0.735***	0.728***
	(0.00673)	(0.0261)	(0.0184)	(0.00901)	(0.0570)	(0.0700)	(0.0297)	(0.00472)	(0.0172)
Gov. Cons	0.0487***	0.0485***	0.0611***	0.0530***	0.0525***	0.0630***	0.0432***	0.0424***	0.0553***
	(0.000291)	(0.000675)	(0.00261)	(0.00124)	(0.00248)	(0.00364)	(0.000558)	(0.000404)	(0.00302)
polity2		-0.0136	-0.0147		0.00640	0.00829		-0.00771	-0.00882
		(0.0286)	(0.0280)		(0.0215)	(0.0200)		(0.0281)	(0.0285)
Inter. Taxes			0.0691***			0.0547***			0.0781***
			(0.0208)			(0.0174)			(0.0204)
_cons	-4.943***	-5.934***	-6.626***	-7.876***	-8.494***	-9.950***	-3.496***	-3.867***	-4.116***
	(0.152)	(0.575)	(0.580)	(0.0751)	(0.321)	(0.439)	(0.0561)	(0.330)	(0.273)
Countries	135	124	122	135	124	122	140	126	124
Observations	368	338	326	368	338	326	381	343	331

*,**,***: significant at 10, 5 and 1 percent. Panel is based on 5-year averages. Standard errors are in parentheses. Errors are corrected for clustering around the MENA countries. Rule of law is normalized to fall between 0 and 1 so that it can be squared. First 3 columns include both transparency & rule of law, columns 4-6 include only transparency, columns 7 to 9 include only rule of law.

Failing the optimality test, we ask whether rule of law can still play an important role, for example, by providing the legal framework for tax enforcement, or whether it may still influence the degree to which transparency itself is allowed by the government. Thus, we test whether controlling for the rule of law impacts the robustness of the transparency results, as well as whether rule of law has a direct impact on transparency.

First, since the datasets are different old equations must be re-estimated even if the variables are the same. This means that the old reduced form equation 25 on the association between oil wealth and tax revenues has to be re-estimated with the modified dataset (the empirical equation which is identical to 25 is not repeated here).[24] Second, the structural equations 26 and 27 are re-estimated with the rule of law variable added, as follows:

$$(trans^*)_{it} = a_5 + b_5(oil\ value/Y)_{it} + c_5(rule\ of\ law)_{it} + \vec{D}_5.\vec{X}_{it} + u_{5i} + v_{5t} + \varepsilon_{5it} \quad (35)$$

$$(tax\ compliance)_{it} = a_6 + b_6(trans)^*_{it} + c_6(trans^*)^2_{it} + d_6(rule\ of\ law)_{it} + e_6(G/Y)_{it} + \vec{F}_6.\vec{X}_{it} + u_{6i} + v_{6t} + \varepsilon_{6it} \quad (36)$$

The model is agnostic about whether increased enforcement is a substitute or a compliment for government transparency, as we presented arguments on both sides, i.e.; $c_5 \gtreqless 0$, so that the sign of the coefficient of rule of law will be determined empirically. Besides the signs of other coefficients already discussed, one would expect $d_6 > 0$, that is, rule of law (a proxy for enforcement) to increase with tax compliance.

Results are reported in the three panels of table A6-2. Results from the first panel confirm the previous SEM results of oil's negative effect on income tax revenues. Next, controlling for the rule of law, the results from the second panel confirm the previous SEM findings of negative and significant link between oil and transparency. Importantly, in this group rule of law is *not* robust in the sense that it is significant only when Polity2 is not included (while the latter remains highly significant). This suggests that rule of law is a more *structural* variable and more likely associated with being more democratic, rather than a policy tool to extract more taxes, pointing to its more or less *exogenous* nature. The last panel of the table, i.e., columns 10-12, show that controlling for the rule of law, transparency still influences tax compliance positively and concavely consistent with the strategic behavior of optimizing citizens and government, while rule of law itself positively and significantly influences tax compliance.

[24] Note that it is not possible to include oil price as instrument for oil values in this panel estimation, given the fact that the 5-year estimation method produces only 3 price points.

Table A6-2. Testing optimum tax payer response to transparency with rule of law as control (panel regressions)

	Direct effect of Oil on taxes			Indirect path of oil on taxes through Transparency								
				Path 1: From Oil to Transparency						Path 2: From Transparency to Taxes		
	Taxes			Transparency						Taxes		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Oil Value	-1.082***	-1.005***	-4.919***	-0.179**	-0.127***	-0.136***	-0.226***	-0.190***	-0.137***			
	(0.366)	(0.359)	(0.639)	(0.0748)	(0.0116)	(0.0371)	(0.0360)	(0.0396)	(0.00613)			
GDP PC	1.175***	1.162***	1.267***	0.0491***	0.0461***	0.0450***	0.0444**	0.0423**	0.0425***	0.770***	0.887***	0.935***
	(0.00380)	(0.0477)	(0.0449)	(0.0149)	(0.0128)	(0.0160)	(0.0176)	(0.0171)	(0.0141)	(0.0301)	(0.0018)	(0.0019)
Gov. Cons	0.0442***	0.0440***	0.0785***			-0.00137***		-0.00118***	-0.00161***	0.0450***	0.0450***	0.0582***
	(0.00334)	(0.00386)	(0.0108)			(0.000459)		(0.000355)	(0.000440)	(0.0000734)	(0.00078)	(0.0023)
polity2		0.0105	0.00396		0.00803***				0.00721***		-0.0127	-0.0143
		(0.0201)	(0.0204)		(0.00134)				(0.000906)		(0.0281)	(0.0273)
Inter. Taxes			0.0591***				-0.00563*	-0.00399	-0.00183			0.0529**
			(0.0167)				(0.00298)	(0.00283)	(0.00268)			(0.0215)
Rule of law				0.0222**	0.00789	0.0320**	0.0216	0.0281**	0.0118	0.758***	0.798***	0.944***
				(0.0104)	(0.00945)	(0.0140)	(0.0138)	(0.0133)	(0.0110)	(0.0995)	(0.173)	(0.216)
Transparency										7.459***	7.764***	8.451***
										(0.498)	(0.934)	(0.619)
Transparency squared										-6.294***	-7.113***	-7.793***
										(0.645)	(1.002)	(0.874)
_cons	-6.286***	-6.161***	-7.437***	0.168	0.164*	0.222*	0.224	0.255*	0.231*	-5.248***	-5.814***	-6.659***
	(0.135)	(0.126)	(0.259)	(0.113)	(0.0985)	(0.128)	(0.142)	(0.142)	(0.123)	(0.408)	(0.237)	(0.268)
Countries	131	125	123	163	155	161	156	155	146	135	124	122
Observations	357	341	329	486	461	475	437	430	407	368	338	326

*,**,***: significant at 10, 5 and 1 percent. Standard errors are in parentheses. Panel is based on 5-year averages. Errors are corrected for clustering around the MENA countries. First 3 columns capture the direct effect of oil on tax compliance, the next 9 capture it's indirect effect through transparency: columns 4-9 capture the effect of oil on transparency (path 1) and columns 10-12 capture the effect of transparency on tax compliance (path 2).