Working Paper 23

Taxation, Non-Tax Revenue and Democracy: New Evidence Using New Cross-Country Data

Wilson Prichard, Paola Salardi and Paul Segal

September 2014

www.ictd.ac

UKaid from the Department for International Development

Norad

ICTD Working Paper 23

Taxation, Non-Tax Revenue and Democracy: New Evidence Using New Cross-Country Data

Wilson Prichard, Paola Salardi and Paul Segal

September 2014

Taxation, Non-Tax Revenue and Democracy: New Evidence Using New Cross-Country Data
Wilson Prichard, Paola Salardi and Paul Segal
ICTD Working Paper 23
First published by the Institute of Development Studies in September 2014

ISBN: 978-1-78118-173-7

A catalogue record for this publication is available from the British Library.

Available from:
The International Centre for Tax and Development
at the Institute of Development Studies,
Brighton BN1 9RE, UK
Tel: +44 (0) 1273 606261 Fax: +44 (0) 1273 621202
E-mail: info@ictd.ac.uk
Web: www.ictd/en/publications

Taxation, Non-Tax Revenue and Democracy: New Evidence Using New Cross-Country Data

Wilson Prichard, Paola Salardi and Paul Segal

Summary

A large body of cross-country econometric research has investigated the possibility of a political resource curse, by which access to extensive natural resources reduces the extent of democracy and accountability. However, this literature has been plagued by problematic data and correspondingly inappropriate model specification. Dominant theories of the political resource curse focus on the political consequences of differences in the composition of government revenue, with greater reliance on non-tax revenue undermining democracy. However, most studies do not actually test this relationship: owing to the poor quality of government revenue data, they have focused instead on the impact of total resource income on democracy – a reasonable, but imperfect, approximation of the actual theory. Meanwhile, the robustness of those few studies that have focused on government revenue specifically is undermined by poor data quality. We overcome this problem by drawing on the newly-created ICTD *Government Revenue Dataset*, which dramatically improves the quality of existing data and allows us to test directly the connection between the composition of government revenue and democracy. Employing this new data we re-test the most compelling econometric approaches from the existing literature, finding support for the existence of a political resource curse.

Keywords: taxation, non-tax revenue, accountability, resource curse.

Wilson Prichard is Assistant Professor in Political Science and Global Affairs at the University of Toronto, and Research Co-Director of the International Centre for Tax and Development.

Paola Salardi is Post-Doctoral Fellow at the Munk School of Global Affairs, University of Toronto.

Paul Segal is Senior Lecturer in Emerging Economies and International Development at King's College, London.

Contents

Acknowledgements

Thanks to the International Centre for Tax and Development for funding this work, and for the ICTD *Government Revenue Dataset* on which it is based. Thanks to Stephen Haber and Victor Menaldo for sharing data from their research, and to David Wiens, Paul Poast and William Clark for sharing information and Stata code from their work. Finally, thanks also to Markus Eberhardt for useful comments, and to Michael Ross for very helpful feedback and advice throughout the process.

Acronyms

CCE-MG	Mean Group-Common Correlated Effects
ECM	Error Correction Model
EECA	Eastern Europe and Central Asia
FE	Fixed Effects
GDP	Gross Domestic Product
GMM	Generalised Method of Moments
ICTD	International Centre for Tax and Development
MENA	Middle East and North Africa
MG	Mean Group
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
OVB	Omitted Variables Bias
PRIO	Peace Research Institute Oslo
UCDP	Uppsala Conflict Data Program
WDI	World Development Indicators

Introduction

The past decade has witnessed a regular stream of cross-country econometric research into the political resource curse, which proposes that governments that rely more heavily on revenue from natural resources[1] are less likely to be democratic and accountable to their citizens. Most research has reported a negative relationship between natural resource production and democracy (Andersen and Ross 2014; Aslaksen 2010; Ross 2001). However, a handful of studies have questioned the robustness of these findings (Haber and Menaldo 2011; Alexeev and Conrad 2009; Dunning 2008; Herb 2005), while several studies have focused instead on the impact of resource wealth on regime durability rather than levels of democracy per se (Smith 2004; Morrison 2009; Wright et al. 2013).

These studies have included the progressive implementation of more sophisticated econometric methods, as well as the development of improved data on the annual value of resource production across countries – referred to here as 'resource income'. However, while these studies have made valuable contributions, they have often paid insufficient attention to basic questions about data quality, model specification and the sensitivity of results to alternative econometric models. We correspondingly seek to go back to basics by drawing on better data and a more inclusive set of econometric tests in order to present more theoretically and empirically robust results.

Central to our approach is an effort to specify and test the theoretical relationship of interest more clearly. The most prominent theories of the political resource curse are ultimately theories about the connections between the *composition of government revenue* and political outcomes. Government reliance on non-tax revenue, primarily from natural resources, is expected to reduce the quality of democracy and accountability by weakening state-society links, facilitating government investments in patronage and repression, and driving expanded political corruption. Meanwhile, reliance on tax revenue may have a conversely positive impact on governance by providing the state with stronger incentives to bargain with their citizens over how public revenue is used (Moore 1998).

However, while theory focuses on the extent of government *revenue* from tax and non-tax sources, most existing studies have, in fact, tested the relationship between *resource income* and democracy (Wiens et al. forthcoming; Haber and Menaldo 2011; Aslaksen 2010). While resource revenue and resource income are, of course, closely related, the precise relationship varies significantly across countries, while tests that focus on resource income take no account of differences across counties in levels of non-resource taxation. The implication is that most existing studies have tested a reasonable, but ultimately imperfect, approximation of the theoretical relationship of interest – a fact implicitly or explicitly acknowledged in several recent studies (Haber and Menaldo 2011; Wiens et al. forthcoming). While a handful of studies have attempted to focus directly on the composition of government revenue, the success of these efforts has been undermined by their reliance on incomplete and problematic data that raises major questions about the robustness of the results (Ross 2004; Morrison 2009).[2]

The central contribution of this paper thus lies in drawing on the newly-created International Centre for Tax and Development *Government Revenue Dataset* (ICTD GRD), which allows us to construct measures of the composition of government revenue that are relatively complete, accurate and theoretically appropriate. The new dataset allows us to distinguish

[1] In this context, natural resources generally refers, more specifically, to non-renewable natural resources that generate significant economic rent – primarily oil production and, to a lesser extent, mining revenue.

[2] See Prichard et al. (2014) for a fuller discussion of these data limitations.

consistently between tax revenue, which is defined *exclusive* of natural resource revenue,[3] and non-tax revenue, which is dominated by natural resource revenue, while also including other forms of non-tax revenue. From these two components we then construct a composite measure of *tax reliance*, which measures the share of government revenue secured through non-resource taxation. The tax revenue and non-tax revenue variables closely mirror the core analytical distinction of interest: between 'earned' revenue, which is expected to lead to increased accountability through bargaining between government and citizens, and 'unearned' revenue, which is expected to reinforce autocracy (Moore 1998).[4]

With this data in hand, we test the relationship between government revenue and democracy by employing several of the most compelling econometric approaches from the existing literature: Generalised Method of Moments (GMM), Mean Group estimators (MG), Random and Fixed Effects (FE) Logit and Error Correction Models (ECMs). This wide-ranging approach is a response to the proliferation of alternative econometric strategies in the literature, and corresponding concerns about comparability and robustness. In the presence of methodological uncertainty, we view the transparent implementation of several complementary, but individually imperfect, estimation strategies as the most compelling approach.

We find strong evidence of a political resource curse across all the econometric estimation strategies. Our composite measure of tax reliance is positively associated with democracy in all specifications, and achieves statistical significance in the strong majority of them. When we disaggregate our measure of tax reliance into its component parts – total tax revenue as a share of GDP and total non-tax revenue as a share of GDP – we find consistent evidence that increases in non-tax revenue have a negative impact on democracy. Critically, the magnitude of these effects is substantial, with countries securing the majority of income from non-tax revenue dramatically less likely to be democratic. The results linking increased taxation to democracy are significantly more mixed – though sometimes positive and significant – which we attribute to the greater complexity of the hypothesised causal relationship. Finally, while the results offer clear support for the existence of a political resource curse, we also intentionally highlight some sensitivity of particular results to changes in model specification and estimation strategies. This echoes Acemoglu et al. (2008), among others, who highlight the difficulty of estimating complex political relationships using cross-country econometric methods. This reinforces the importance of carefully exploring the sensitivity of key results.

The paper proceeds in seven parts. The next section presents a brief review of the relevant literature, focusing on going back to basics in ensuring appropriate data to test the theoretical propositions of interest directly. The second section presents the new data and the construction of the revenue variables. The third section presents the empirical strategy, reviewing the rationale for a range of alternative econometric models. The fourth section presents the core results, and the fifth section presents robustness checks. The sixth section offers a discussion of the core results, including how best to interpret the relatively ambiguous results linking tax collection and democracy. The final section concludes.

---

[3] As described in greater detail below, earlier government revenue datasets frequently conflate resource and non-resource sources of taxation, despite them being analytically distinct in relation to theories of the resource curse.

[4] While the data does not allow for further disaggregating non-tax revenue into its natural resource and non-resource components, we view this is a minor problem: most of the variation in non-tax revenue is driven by natural resource revenue, many types of non-tax revenue are analytically similar to natural resource revenue, and, most critically, the inclusion of any 'tax-like' elements in the non-tax category will bias our analysis *against* finding a significant relationship between non-tax revenue and democracy.

1 Going back to basics: model specification, data and methods

Studies of the political resource curse propose that countries that rely more heavily on revenue from natural resources are less likely to be democratic than those that rely more heavily on tax revenue. This is most commonly attributed to two related, but distinct, causal mechanisms. On one hand, governments that rely heavily on tax revenue are expected to be forced to bargain with mobile taxpayers, conceding greater accountability and democracy in exchange for tax revenue (Levi 1988). On the other hand, governments with access to captive sources of non-tax revenue, primarily from natural resources, will be empowered to resist political opposition through repression or through expanded public spending and patronage. Ulfelder (2007) and others refer to these alternative processes as the 'demand-side' and 'supply-side' of the political resource curse respectively.

While these ideas were initially grounded in country-level studies of politics in resource-rich states, the past decade in particular has seen the rapid proliferation of cross-country econometric tests. Beginning with Ross (2001), most of these studies report the expected negative relationship between dependence on natural resources and the level of democracy. These studies variously measure natural resource dependence by the share of natural resources in exports, the value of natural resource production as a share of GDP, the annual value of natural resource production in absolute terms, and the overall stock of natural resources (Wantchekon 2002; Jensen and Wantchekon 2004; Tsui 2010; Aslaksen 2010; Ramsay 2011; Ross 2012; Andersen and Ross 2014). Smith (2004) and Morrison (2009) make the slightly different argument that natural resource wealth may increase regime durability in both autocracies and democracies, though Andersen and Aslaksen (2013) and Wiens et al. (forthcoming) contend that natural resources increase authoritarian durability, while having little effect on democracies.

There have been several challenges to these findings. Herb (2005) reports that the negative association between resource wealth and accountability disappears if GDP – which is expected to be an important determinant of democracy – is calculated exclusive of natural resource wealth. Dunning (2008) argues that the political resource curse does not hold where levels of inequality are high, as resource wealth may reduce opposition to democracy among elites by reducing the need for redistributive taxation. The most high profile challenge to the existence of a political resource curse has come from Haber and Menaldo (2011), who constructed an extensive new cross-country dataset of national resource income dating back to 1800, and report no consistent support for the existence of a political resource curse. However, these new results quickly elicited a response from Andersen and Ross (2014), who contend – using the same data – that evidence of the political resource curse reappears if analysis is restricted to the period after 1980, and if the impact of resource wealth is allowed to occur over a period of three or five years.

While recent years have thus seen the proliferation of econometric studies, employing progressively more sophisticated econometric techniques, this increasing complexity has overshadowed persistent problems related to data and model specification. We thus emphasise a need to go back to basics in ensuring a solid foundation for applying progressively more complex econometric approaches. The most prominent theories of the political resource curse focus on the political implications of differences in the *composition of government revenue*, rather than on the implications of the *total value of resource production* (resource income). A government that is able to capture extensive revenue from the exploitation of natural resources within its borders will be able to spend that revenue on patronage and repression, while it will have less need to bargain with taxpayers. Critically,

the same effects will *not* hold true if there is extensive resource production, but only a small proportion of resource rent is captured by the government as resource revenue.[5]

However, despite this theoretical focus on the composition of government revenue, most studies have focused on measures of the total value of natural resource production – resource income (Dunning 2008; Ramsay 2011; Ross 2012; Wiens et al. forthcoming; Aslaksen 2010; Haber and Menaldo 2011; Jensen and Wantchekon 2004; Ross 2001; Smith 2004; Ulfelder 2007). While the level of natural resource income is an important determinant of government resource revenue, this relationship is highly variable across countries and over time.[6] Tellingly, Wiens et al. (forthcoming) explicitly note that a measure of government revenue from natural resources would be theoretically preferable to their reliance on a measure of resource income as a share of GDP.

However, studies relying on measures of government revenue have remained rare, and their analysis somewhat unconvincing, owing to the inadequacy of cross-country government revenue data. Haber and Menaldo (2011) employ the best measure of fiscal reliance, which is measured as the share of natural resource revenue in total government revenue, but employ data for a sample of only eighteen resource-dependent countries. Meanwhile, earlier studies by Ross (2001), Herb (2005) and Morrison (2009) have relied on data from the IMF and World Bank, both of which are highly incomplete and widely understood to include important inaccuracies for analytical purposes, thus casting doubt on the robustness of those results (Prichard et al. 2014). Against this background, the central contribution of this study lies in introducing the ICTD *Government Revenue Dataset*, which contains more complete and accurate data on the composition of government revenue. This allows us to test the impact of changes in the composition of government revenue on levels of democracy and accountability, thus capturing the core theoretical propositions underpinning the political resource curse.

Finally, a brief note is necessary about our reliance on measures of the relative, rather than absolute, extent of revenue from non-tax sources. In recent years a variety of studies of the political and economic impacts of natural resources have focused on the absolute level of natural resource income, rather than measuring resource dependence as a share of GDP or of total government revenue. This decision has been grounded in plausible methodological concerns about endogeneity: low-income countries may be highly dependent on resource revenue not because the resource sector is large, but because the rest of the economy is small and tax collection is weak (Dunning 2008; Brunnschweiler and Bulte 2008).

However, while this implies that relative measures of resource dependence may be endogenous with respect to measures of economic performance (such as long-run economic growth), it applies to accountability and democracy only if accountability and democracy are themselves determined by economic performance. This is at best contentious, and we find, like Acemoglu et al. (2008), that this is not the case. Moreover, the solution is worse than the original problem: employing a measure of the absolute, rather than the relative, level of resource income fundamentally distorts the research question.[7] Theories of the political resource curse are explicitly arguments about the extent of non-tax government revenue

---

5 Of course, the value of resource production may still be relevant to alternative development outcomes. High levels of resource income may undermine economic growth or increase the risk of conflict, irrespective of the ability of the government to collect revenue from resource production (e.g. van der Ploeg 2011; Humphreys 2005). However, these outcomes are relatively distinct from the causal arguments that are the core focus here and in the literature on the political resource curse.

6 These differences may arise, among others, from differences in the characteristics of the resources and costs of extraction, in the ownership structure of resource companies, in national tax policies, or in the effectiveness of tax enforcement. Illustratively, Haber and Menaldo (2011) construct a measure of 'fiscal reliance' – that is, the share of nature resource revenue in total government revenue – for a subset of 18 countries, and it has a correlation of only 0.65 with levels of resource income in the same countries (Wiens et al. forthcoming).

7 It is equally worth noting that reliance on the absolute size of the resource sector does not eliminate the risk of endogeneity, as resource discoveries and exploitation are also more likely in more developed countries.

relative *to* other sources of revenue, and the overall size of the economy (Wiens et al. forthcoming). The resource curse is premised on the idea that access to resource wealth will alter the incentives and behaviour of political leaders and citizens. The size of these incentive effects should, in turn, depend on the relative, rather than absolute, size of resource revenue.

2 Data and variables

The ICTD *Government Revenue Dataset* has several key advantages relative to earlier government revenue datasets, which are summarised here and described in additional detail in Prichard et al. (2014). Previous research has relied primarily on data from the IMF and World Bank, both of which are plagued by missing data (Ross 2004; Morrison 2009), or on datasets that offer relatively more complete data coverage for only a very limited subset of countries (Haber and Menaldo 2011; Baskaran and Bigsten 2013). By contrast, the ICTD GRD covers 188 countries and a total of 3,342 country-year observations during the period 1990-2010. This includes almost 70 per cent more observations for developing countries than the IMF *Government Finance Statistics* during the same period.[8] Whereas most previous datasets have focused exclusively on central government data, the ICTD GRD employs general government data where available, in order to avoid underestimating revenue collection in federal states.[9] The analysis here focuses on data beginning in 1990, as it offers the most complete data coverage, though the results are robust to employing earlier but less complete data as well.

At least as important are substantial improvements in data accuracy, particularly in relation to the classification of tax and non-tax revenue.[10] The key analytical distinction for our purposes is between what Moore (1998) terms earned income, which is collected from citizens and expected to prompt bargaining and accountability, and unearned income, which is collected from narrow and relatively captive sources, and is expected to make governments more autonomous from citizens. Broadly speaking, most domestic taxes can be thought of as earned revenue, while revenue collected from natural resource operations are more akin to unearned revenue.

Unfortunately, this analytical distinction is frequently violated within existing international datasets. In general, governments collect revenue from resource companies through two broad channels, corporate taxes and various types of royalties, alongside a more widely varied range of smaller taxes and levies.[11] Motivated by an accounting logic, international databases frequently group corporate taxes on resource firms along with all other types of

[8] This excludes a small set of observations that are excluded from the analysis owing to: (a) an inability to deal effectively with resource revenue; (b) highly irregular data resulting, most often, from hyperinflation; and (c) data for years in which there were exceptionally large fluctuations in the composition of government revenue during transitions from centrally-planned economies. These exclusions are described in detail in Prichard et al. (2014), while the results reported here are in any case robust to including countries transitioning from centrally-planned economies.

[9] In principle this approach risks creating a new set of distortions by ignoring sub-national revenue sources in highly centralised countries. However, in practice sub-national tax revenue generally accounts for less than 5% of total tax revenue, and often a significantly smaller proportion, in highly centralised (generally low-income) countries (Bird 2011). As such, any distortion of this type is likely to be negligible, whereas the distortions implied by relying solely on central government data are potentially large. Indeed, we find that the results to follow are more robust, and more in line with expectations, when relying on the dataset that merges central and general government data.

[10] Aside from improvements in the classification of tax and non-tax revenue, greater accuracy results from systematic comparison of alternative sources to identify problematic observations and from standardising the recording of social security contributions across countries and over time (Prichard et al. 2014).

[11] For the sake of simplicity, this term is used very broadly to include royalties as such, profit sharing, the auction values of exploration and extraction rights, and any other legal mechanism of transferring natural resource rents to the government.

domestic tax revenue, while royalties are classified as non-tax revenue.[12] However, from a political economy perspective, taxes on resource firms are no different from royalties, and both are analytically distinct from other types of tax revenue. The result is that the 'taxes' and 'non-tax revenue' categories in international databases often do not correspond to the theoretical quantities of interest, earned and unearned income.

Against this background, the key innovation of the ICTD GRD lies in consistently distinguishing the resource and non-resource components of tax revenue.[13] This makes it possible to construct variables that are more analytically accurate. The first variable captures total tax collection as a share of GDP (*tottax*) and, critically, is exclusive of resource revenue, thus making it a more consistent proxy for earned income. The second variable captures total non-tax revenue as a share of GDP (*totnontax*), and is a proxy for unearned income. This variable incorporates all revenue from natural resource production, be it formally labelled royalties, taxes or any alternative, as well as additional non-tax revenue, including revenue from state enterprises and licences. We exclude aid grants from non-tax revenue owing both to problems of data availability and accuracy, and reflecting significant evidence elsewhere that aid affects government incentives differently to natural resource rents.[14] Finally, we are able to combine these two variables to construct a measure of tax reliance (*tax_rel*), which measures the share of non-resource tax revenue in total government revenue.[15] All our econometric tests first look at the impact of tax reliance (*tax_rel*) on democracy, following earlier studies (Ross 2004). We then disaggregate tax reliance into its component parts, *tottax* and *totnontax*, in order to attempt to disentangle the respective roles of taxation and non-tax revenue in shaping outcomes.[16]

It would, in principle, be desirable to disaggregate non-tax revenue into its resource and non-resource components, in order to focus attention separately on the political implications of, respectively, government revenue from natural resources and all other sources of non-tax revenue. In practice, however, available data sources to not allow for making this distinction consistently. This does not call any of the core results into question, but does result in some ambiguity about whether our results are capturing the impact on democracy of natural resource revenue specifically, or of non-tax revenue more generally.

At a minimum it is clear that our results are driven primarily by natural resource revenue, which comprises the vast majority of variation in total non-tax revenue. All the countries in the analysis that collect more than 10 per cent of GDP in non-tax revenue are major resource producers, among which non-tax revenue averages 25 per cent of GDP and reaches as high as 60 per cent of GDP. The impact of other sources of non-tax revenue is, by contrast, more ambiguous. Several of the non-resource components of non-tax revenue in developing

---

12 This is not universally the case, as taxes on resource firms are sometimes classified as non-tax revenue, while classification is more unpredictable and varied in countries in which resource extraction is controlled by state-owned firms.

13 This process draws primarily on IMF Article IV reports. Occasional country-year observations for which this distinction cannot be successfully drawn are excluded from the analysis.

14 With respect to data, it is frequently impossible to distinguish 'on budget' grants – which would be the quantity of interest – from other types of aid, making it impossible to construct a truly consistent and comparable measure of aid revenue across countries. Analytically, aid is expected to have distinct governance implications owing to the fact that it is frequently targeted for specific purposes and is tied to external monitoring, conditionality and technical assistance (Collier 2006). Empirically, this intuition is confirmed by evidence that natural resource revenue leads to sharply reduced domestic tax collection, while the same is not true of aid grants (Bornhorst et al. 2008; Morrissey et al. 2014). While Morrison (2009) includes aid revenue in his measure of non-tax revenue, we are sceptical of both the precision of the data and the analytical appropriateness of grouping these alternative revenue sources.

15 Mathematically, *tax_rel* = *tottax* / (*tottax*+*totnontax*)

16 In recent years it has been increasingly recognised that GDP is substantially underestimated in many developing countries owing to infrequent GDP rebasing, with a resultant overestimation of tax and revenue measured as a share of GDP (Jerven 2013; Prichard et al. 2014). To some extent this offers an argument in favour of the *tax_rel* variable, as it is unaffected by any potential underestimation of the results. At a minimum, *tax_rel* offers an important robustness check in light of concerns about GDP figures. The results employing *tottax* and *totnontax* could potentially be affected by underestimation of GDP, though reliance on within country estimators reduces this risk.

countries share at least some features associated with unearned revenue, most notably from state investment funds, state monopolies and enterprises, or revenue from fishing or forestry licences.[17] However, other revenue sources included in non-tax revenue involve at least some element of bargaining with citizens, as with user fees. It is possible that non-resource types of non-tax revenue have a negative impact on democracy similar to that of resource revenue, but we cannot isolate this relationship. With that said, the inclusion of more 'tax-like' elements of non-tax revenue is expected to bias our results *against* finding evidence of a political resource curse, thus eliminating any concern that the conflation of different types of non-tax revenue is leading to misleadingly significant results.[18]

Our primary dependent variable is democracy, which is constructed by employing the *Polity2* measure of democracy from the *Polity IV* dataset, and normalising it to the range 0 to 100 (Marshall et al. 2012). This follows the dominant trend in the literature, as the polity measure of procedural democracy is reliable and readily available, while democracy is a useful proxy for the broader quality of accountability and governance. In most of our results we adopt a similar set of control variables to those employed in recent work by Haber and Menaldo (2011). Despite sceptical recent econometric evidence (Acemoglu et al. 2008), the most widely-cited predictor of accountability remains national income per capita, and we correspondingly include a measure of log GDP per capita (*lgdp*), drawn from the *World Economic Outlook* database, in all our core regression results.[19] We additionally control for population size (*pop*, from the *World Development Indicators* (WDI)), the occurrence of violent conflict (*Civil_War*, which we draw from Haber and Menaldo [2011] and extend to 2010 using the UCDP/PRIO *Armed Conflict Dataset*) and the regional diffusion of democracy (*Regional_Dem_Diffuse*, constructed from *Polity2* data).[20] We also include a measure of annual economic growth per capita (*growthpc*, from WDI) in our logit regressions, in order to mirror the controls employed by Wiens et al. (forthcoming).

Figures 1, 2 and 3 present simple correlations between average values over the period 1990-2010 of our dependent variable, democracy, and our three alternative revenue measures: *Tax_rel* (tax reliance), *Tottax* (total tax revenue) and *Totnontax* (total non-tax revenue). The figures are based on a final total of 2,866 observations across 159 countries, reflecting the overlapping data coverage offered by the dependent and independent variables.

---

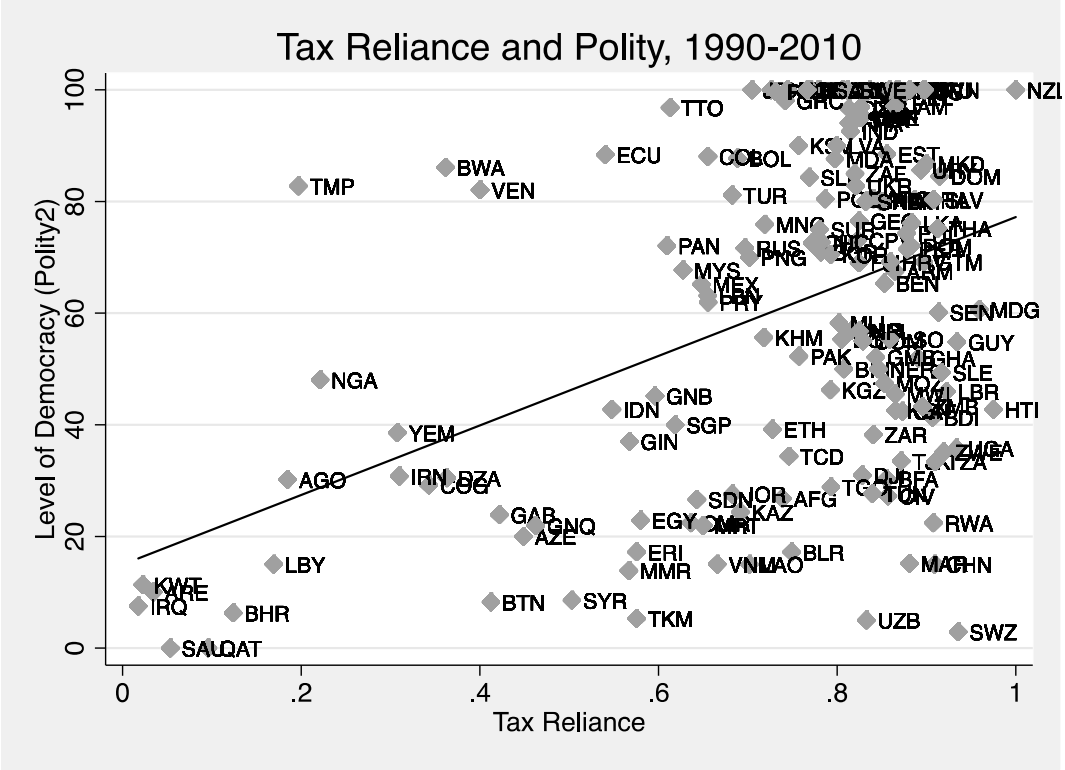
17 Morrison (2009) has previously argued, though on the basis of questionable empirical data, that the different components of non-tax revenue are likely to have similar implications for regime stability.

18 Illustratively, the highest levels of non-tax revenue in the dataset, other than in oil-producing countries, occur in OECD countries.

19 We rely on the WEO GDP figure because it is the same figure employed in calculating the shares of government revenue in GDP in the ICTD dataset, as described in Prichard et al. (2014).

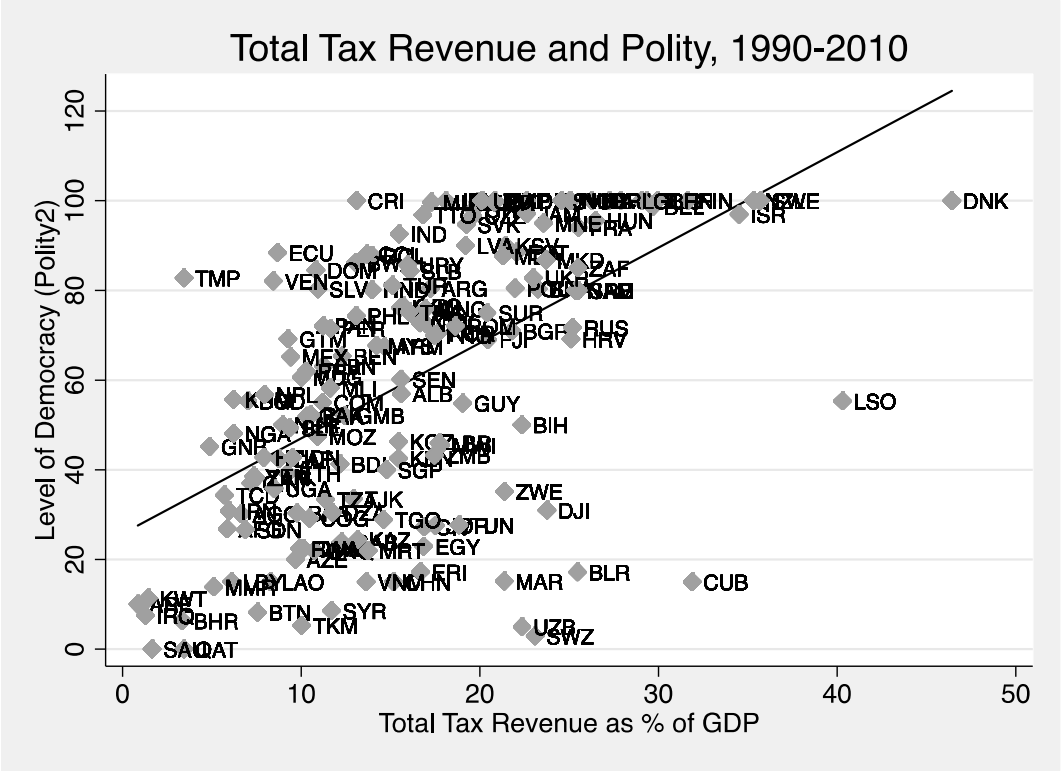
20 Haber and Menaldo (2011) also employ a control for global democratic diffusion, but we do not include it in our core results as it should be captured by our time dummies and is never significant when we experiment with its inclusion.

Figure 1 Relationship between tax reliance and polity using country averages, 1990-2010



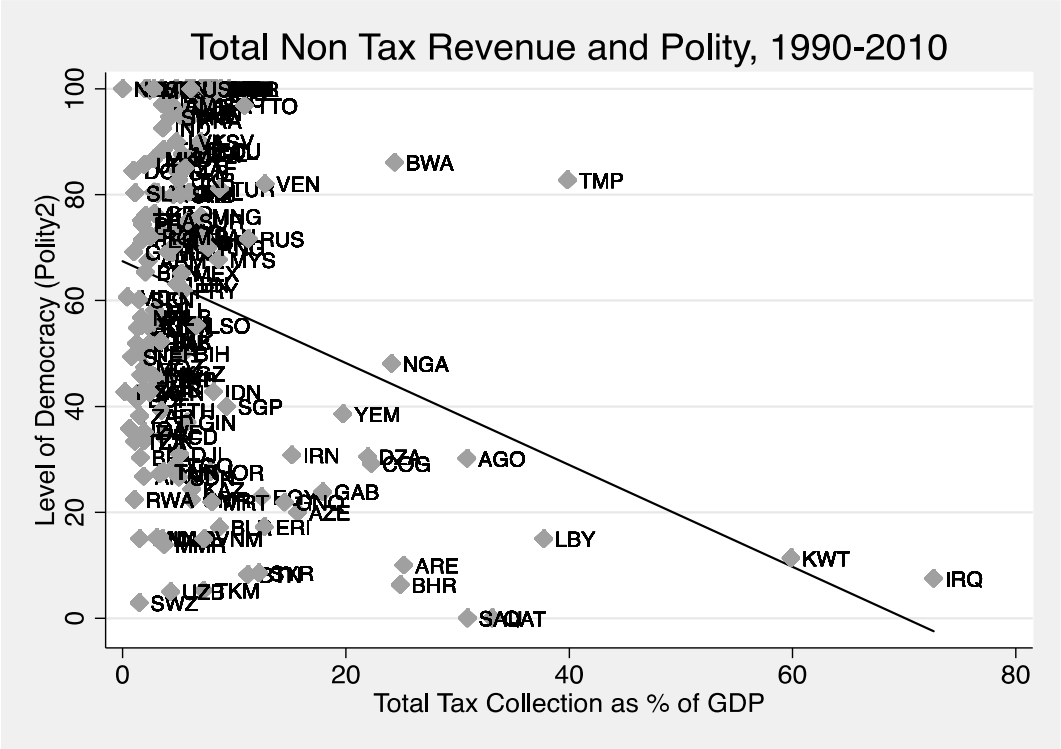
Source: Authors' computations using ICTD GRD (2014). Data from 1990, using period averages for each country in the dataset.

Figure 2 Relationship between total tax revenue and polity using country averages, 1990-2010



Source: Authors' computations using ICTD GRD (2014). Data from 1990, using period averages for each country in the dataset.

Figure 3: Polity vs. total non-tax revenue



Source: Authors' computations using ICTD GRD (2014). Data from 1990, using period averages for each country in the dataset.

The broad patterns in the descriptive data are as expected: countries that are more reliant on tax revenue and less reliant on non-tax revenue tend to be more accountable. Particularly striking are patterns among resource producers. Looking at Figure 3, there are eighteen countries[21] that have collected an average of at least 13 per cent of GDP in non-tax revenue. If we define democracy very liberally, as a country with a score of at least 50 (out of 100) on our adjusted polity variable,[22] only two countries in the group – Botswana and Timor Leste – have, on average, been democracies during the same period, as compared to half of all developing countries for which we have data. Focusing on Figure 1, the pattern is even more stark: of twenty-nine countries[23] with levels of tax reliance below 60 per cent – that is, countries for which less than 60 per cent of total government revenue comes from non-resource taxes – only three have, on average, met even the very relaxed definition of democracy above (Venezuela, Botswana and Timor Leste). Of course, these are merely indicative cross-country correlations, and we turn now to the econometric analysis in order to gain a more robust picture of these relationships.

3 Empirical strategy

Seeking to move beyond simple correlations, we estimate several alternative econometric models that draw on the existing literature: a System-GMM panel estimator, a Mean Group estimator to exploit the time series dimension of our macro panel more fully, and, finally, random and fixed effects logit models designed to focus attention on the likelihood of

[21] Angola, UAE, Azerbaijan, Bahrain, Botswana, Republic of the Congo, Algeria, Gabon, Equatorial Guinea, Iran, Iraq, Kuwait, Libya, Nigeria, Qatar, Saudi Arabia, Timor-Leste and Yemen.

[22] Equivalent to a 0 value for the unadjusted *Polity2* variable from the *Polity IV* dataset, and much lower than the cut-off for defining democracy in most studies.

[23] Angola, UAE, Azerbaijan, Bahrain, Bhutan, Botswana, Republic of the Congo, Algeria, Ecuador, Egypt, Eritrea, Gabon, Guinea, Guinea-Bissau, Equatorial Guinea, Indonesia, Iran, Iraq, Kuwait, Libya, Myanmar, Nigeria, Qatar, Saudi Arabia, Syria, Turkmenistan, Timor-Leste, Venezuela and Yemen.

transitions to (or from) democracy. Each estimator is subject to particular but distinct limitations. and we view the parallel implementation of all three as the best means to account for potential weaknesses in any individual estimation strategy.

3.1 Pooled Ordinary Least Squares (OLS), Fixed Effects (FE) and General Method of Moments (GMM)

We begin with the following simple dynamic specification:

$$Democracy_{i,t} = \delta Democracy_{i,t-l} + \beta_1 Tax\_rel_{i,t-l} + X'_{i,t-l}\ \gamma + \alpha_i + \lambda_t + e_{i,t} \quad [1]$$

In the basic specification, $Democracy_{i,t}$ is the measure of accountability for country *i* in period *t*. On the right-hand of equation [1], we include a lagged value of the dependent variable both to model the persistence of democracy over time, and to account for potential mean reverting dynamics. The key independent variable of interest is the lagged value of Tax_rel_i. In our main specifications we consider a single year lag, t-l where l=1, while we later also run robustness checks using *l*=3 and 5. The estimated parameter β_1 captures the impact of tax reliance on the extent of accountability. The vector $X'_{i,t}$ includes our time-varying control variables. We add country fixed effects, α_i, to control for omitted country-specific factors that do not change over time, and time dummies, $\lambda_{t,}$ to control for time-varying shocks that are common across countries. In all regressions, the error term, $e_{i,t}$, captures all other omitted factors.

After initially estimating the model using *tax_rel* as our independent variable, we disaggregate tax reliance into its component parts – tax revenue as a share of GDP (*tottax*) and non-tax revenue as a share of GDP (*totnotax)* – in order to attempt to disentangle competing causal arguments. Thus the second specification is:

$$Democracy_{i,t} = \delta Democracy_{i,t-l} + \beta_1 tottax_{i,t-l} + \beta_2 totnontax_{i,t-l} + X'_{i,t-l}\ \gamma + \alpha_i + \lambda_t + e_{i,t} \quad [2]$$

where the parameters β_1 and β_2 now capture the effects of total tax revenue and total non-tax revenue separately.

An important concern in relation to the tax reliance (*tax_rel*) and total tax revenue (*tottax*) variables is the possibility of reverse causation: while increased taxation may drive increased democracy, increased democracy might also lead to increased taxation by encouraging greater quasi-voluntary tax compliance, or spurring expanded demands for government services (Levi 1988; Timmons 2010). This concern does not appear to apply to non-tax revenue, as there is no clear reason to expect democracy to shape non-tax revenue collection. Ross (2004: 238) argues that the use of a lagged dependent variable should address the problem of reverse causality, based on the assumption that, 'a change in the independent variables (taxes and government services) should precede a change in the dependent variable (regime type)'.

However, this is a highly imperfect solution, as it implies an a priori assumption about the direction and timing of causation. A more econometrically ideal approach would be to identify a source of exogenous variation in tax and non-tax revenue by adopting a quasi-experimental design, or by employing an instrumental variables approach. However, over more than a decade of econometric research neither strategy has been implemented fully convincingly, with only highly imperfect instruments available.[24] We therefore follow the bulk

[24] In an interesting contribution, Ramsey (2011) employs natural disasters, which are expected to drive increased oil prices, as an instrument for increased oil revenue. However, this approach focuses attention primarily on the impact of short-term changes in oil prices rather than the long-term impacts of resource dependence, while, like most studies, his is an instrument for resource income rather than the level of revenue accruing to governments. Baskran and Bigsten

of the existing literature in concluding that panel data methods offer the best available strategy for estimating the relationship of interest.

For the sake of completeness we first follow Ross (2004) by implementing a Pooled OLS model. However, the Pooled OLS model omits country fixed effects, and is likely to suffer from omitted variables bias (OVB). We subsequently run the FE model in order to control for time-invariant country heterogeneity. However, FE estimates that include a lagged dependent variable as a regressor, as in our case, are biased because of the correlation between the lagged dependent variable and the error term (Nickell 1981; Bond 2002). While neither result is thus robust, Bond (2002) explains that the initial naïve OLS regression is biased upwards, while the FE regression is biased downwards, implying that together they provide bounds on the actual values of the lagged dependent variable, thus offering a useful check on any subsequent results.

In order to estimate a dynamic process with fixed effects we thus follow Aslaksen (2010) in employing GMM estimators that use instruments generated from the time series. The GMM models are considered the best estimator for dynamic processes with fixed effects where there are potentially endogenous regressors and idiosyncratic disturbances (country-specific patterns of heteroscedasticity and autocorrelation). Difference-GMM uses first differences as instruments, while system-GMM uses both first differences and lagged levels. Sys-GMM is preferable when – as in this case – there is a high degree of persistence of the variables of interest, as in these cases Diff-GMM suffers from weak instrumentation (Blundell and Bond 1998). By using lagged levels of the variables as instruments, Sys-GMM preserves information on cross-country differences that is lost when only the first differenced equation is estimated, and is therefore more efficient (Aslaksen 2010).

3.2 Mean Group-Common Correlated Effects (CCE-MG) estimator

Reliance on Sys-GMM has become increasingly common across a wide range of topics in political science and economics, but is not without its limits. Most notably, its comparative advantage is with large-N but small-T datasets, whereas we have up to twenty-one years of data from 1990-2010. Haber and Menaldo (2011) correspondingly employ panel ECM regressions in order to make use of the time series dimension of their data, and we test their approach in the robustness tests to follow. However, in our view a more appropriate method for exploiting the time series dimension of the data is to use the CCE-MG estimator, which has been developed specifically to deal with such macro panels.[25] Relative to conventional panel techniques it addresses both parameter heterogeneity and cross-section dependence, which we describe in turn (e.g. Holly et al. 2010; Cavalcanti et al. 2011; Fayad et al. 2011; Eberhardt and Teal 2013).

Parameter heterogeneity concerns the slope coefficients in our estimated relationship. Pooled OLS, FE, GMM and panel ECM estimators all assume that these slope coefficients are constant across countries. It is at least plausible that they are not, and Pesaran and Smith (1995) show that in this case none of the methods mentioned produces a consistent estimate of the cross-country average slope. Notably, instrumental variables cannot help us: any potential instrument that is correlated with the regressor will necessarily also be correlated with the error term. The point therefore applies equally to the GMM estimates, which depend on lagged variables as instruments.

(2013) employ the manufacturing share of exports as an instrument for tax revenue, but existing empirical research suggests that a larger manufacturing sector is only a weak predictor of tax collection (e.g. Gupta 2007), while there are equal concerns that manufactured exports may be an independent cause of improved governance, as a large middle class has been widely argued to be a key determinant of democratisation (Moore 1966; Birdsall 2007).

[25] Haber and Menaldo (2011) also employ a dramatically longer time series, dating back to 1900, making their study more appropriate to time series methods.

Pesaran and Smith's (1995) MG estimator overcomes this problem very simply by estimating the relationship between the levels of the dependent variable and the independent variable for each country individually, and then taking an average of the country-specific coefficients. However, this comes at a cost: the MG estimator does not model dynamics, and does not account for reverse causality. As such, if there is reverse causality then MG estimates will include the effects of both directions of causality, and cannot be interpreted simply as the impact of the independent variables on the dependent variable. Indeed, we find that our democracy variable does Granger-cause our revenue variables, suggesting that there is reverse causality.[26] GMM-based estimates, while suffering from problematic instruments in the presence of heterogeneous coefficients, do not suffer from this problem.

The second related issue addressed by the MG estimator is that both global factors that affect all countries, and spillover effects from one country to another, can lead to cross-section correlation, confounding standard regression estimates (Pesaran 2006). While year dummies could account for the first type of effect, they could not account for the second. In our case, spillover effects are highly plausible: democratic transition in one country seems likely to increase pressure on its neighbours to follow suit, as evinced by the recent 'Arab Spring' and arguably the wave of democratisation in Latin America in the 1980s. Pesaran (2006) develops CCE estimators to account for this cross-section correlation. These estimators augment the regression with year-specific cross-section (in our case, cross-country) averages of the variables as proxies for unobserved global and local factors, allowing the coefficients on these averages to vary across countries. This technique can be combined with the MG estimator above, giving us the CCE-MG estimator.

3.3 Logistic regressions focusing on regime transitions

The GMM and CCE-MG estimators offer robust and complementary approaches to estimating the relationship between the composition of government revenue and continuous measures of the level of democracy. However, several studies have argued for a more narrow focus on the impact of the composition of government revenue on the likelihood of *transition* from autocracy to democracy or, alternatively, from democracy to autocracy.

At the level of theory, Smith (2004), Morrison (2009) and others have proposed that resource wealth may not be inherently pro- or anti-democratic, but may simply make any existing form of government more resistant to change. That is, regime transitions from either autocracy or democracy may become less likely with greater access to resource wealth, as resource wealth allows regimes to secure political support and confront political resistance. In this case, we would not expect to find a positive or negative relationship between resource dependence and democracy, but, rather, an impact of resource dependence on the likelihood of political transition.

A focus on regime transitions equally addresses a pragmatic estimation problem: many resource-rich countries were already very undemocratic prior to acquiring large-scale resource revenue (or prior to the expansion of existing resource revenue). As a result, it is often impossible for increased resource revenue to lead to a significant further decrease in the level of democracy. Instead, in these cases we expect increased resource wealth to result in greater persistence of autocracy, reflected in a reduced likelihood of a transition towards greater democracy.

---

[26] We test Granger causality by regressing variable *y* on its own first and second lag, and on the lag of variable *x,* where *x* Granger-causes *y* if the coefficient on the lag of *x* is significant. We find that *polity* Granger-causes *totnotax* with p=6.7%, *polity* Granger-causes *tottax* with p=1.5%, *totnotax* Granger-causes *polity* with p=0.0%, and *tottax* Granger-causes *polity* with p=1.3%. In all cases the sign is as expected: *polity* is negatively correlated with *totnotax* and positively correlated with *tottax*. Further details available from the authors.

We thus conclude the empirical analysis by following Wiens et al. (forthcoming) in estimating a dynamic logit model that captures the impact of the composition of government revenue on the likelihood that a country will transition between democracy and autocracy. Following Wiens et al. (forthcoming), our prediction is that higher non-tax revenue and reduced tax reliance will reduce the likelihood of transition to democracy, but will have little, if any, impact on the likelihood of transition from democracy to autocracy. The latter reflects arguments that the political resource curse is likely to be less important where democracy is already established, and political institutions are comparatively resilient (Mehlum et al. 2006).

In order to test these predictions the model differs somewhat from the discussion so far. When we test the impact of *tax_rel* it is:

$$\Pr(\text{Regime}_{i,t}) = \Lambda\Big[\delta\text{Regime}_{i,t-1} + \beta_1\text{Tax\_rel}_{i,t-1} + \beta_2\text{Regime} \times \text{Tax\_rel}_{i,t-1} + \beta_3 X'_{i,t-1} + \beta_4\text{Regime}_{i,t-1} \times X'_{i,t-1} + \lambda_t + e_{i,t}\Big] \quad [3]$$

While separating *tax_rel* into its component parts, *tottax* and *totnontax*, gives us:

$$\Pr(\text{Regime}_{i,t}) = \Lambda\Big[\delta\text{Regime}_{i,t-1} + \beta_1\text{Tax\_rel}_{i,t-1} + \beta_2\text{Regime} \times \text{Tax\_rel}_{i,t-1} + \beta_3\text{Tax\_rel}_{i,t-1} + \beta_4\text{Regime} \times \text{Tax\_rel}_{i,t-1} + \beta_5 X'_{i,t-1} + \beta_6\text{Regime}_{i,t-1} \times X'_{i,t-1} + \lambda_t + e_{i,t}\Big] \quad [4]$$

In both equations the new measure of democracy is the binary variable *Regime*, which is coded 1 if the country is a democracy, and 0 if the country is an autocracy. The variable is based on the variable presented in Przeworski et al. (2000) and extended through 2008 by Cheibub et al. (2010). Our dependent variables, tax reliance (*tax_rel)*, total tax revenue (*tottax)* and total non-tax revenue (*totnnotax)*, are unchanged from earlier models. *X'* is our group of control variables. Our primary results employ the same controls employed by Wiens et al. (forthcoming) for the sake of comparability: *lgdp*, *civil_war* and *growthpc*. However, we also present results using the same battery of control variables employed in the Sys-GMM and CCE-MG estimations, and the results are essentially unchanged. λ_t are time dummies for every year in the sample, and $e_{i,t}$ are the error terms.

The key innovation in Wiens et al. (forthcoming), and replicated here, is the inclusion of interaction terms between lagged *Regime* and each of the variables on the right-hand side of the equation. The interaction terms condition the effect of each of the independent variables on the presence or absence of democracy, and allow us to identify whether the revenue variables have distinct impacts in democracies and autocracies. The total effect of the revenue variables in democracies is thus the sum of the coefficients on the revenue term, and on the interaction terms between the revenue terms and regime.

As with earlier models, we wish to account for the effects of unobserved country heterogeneity. However, the inclusion of fixed effects when employing a binary measure of democracy results in dropping from the analysis any country that did not experience a political transition between 1990 and 2010. This leaves only twenty-nine countries, while excluding almost all major resource producers.[27] Wiens et al. (forthcoming) address this problem by employing a random effects specification in order to account for unobserved country heterogeneity as far as possible while retaining necessary country coverage. This is an imperfect solution to the problem of unobserved heterogeneity, but is a potentially

[27] This figure includes 16 transitions to autocracy, and 46 transitions to democracy. When the data is extended back to include the period 1980-2010 in our robustness checks, the analysis includes 24 transitions to autocracy and 63 transitions to democracy, across 41 countries.

sensible econometric compromise in the interest of testing a more theoretically attractive model.

In order to test the validity of adopting a random effects specification, we run a Hausman test of the null hypothesis that there are no systematic differences between the coefficients in the fixed effects and the random effects models (Hausman 1978). The test returns a negative value, which the literature generally interprets as a valid basis for employing random effects (Hausman and McFadden 1984). However, the robustness of simply accepting a negative Hausman statistic has been questioned (Vijverberg 2011), and, in order to err on the side of caution, we report the first set of core results employing both the random effects and fixed effects estimations, as the key messages are broadly consistent across both models.[28]

4 Results

We first present results employing our array of panel data techniques, before turning to the results of the dynamic logit regressions. Table 1a presents estimates for Pooled OLS, Fixed Effects, Diff- and Sys-GMM and CCE-MG models, with tax reliance (*tax_rel*) as the independent variable. Table 1b presents estimates when we instead include total tax revenue (*tottax*) and total non-tax revenue (*totnontax*) as distinct regressors. In both cases we initially adopt a parsimonious specification of our model, controlling only for the log of GPD per capita (*lgdp*). [29]

Table 1a Effect of tax reliance on polity, across different estimation methods

	(1) Pooled OLS b/se	(2) FE b/se	(3) Diff-GMM two-step b/se	(4) Sys-GMM two-step b/se	(5) CCE-MG w/o trend b/se	(6) CCE-MG with trend b/se
L.Polity_s	0.9410*** (0.0113)	0.7231*** (0.0289)	0.8040*** (0.0422)	0.8741*** (0.0359)		
L.Tax_Rel	4.4130*** (1.1760)	0.3130 (2.9529)	2.9361 (8.2961)	18.0684*** (5.3422)		
Tax_Rel					14.3193** (7.2816)	16.2505** (7.1634)
L.lgdp	0.2371 (0.1607)	-0.5221 (1.1645)	5.7623* (2.9603)	-0.8896 (1.2806)		
lgdp					9.6114* (5.1366)	11.1951 (10.2255)
trend						0.3168 (0.6178)
AR(1)			0.000	0.000		
AR(2)			0.623	0.648		
Hansen			0.222	0.376		
No. of obs.	2579	2579	2424	2579	2682	2674
No. of countries		155.000	155.000	155.000	154.000	152.000
r2	0.937	0.602				

Notes: All regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. Residuals of the CCE-MG regressions were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend). See text for discussion.

Source: ICTD GRD (2014).

[28] While not reported here, all the subsequent robustness checks are also implemented using fixed effects, and the pattern of results is consistent. Results available on request.

[29] Our core results treat *lgdp* as endogenous given the possibility that democracy may affect economic growth, though the results remain positive and significant when lgdp is treated as exogenous.

Table 1b Effect of total tax and non-tax revenue on polity, across different estimation methods

	(1) Pooled OLS b/se	(2) FE b/se	(3) Diff-GMM two-step b/se	(4) Sys-GMM two-step b/se	(5) CCE-MG w/o trend b/se	(6) CCE-MG with trend b/se
L.Polity_s	0.9422*** (0.0112)	0.7244*** (0.0285)	0.7985*** (0.0369)	0.8612*** (0.0280)		
L.TotTax	3.8814 (2.7511)	-17.4542** (6.8917)	-18.3505 (16.2196)	-11.7086 (15.5717)		
L.totnontax	-9.0711*** (2.4282)	-6.8190 (4.4497)	-4.3349 (9.6489)	-26.0380*** (8.3181)		
TotTax					73.7190** (30.8115)	54.0306* (29.3426)
totnotax					-137.2164** (57.2204)	-156.036*** (58.4479)
L.lgdp	0.1472 (0.1579)	-0.4235 (1.1072)	2.8651 (2.7534)	-1.6612 (1.3640)		
lgdp					10.7081* (6.4542)	14.3815 (11.0953)
trend						-0.0880 (0.5124)
AR(1)			0.000	0.000		
AR(2)			0.647	0.650		
Hansen			0.140	0.291		
No. of obs.	2579	2579	2424	2579	2674	2669
No. of countries		155.000	155.000	155.000	152.000	151.000
r2	0.937	0.603				

Notes: All regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. Residuals of the CCE-MG regressions were tested for non-stationarity. In both cases the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend). See text for discussion.

Source: ICTD GRD (2014).

As can be seen in Table 1a, pooled OLS produces a strongly significant positive coefficient on (lagged) tax reliance, but the coefficient is insignificant when we take country-specific fixed effects into account with the FE model. Turning to the GMM estimates, the results remain insignificant when employing the Diff-GMM estimator, but are again positive and significant at the 5 per cent level when we use Sys-GMM.

The estimates employ two-step Windmeijer corrected standard errors, which correct for downward bias in the standard errors in a finite sample (given that the GMM framework is based on asymptotically large samples), thus providing larger and more reliable confidence intervals. Following Roodman (2006, 2009) the Sys-GMM results also include tests for second-order autocorrelation, as well as results of the Hansen test of instrument validity, and in both cases we fail to reject the null hypothesis that the results are valid.[30] The overall pattern of results is entirely consistent with results reported by Aslaksen (2010) in her tests of the relationship between resource wealth and democracy. As in her case, we attribute the stronger results when employing Sys-GMM to the fact that the Sys-GMM estimator makes use of a wider variety of information, and is thus more efficient when the key variables change slowly.

[30] Roodman (2009) urges caution in interpreting Hansen test statistics, for which implausibly high values approaching 1 can indicate the failure of the test owing to the excessive proliferation of instruments. For this reason he urges reliance wherever possible on more parsimonious specifications, which is the strategy that we adopt here.

The CCE-MG estimates also produce positive coefficients on tax reliance, significant at the 5 per cent level both with and without trend, as the trend itself is highly insignificant. For these regressions we tested the residuals for stationarity. The Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1 per cent level, while the Pesaran (2007) panel unit root test (CIPS) fails to reject non-stationarity. The primary difference between the two is that Maddala and Wu (1999) assumes cross-section independence, while CIPS does not. Since the residuals are the result of regressions that already control for cross-section correlation (the CCE part of the CCE-MG estimator), they are indeed cross-sectionally independent, so the Maddala and Wu (1999) test is valid and we can accept its rejection of non-stationarity.

Table 1b disaggregates tax reliance into tax revenue (*tottax*) and non-tax revenue (*totnontax*), and the results suggest that the positive and significant coefficients on tax reliance are driven primarily by the effect of non-tax revenue, though the results are not uniform. In the Sys-GMM regression the coefficient on *totnontax* is negative and significant at the 1 per cent level, while the coefficient on *tottax* is insignificant and has an unexpectedly negative sign. By partial contrast, in the CCE-MG regressions both *tottax* and *totnontax* are significant and with the expected sign. As with the tax reliance variable, the Maddala and Wu (1999) test confirms that the residuals are stationary. The greater significance of the *tottax* variable in the CCE-MG estimations may be explained by the fact that the CCE-MG estimator captures the long-run relationship in levels between tax collection and democracy. Because it does not account for reverse causation, the result may include a positive impact of democracy on tax collection – consistent with our finding of Granger causation in both directions. Meanwhile, the CCE-MG model also relaxes assumptions about the dynamic structure of the data – a point to which we return in the discussion.

Having established this baseline, Tables 2a and 2b progressively introduce additional control variables for the Sys-GMM and CCE-MG models in turn. Table 2a reports results for the Sys-GMM estimations, and we find that the results are essentially unchanged with the inclusion of the new control variables, with the exception of *tottax* becoming marginally significant and again negative in Column 10. Table 2b reports results for the CCE-MG estimations, though we do not included the control for regional democratic diffusion given that CCE-MG already controls for cross-section correlation.[31] Again the bulk of the results prove robust to alternative control variables. The coefficients on both *tax_rel* and *totnontax* become larger and more significant with additional controls. However, the positive coefficient on the total tax variable in our baseline specification – which includes only log GDP as a control – is not robust to the addition of the control variables. The overall value of adding the control variables is open to question as they are almost universally insignificant, but the sensitivity of the total tax variable is notable.

[31] The pattern of significance is unchanged when regional democratic diffusion is included, but the point estimates become somewhat more imprecise (though also somewhat larger).

Table 2a Effect of tax reliance, total tax and non-tax revenue on polity, controlling for a full set of covariates

	(1) Sys-GMM two-step b/se	(2) Sys-GMM two-step b/se	(3) Sys-GMM two-step b/se	(4) Sys-GMM two-step b/se	(5) Sys-GMM two-step b/se	(6) Sys-GMM two-step b/se	(7) Sys-GMM two-step b/se	(8) Sys-GMM two-step b/se	(9) Sys-GMM two-step b/se	(10) Sys-GMM two-step b/se
L.Polity_s	0.8905***	0.8741***	0.8726***	0.8658***	0.8671***	0.8652***	0.8612***	0.8642***	0.8523***	0.8624***
	(0.0361)	(0.0359)	(0.0371)	(0.0365)	(0.0353)	(0.0309)	(0.0280)	(0.0271)	(0.0288)	(0.0278)
L.Tax_Rel	11.4557**	18.0684***	17.3664***	14.9164**	15.8768***					
	(5.8238)	(5.3422)	(5.5898)	(5.9211)	(5.2605)					
L.TotTax						-10.5979	-11.7086	-10.7905	-16.8992	-25.5627*
						(13.5838)	(15.5717)	(14.8840)	(15.4091)	(15.2234)
L.totnotax						-17.2002*	-26.0380***	-27.0054***	-28.5379***	-29.7871***
						(9.1191)	(8.3181)	(8.2345)	(9.0730)	(10.1836)
L.lgdp		-0.8896	-0.6360	-1.0146	-1.7376		-1.6612	-1.2962	-0.8832	-2.5850
		(1.2806)	(1.2099)	(1.0250)	(1.5148)		(1.3640)	(1.3039)	(1.3268)	(1.6874)
L.Pop			-0.0000	-0.0000	-0.0000			-0.0000	-0.0000	-0.0000
			(0.0000)	(0.0000)	(0.0000)			(0.0000)	(0.0000)	(0.0000)
L.Civil_War				-1.0229	-1.3181				-1.7178	-1.7057
				(1.3300)	(1.4771)				(1.2118)	(1.3935)
L.Region_Dem Diffuse					0.0755					0.1562***
					(0.0501)					(0.0503)
Implied long-term effect	104.6182	143.5139	136.3140	111.1505	119.4643	-127.597	-187.593	-198.862	-193.215	-216.476
AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.644	0.648	0.650	0.664	0.693	0.651	0.650	0.648	0.682	0.731
Hansen	0.560	0.376	0.351	0.344	0.521	0.207	0.291	0.295	0.192	0.259
N	2648	2579	2576	2544	2536	2648	2579	2576	2544	2536
N_g	157.000	155.000	155.000	153.000	152.000	157.000	155.000	155.000	153.000	152.000

Notes: All regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. Implied long-term effect calculated using the formula $\beta/(1 - \omega)$, where ω is the coefficient on the lagged dependent variable, and β is the coefficient on the revenue variable of interest. Long term effects in Columns (1) – (5) is for *tax_rel*, while for columns (6) – (10) it is for *totnontax*.

Source: ICTD GRD (2014).

Table 2b Effect of tax reliance, total tax and non-tax revenue on polity, controlling for a full set of covariates

	(1) CCE-MG w/o trend b/se	(2) CCE-MG w/o trend b/se	(3) CCE-MG w/o trend b/se	(4) CCE-MG w/o trend b/se	(5) CCE-MG w/o trend b/se	(6) CCE-MG w/o trend b/se	(7) CCE-MG w/o trend b/se	(8) CCE-MG w/o trend b/se
Tax_Rel	14.6396*	14.3193**	18.3955***	20.6028***				
	(8.6574)	(7.2816)	(6.8242)	(7.6264)				
TotTax					19.4631	73.7190**	41.4703	51.7679
					(28.3911)	(30.8115)	(29.0371)	(35.1532)
totnotax					-136.4237**	-137.2164**	-158.5317***	-155.2035**
					(63.9151)	(57.2204)	(61.3340)	(66.4175)
lgdp		9.6114*	5.6048	5.5947		10.7081*	9.1711	4.7162
		(5.1366)	(8.5748)	(8.3597)		(6.4542)	(10.8602)	(10.6713)
Pop			0.0000	0.0000			-0.0000	-0.0000
			(0.0000)	(0.0000)			(0.0000)	(0.0000)
Civil_War				-0.1046				-0.5235
				(0.4150)				(0.3874)
N	2754	2682	2671	2634	2751	2674	2666	2604
N_g	156.000	154.000	152.000	149.000	155.000	152.000	151.000	144.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. Residuals of the CCE-MG regressions were tested for non-stationarity. In all cases, the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend). See text for discussion.

Source: ICTD GRD (2014).

The magnitude of the effects of tax reliance and non-tax revenue on democracy is substantial. Looking first at the Sys-GMM estimates, we draw on the results that include the full set of controls (columns 5 and 10 of Table 2a). The coefficient of 15.89 on *tax_rel* implies that a 10 percentage point increase in *tax_rel* will produce a short-term increase in democracy of 1.59 percentage points. To place these magnitudes in perspective, a move from the level of tax reliance of Senegal (average tax_rel = 0.91) to that of Nigeria (average tax_rel = 0.22) produces a relatively modest 11 percentage point decline in the level of democracy.

However, we can also calculate the implied long-term effect, reported near the bottom of Table 2a, and it is substantially larger: over the long term, the same 70 percentage point decline in tax reliance is expected to reduce the level of democracy by 80 percentage points – the equivalent of going from the United States to Kazakhstan or Morocco.[32] The individual impact of changes in non-tax revenue on democracy is substantively similar: a change from the level of non-tax revenue in Senegal (average non-tax revenue = 1.5 per cent of GDP) to that of Nigeria (average non-tax revenue = 22 per cent of GDP) is expected to generate a short-term decline in democracy of 6 percentage points, and a long-term decline of 43 percentage points. An increase in non-tax revenue to a level of 43 per cent of GDP – the level in Angola in 2008 – would imply a long-term decline in democracy of almost 90 percentage points.

The CCE-MG results provide estimates exclusively of the long-term relationship, and we focus on the results in Columns 4 and 9. The magnitude of the relationship between *tax_rel* and democracy is comparable to the short-run estimates for the Sys-GMM results, and substantially smaller than long-run estimates from those regressions. The coefficient of 20.6 on tax reliance implies that a rise of 10 percentage points in tax reliance is associated with a rise of 2.06 percentage points in the *polity* score. Taking the same example as previously, a shift from the level of tax reliance in Senegal to that in Nigeria would imply a 14.2 per cent decline in democracy – the equivalent in 2010 of moving from Malawi (polity = 80) to Niger (polity = 65). However, the coefficient on non-tax revenue is substantially larger, and more in line with the Sys-GMM results: it implies that a 10 percentage point increase in non-tax revenue as a share of GDP is associated with a 15.5 percentage point decrease in democracy. Again using the earlier example, an increase from the level of non-tax revenue in Senegal to that of Nigeria would be associated with a 31 percentage point decline in democracy, while non-tax revenue at the level of Angola would imply a decline in democracy of greater than 60 percentage points.

Having reviewed the Sys-GMM and CCE-MG results, Tables 3a and 3b present the results employing the dynamic logit model, reporting, in turn, results using random effects and fixed effects. The results estimate the effect of the composition of government revenue on the likelihood of transition between autocracy and democracy. Columns 1 and 2 in each table report results without controls, columns 3 and 4 report results employing the same controls employed in the Sys-GMM and CCE-MG estimates, and columns 5 and 6 report results including the same controls as Wiens et al. (forthcoming): *lgdp*, *civil_war* and *growthpc*.

[32] Following Aslaksen (2010) the long-term effect is calculated as $\beta/(1 - \delta)$, where δ is the coefficient on the lagged dependent variable, and β is the coefficient on the revenue variable of interest.

Table 3a Effect of tax reliance, total tax and non-tax revenue on regime using random effects logit model

	(1) Logit - RE b/se	(2) Logit - RE b/se	(3) Logit - RE b/se	(4) Logit - RE b/se	(5) Logit - RE b/se	(6) Logit - RE b/se
L.regime	8.7893*** (1.4890)	4.0642*** (0.9087)	-3.8956 (4.0110)	1.2185 (3.3445)	-5.2888 (3.5544)	-0.6607 (2.8037)
L.Tax_Rel	2.0738** (0.8913)		0.9573 (1.1176)		1.5100 (1.0490)	
L.regimextaxrel	-0.1405 (1.8853)		2.4272 (1.8739)		1.9235 (1.8691)	
L.TotTax		-5.5241* (3.1807)		-6.2322 (3.8912)		-6.2755* (3.7480)
L.TotNonTax		-9.8712*** (3.6259)		-10.5769* (5.4163)		-12.5968** (5.3654)
L.regimextottax		27.4911*** (6.5148)		17.9933** (8.8547)		16.7894** (8.1804)
L.regimextotnontax		13.6779** (6.8268)		7.2707 (7.4311)		8.6665 (7.4713)
L.lgdp			-0.4881* (0.2557)	-0.1096 (0.2902)	-0.2379 (0.2263)	0.1238 (0.2608)
L.Pop			-0.0000 (0.0000)	-0.0000 (0.0000)		
L.Civil_War			-0.9964 (0.7721)	-1.1218 (0.7755)	-1.0092 (0.7729)	-1.1202 (0.7740)
L.Region_Dem_Diffuse			0.0243** (0.0117)	0.0200* (0.0114)		
L.Growthpc					-0.0333 (0.0303)	-0.0336 (0.0295)
L.regimexlgdp			1.2304** (0.4999)	0.4326 (0.5652)	1.5121*** (0.3810)	0.7510* (0.4431)
L.regimexpop			0.0000 (0.0000)	0.0000 (0.0000)		
L.regimexcivilwar			-0.2575 (1.1924)	-0.1232 (1.1928)	0.2652 (1.1874)	0.3397 (1.1855)
L.regimexdemdiffuse			0.0044 (0.0220)	0.0086 (0.0216)		
L.regimexgrowthpc					0.0917 (0.0681)	0.0911 (0.0685)
Aggregate effects in democracies						
$\beta 1 + \beta 2$	1.9333 (1.6552)		3.3844** (1.5016)		3.4334** (1.5399)	
$\beta 1 + \beta 3$		21.9669*** (5.7284)		11.7610 (7.9517)		10.51384 (7.2802)
$\beta 2 + \beta 4$		3.8067 (5.7743)		-3.3062 (5.1022)		-3.9303 (5.1777)
Random effects model descriptors						
σ_μ	.0001892	.0001973	.0021279	.0020675	.0020783	.001475
ρ	1.09e-08	1.18e-08	1.38e-06	1.30e-06	1.31e-06	6.61e-07
N	2807	2807	2322	2322	2305	2305
N_g	183.000	183.000	155.000	155.000	157.000	157.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. $\beta 1 + \beta 2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta 1 + \beta 3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta 2 + \beta 4$ captures the joint significance of *totnontax* and *regime x totnontax*.

Source: ICTD GRD (2014).

Table 3b Effect of tax reliance, total tax and non-tax revenue on regime using fixed effects logit model

	(1) Logit - FE b/se	(2) Logit - FE b/se	(3) Logit - FE b/se	(4) Logit - FE b/se	(5) Logit - FE b/se	(6) Logit - FE b/se
L.regime	5.0349***	4.9087	9.8169	11.3485*	4.3906	4.9988
	(1.7079)	(1.1528)	(7.4420)	(5.9689)	(5.0130)	(3.8021)
L.Tax_Rel	6.4060**		7.8037**		7.6642**	
	(3.0239)		(3.3873)		(3.3890)	
L.regimextaxrel	-.67088		-0.5483		-0.2264	
	(2.1514)		(2.8489)		(2.3673)	
L.TotTax		25.2865*		26.3396		21.6904
		(14.7148)		(16.2750)		(15.8069)
L.TotNonTax		-10.7415		-19.5740		-15.0584
		(8.2790)		(12.2259)		(10.0924)
L.regimextottax		-4.1663		-2.0454		-4.5262
		(8.6561)		(10.3500)		(9.6103)
L.regimextotnontax		1.5316		1.1267		2.0054
		(7.0117)		(11.9882)		(8.0799)
L.lgdp			-0.0966	-0.6407	0.9682	0.9285
			(2.1380)	(2.3219)	(1.8678)	(1.9143)
L.Pop			0.0000	0.0000		
			(0.0000)	(0.0000)		
L.Civil_War			-0.7234	-0.8511	-0.8119	-0.7654
			(1.1417)	(1.1757)	(1.0971)	(1.1020)
L.Region_Dem_Diffuse			0.0855*	0.0682		
			(0.0519)	(0.0522)		
L.Growthpc					0.0173	0.0061
					(0.0389)	(0.0420)
L.regimexlgdp			-0.9150	-1.2311	0.0445	-0.0090
			(0.9292)	(0.9771)	(0.5593)	(0.5763)
L.regimexpop			0.0000*	0.0000*		
			(0.0000)	(0.0000)		
L.regimexcivilwar			-1.3783	-1.2913	-0.4830	-0.4226
			(1.4852)	(1.5142)	(1.3719)	(1.3703)
L.regimexdemdiffuse			0.0358	0.0504		
			(0.0291)	(0.0308)		
L.regimexgrowthpc					0.0759	0.0797
					(0.0848)	(0.0853)
Aggregate effects in democracies						
$\beta 1 + \beta 2$	5.7352*		7.2553*		7.4378*	
	(3.4097)		(4.1093)		(3.8090)	
$\beta 1 + \beta 3$		21.1202		24.2942		17.1641
		(15.6522)		(17.3369)		(16.4619)
$\beta 2 + \beta 4$		-9.2098		-18.4473		-13.0529
		(8.8622)		(13.9366)		(8.7285)
N	479	479	458	458	458	458
N_g	29.000	29.000	28.000	28.000	28.000	28.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. $\beta 1 + \beta 2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta 1 + \beta 3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta 2 + \beta 4$ captures the joint significance of *totnontax* and *regime x totnontax*.

Source: ICTD GRD (2014).

The overall results are again generally consistent with expectations, though they are somewhat 'noisy' and suggest additional nuance. Focusing on the random effects estimates, the most consistent finding is a negative and significant relationship between *totnontax* and *regime*. Because of the interaction terms in the model, the coefficient on *totnontax* (and those on the other revenue variables) tells us that an increase in non-tax revenue reduces the likelihood that a non-democratic state will transition to democracy. The tax reliance variable, *tax_rel*, likewise follows expectations, as it is positive and significant when no controls are included in the model, though it loses significance at conventional levels as controls are added. The comparative weakness of the *tax_rel* variable is, in turn, attributable to the total tax revenue variable, *tottax*, which is unexpectedly negative, and significant at the 10 per cent level in some specifications. As with previous results, the overarching message is that *totnontax* has a significantly negative impact on the likelihood of transition to democracy, while the impact of tax revenue is very mixed.

However, the subsequent results complicate this simple story. First, the fixed effects results – which act as an important robustness check given the ambiguous Hausman test results noted earlier – follow a somewhat different pattern from the random effects estimates. The coefficient on *totnontax* remains negative, with a magnitude comparable to the random effects results, but falls slightly short of statistical significance. This can in part be explained by the dramatically reduced sample size. Meanwhile, *tottax* is now both positive in all specifications, and significant when no controls are included. This, in turn, results in the coefficient on *tax_rel* being positive and significant across all the fixed effects results. Both the fixed and random effects results thus indicate the existence of a political resource curse, but they tell a mixed story about the respective roles of tax and non-tax revenue in driving this relationship.

Adding further complexity are the interaction terms, which allow us to discern the distinct effect of the revenue variables in democracies. The interaction term itself tells us whether the relationship is statistically different in democracies than in autocracies, while the overall effect in democracies is calculated as the sum of the coefficient on the interaction term and the coefficient on the revenue term (reported below the core results).[33] As noted earlier, Wiens et al. (forthcoming) find no effect of natural resource wealth in democracies, while Morrison (2009) argues that greater non-tax revenue in fact strengthens existing democracies.

Our results mirror those of Wiens et al. (forthcoming) – and contradict those of Morrison (2009) – as we find that *totnontax* does not have a significant effect in democracies, though the coefficient remains negative (row $\beta 2 + \beta 4$ in Table 3a). Like Wiens et al. (forthcoming), we attribute this to the fact that democratic institutions are able to mitigate the negative political costs of resource wealth. However, we also find that the effect of *tax_rel* is positive and significant in democracies, in both the random effects and fixed effects models (row $\beta 1 + \beta 3$ in Table 3a). The implication of the random effects results is that while the negative impact of non-tax revenue is comparatively muted in existing democracies, the political benefits of tax collection are enhanced (with the *regime x tottax* interaction term positive and significant).[34]

To summarise, two conclusions appear relatively robust. First, we have relatively consistent evidence that increases in non-tax revenue result in a reduced likelihood of democratic transition. Second, we find a generally positive association between tax reliance and democracy in *both* autocracies and democracies, though the fixed effects and random effects

[33] For example, the impact of *tax_rel* in democracies is equal to the joint significance of *tax_rel* and *regime x taxrel*.

[34] The fixed effects results are consistent with this story, but suggest that autocracies and democracies are quite similar, with taxation moderately encouraging democracy and non-tax revenue moderately discouraging tax revenue in both cases.

estimates vary in the extent to which they attribute this finding to the democratising effects of taxation or the anti-democratic effects of non-tax revenue, highlighting the need for caution in drawing strong conclusions.

Because the logit estimates are non-linear, we separately calculate the magnitude of the effects, reported in Tables 3c and 3d, drawing on the results in columns 5 and 6 of Table 3a. The top half of Table 3c reports the likelihood that a country that is an autocracy in period *t-1* will become a democracy at time *t*, with different levels of non-tax revenue, no civil war and all other variables set to the mean level for developing countries.[35] The bottom half of the table, in turn, reports the likelihood of becoming a democracy at time *t* at different levels of tax reliance.[36] Table 3d reverses the question of interest, and reports the likelihood that a democratic country will remain a democracy at time *t*, for different levels of non-tax revenue and tax reliance.[37]

Table 3c Likelihood of transition to democracy

	Level of non-tax revenue as % of GDP				
	2%	**5%**	**10%**	**20%**	**30%**
Pr(Democracy) at *t*	0.0422	0.0292	0.0157	0.0044	0.0012
95% confidence interval (upper)	0.0652	0.0425	0.0266	0.0128	0.0046
95% confidence interval (lower)	0.0192	0.0159	0.0047	-0.0029	-0.0021
	Level of tax reliance (tax revenue share of total government revenue)				
	0.2	**0.4**	**0.6**	**0.75**	**0.9**
Pr(Democracy) at *t*	0.0114	0.0155	0.0211	0.0265	0.0333
95% confidence interval (upper)	0.0247	0.0280	0.0321	0.0387	0.0524
95% confidence interval (lower)	-0.0019	0.0031	0.0101	0.0143	0.0142

Notes: Likelihood of transition to democracy at different values of non-tax revenue and tax reliance. Calculated based on Columns 5 and 6 of Table 3a, using *margins* command in Stata, with *regime* = 0, *civil_war* = 0 and other variables set to their mean values excluding OECD countries. The mean level of *lgdp* is 7.89 (equivalent to per capita income of $2670), the mean level of *growthpc* is 2.51 and the mean level of *tottax* is 0.136 (13.6% of GDP).

Source: ICTD GRD (2014).

35 To calculate this likelihood we employ the *margins* command in Stata and set *regime* = 0, *civilwar* = 0 and the other variables at their mean values excluding OECD countries, given that most of the variation in *totnontax* and *tax_rel* occurs within this group. The mean level of lgdp is 7.89 (equivalent to per capita income of $2670), the mean level of *growthpc* is 2.51, and the mean level of total tax revenue as a share of GDP is 13.6%.

36 Though recognising that the coefficient on the tax reliance variable varies significantly across models.

37 Bearing in mind that the non-tax revenue variable is insignificant in democracies.

Table 3d Likelihood of transition to autocracy

	Level of non-tax revenue as % of GDP				
	2%	**5%**	**10%**	**20%**	**30%**
Pr(Democracy) at *t*	0.9897	0.9884	0.9861	0.9795	0.9702
95% confidence interval (upper)	0.9981	0.9970	0.9982	1.0128	1.0456
95% confidence interval (lower)	0.9814	0.9780	0.9739	0.9463	0.8949
	Level of tax reliance (tax revenue share of total government revenue)				
	0.2	**0.4**	**0.6**	**0.75**	**0.9**
Pr(Democracy) at *t*	0.9265	0.9622	0.9809	0.9886	0.9932
95% confidence interval (upper)	1.0393	1.0040	0.9956	0.9970	0.9996
95% confidence interval (lower)	0.8136	0.9204	0.9663	0.9803	0.9869

Notes: Likelihood of transition to democracy at different values of non-tax revenue and tax reliance. Calculated based on Columns 5 and 6 of Table 3a, using *margins* command in Stata, with *regime* = 0, *civil_war* = 0 and other variables set to their mean values excluding OECD countries. The mean level of *lgdp* is 7.89 (equivalent to per capita income of $2670), the mean level of *growthpc* is 2.51 and the mean level of *tottax* is 0.136 (13.6% of GDP).

Source: ICTD GRD (2014).

As with the earlier results, the magnitude of the effects is substantial. With all other variables at their means for developing countries, an autocratic country with non-tax revenue of 2 per cent of GDP (e.g. Mozambique or Gambia in 2008[38]) has a 4.2 per cent probability of transitioning to democracy in a given year. By contrast, an autocratic country with non-tax revenue of 10 per cent of GDP (e.g. Malaysia) has only a 1.6 per cent probability of a democratic transition, while non-tax revenue of 20 per cent of GDP (e.g. Gabon or Chad) drops that probability to 0.44 per cent, though at this level the values become more imprecise. Put differently, the likelihood of a democratic transition declines by almost a third as non-tax revenue increases to 10 per cent of GDP, and by almost 90 per cent when non-tax revenue increases to 20 per cent. The effect of changes in tax reliance is somewhat smaller (based on the random effects results): a country where tax revenue makes up 90 per cent of total government revenue (e.g. Burkina Faso) has a 3.3 per cent probability of transitioning to democracy in any given year, while that number drops to 1.1 per cent in countries where only 20 per cent of government revenue arrives through taxation (e.g. Equatorial Guinea or Azerbaijan).

These patterns are somewhat different for democracies, where changes in tax reliance have a larger impact than changes in non-tax revenue alone. This suggests that democratic countries are particularly likely to become autocracies when they combine high non-tax revenue with very weak tax collection. Thus, a democratic country with tax reliance of 0.9 (e.g. the Philippines), and other variables at their developing country means, has a less than 1 per cent likelihood of transitioning to autocracy in any given year, while an otherwise identical country with tax reliance of 0.2 (e.g. Nigeria) has an almost 7 per cent likelihood of becoming an autocracy. By contrast, a focus only on the role of non-tax revenue alone yields substantially smaller effects: a democratic country with non-tax revenue of 20 per cent of GDP (e.g. Nigeria), but an average level of tax collection, has a 2.2 per cent probability of transitioning to autocracy, while this value declines only modestly to 1.1 per cent in a country with non-tax revenue of only 2 per cent (e.g. Senegal).

The bottom line is that increased non-tax revenue and decreased tax reliance have very large impacts on the likelihood that a country will transition to, or remain, a democracy. This

[38] Note that these country examples are strictly to illustrate countries at these magnitudes of the revenue variables. The likelihood of any individual country experiencing a regime transition will be somewhat different from the averages reported here, owing to values on the control variables that differ from developing country averages as well as unobserved country-specific factors.

has important real world implications, as it implies that the recent discoveries of oil in a range of low-income countries carry with them substantial future governance risks that demand focused policy attention.

5 Robustness checks

Having presented the core results, we test their robustness to changes in the length of the lag on the independent variables, the use of alternative dependent variables, the use of a longer (but less complete) time series, the exclusion of different groups of countries and reliance on central government data only. We also subject our new data to the methods recommended by Haber and Menaldo (2011). In all cases the results prove robust to these alternatives, though useful additional insights emerge.

5.1 Alternative lag lengths

Ross (2004) and Andersen and Ross (2014) have argued that any causal impact of changes in the composition of government revenue will likely take longer than a single year to take effect, and have correspondingly argued for employing longer lags on the explanatory variables in order to more accurately model the causal relationship. Consistent with this logic, Tables 4a and 4b report results using one, three and five year lags on the independent variables for the Sys-GMM and logit estimations, respectively. A similar exercise would not make sense for the CCE-MG results, which already (and only) capture the long-run relationship.

Table 4a Effect of tax reliance, total tax and non-tax revenue on polity, using different lags

	(1) Sys-GMM two-step b/se	**(2) Sys-GMM two-step b/se**	**(3) Sys-GMM two-step b/se**	**(4) Sys-GMM two-step b/se**	**(5) Sys-GMM two-step b/se**	**(6) Sys-GMM two-step b/se**
	λ=1	λ=3	λ=5	λ=1	λ=3	λ=5
Lλ.Polity_s	0.8741*** (0.0359)	0.5991*** (0.0811)	0.5008*** (0.0915)	0.8612*** (0.0280)	0.5904*** (0.0768)	0.4316*** (0.1088)
Lλ.Tax_Rel	18.0684*** (5.3422)	19.4390** (7.9695)	23.8694** (10.0033)			
Lλ.TotTax				-11.7086 (15.5717)	7.0397 (31.2309)	22.2667 (39.0345)
Lλ.totnotax				-26.0380*** (8.3181)	-39.0786** (17.1580)	-35.7187 (32.7934)
Lλ.lgdp	-0.8896 (1.2806)	0.6934 (2.0321)	-1.6613 (2.4600)	-1.6612 (1.3640)	0.0153 (1.9998)	-0.5000 (2.6180)
AR(1)	0.000	0.150	0.316	0.000	0.208	0.436
AR(2)	0.648	0.470	0.390	0.650	0.420	0.285
Hansen	0.376	0.228	0.374	0.291	0.181	0.054
N	2579	2308	2024	2579	2308	2024
N_g	155.000	155.000	152.000	155.000	155.000	152.000

Notes: Columns (1), (4) use lags t-1; columns (2), (5) use lags t-3; columns (3), (6) use lags t-5. All regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation.

Source: ICTD GRD (2014).

The results follow the expected pattern. The Sys-GMM results using *tax_rel* become moderately larger and more significant over longer lag lengths. When we disaggregate

tax_rel into its component parts the same is generally true. *Tottax* is always insignificant, but becomes more positive as the lag grows longer, while *totnontax* become more negative and significant moving from a one-year to a three-year lag.

The only incongruous result is a loss of significance on the *totnontax* variable with a five-year lag, along with a sharp increase in the coefficient on *tottax*. However, both results should be discarded, as the Hansen test statistic indicates that the results are econometrically unreliable owing to the excessive proliferation of instruments – a common risk in relatively complex Sys-GMM specifications.

Table 4b Effect of tax reliance, total tax and non-tax revenue on regime, using random effects logit model and different lags

	(1) **Logit** **b/se** **λ=1**	**(2)** **Logit** **b/se** **λ=3**	**(3)** **Logit** **b/se** **λ=5**	**(4)** **Logit** **b/se** **λ=1**	**(5)** **Logit** **b/se** **λ=3**	**(6)** **Logit** **b/se** **λ=5**
Lλ.regime	-5.2888 (3.5544)	-18.1924*** (3.9408)	-25.7939*** (4.7138)	-0.6607 (2.8037)	-12.5082*** (3.2157)	-20.0787*** (3.9071)
Lλ.Tax_Rel	1.5100 (1.0490)	4.4561*** (1.6442)	7.2327*** (2.0780)			
Lλ.regimextaxrel	1.9235 (1.8691)	2.7075 (1.9330)	3.6250 (2.2226)			
Lλ.TotTax				-6.2755* (3.7480)	-2.3642 (5.0183)	7.8482 (6.7844)
Lλ.TotNonTax				-12.5968** (5.3654)	-16.4260*** (5.7189)	-20.2697*** (6.9277)
Lλ.regimextottax				16.7894** (8.1804)	12.3545 (8.1421)	3.7831 (9.8485)
Lλ.regimextotnontax				8.6665 (7.4713)	3.1397 (6.4659)	-2.0186 (6.9634)
Aggregate effects in democracies						
$\beta 1 + \beta 2$	3.4334*** (1.5399)	7.1636*** (2.1194)	10.8576*** (2.4668)			
$\beta 1 + \beta 3$				10.5138 (7.2802)	9.9903 (7.6084)	11.6313 (9.1243)
$\beta 2 + \beta 4$				-3.9303 (5.1777)	-13.2862** (5.7665)	-22.2882*** (6.8447)
Random effects model descriptors						
σ_μ	.0020783	2.372898	3.583723	.001475	2.110443	3.522656
ρ	1.31e-06	.631202	.7960776	6.61e-07	.575163	.7904406
N	2305	2025	1730	2305	2025	1730
N_g	157.000	156.000	152.000	157.000	156.000	152.000

Notes: Columns (1), (4) use lags t-1; columns (2), (5) use lags t-3; columns (3), (6) use lags t-5. All regressions include a standard set of control variables, which are excluded in order to conserve space: *lgdp*, *civil_war* and *growthpc*. * p<0.10, ** p<0.05, *** p<0.01. $\beta 1 + \beta 2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta 1 + \beta 3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta 1 + \beta 3$ captures the joint significance of *totnontax* and *regime x totnontax*.

Source: ICTD GRD (2014).

The results for the logit estimations are similarly consistent with expectations, as our explanatory variables universally become larger and more significant as the lag length is increased for both autocracies and democracies. Interestingly, with a five-year lag we find

evidence that non-tax revenue may increase the likelihood of autocratic transitions in existing democracies, though this finding is not robust to employing fixed effects.[39]

The fact that most of the results become larger and more significant over longer lags is consistent with the dynamics of the model, as the one-year effects reported earlier should be magnified over multiple years. More notable are the cases where the sign on the results is reversed over a longer lag, as is the case for the *tottax* variable in both models, as this may indicate that the longer lags better capture the causal process of interest, as initially proposed by Ross (2004).

5.2 Alternative dependent variables

We also test the robustness of our results to employing alternative measures of democracy. To begin we run our Sys-GMM and CCE-MG results when employing several other common measures: the Democracy measure from the *International Country Risk Guide* (*icrg_dem*), the Freedom House measure of democracy (*fh_dem*), and the Accountability measure from the World Bank's *World Governance Indicators* (*wgi_acc*).

Tables 5a and 5b report results using Sys-GMM, employing multiple lag lengths. The diagnostic statistics indicate invalid instrument and/or persistent second order autocorrelation in several specifications, and whether the results are econometrically valid is signalled in a row at the bottom of both tables. The results generally lose significance in these (invalid) specifications. By contrast, the *tax_rel* variable is positive and significant in four of the five valid specifications (employing *wgi_acc* and *fh_dem* as dependent variables), while falling just short of significance in the other. Meanwhile, there are only three valid specifications in Table 5b, but the *totnontax* variable is negative and significant in all of them, while the *tottax* variable is always insignificant.

[39] Fixed effects results available on request.

Table 5a Robustness check using tax_rel and different measures of governance

	(1) Sys-GMM two-step b/se	(2) Sys-GMM two-step b/se	(3) Sys-GMM two-step b/se	(4) Sys-GMM two-step b/se	(5) Sys-GMM two-step b/se	(6) Sys-GMM two-step b/se	(7) Sys-GMM two-step b/se	(8) Sys-GMM two-step b/se	(9) Sys-GMM two-step b/se
	ICRG_Dem	ICRG_Dem	ICRG_Dem	WGI_Acc	WGI_Acc	WGI_Acc	Fh_dem	Fh_dem	Fh_dem
	λ=1	λ=3	λ=5	λ=1	λ=3	λ=5	λ=1	λ=3	λ=5
Lλ.DV	0.9779***	0.7356***	0.4333***	0.8902***	0.3752***	0.2421**	0.8789***	0.6537***	0.4538***
	(0.0353)	(0.0683)	(0.0907)	(0.0574)	(0.1061)	(0.1178)	(0.0294)	(0.0598)	(0.0873)
Lλ.Tax_Rel	1.1614	0.8268	2.0694***	0.3367**	1.1057***	1.0033**	0.3488	1.7766***	1.9167***
	(0.8272)	(0.6878)	(0.7894)	(0.1479)	(0.2912)	(0.4261)	(0.2269)	(0.4295)	(0.6153)
Lλ.lgdp	-0.0525	0.1197	0.3726	-0.0264	-0.0167	0.0874	0.0426	0.1367	0.0229
	(0.1913)	(0.2189)	(0.3724)	(0.0710)	(0.1005)	(0.1412)	(0.0802)	(0.1495)	(0.2738)
AR(1)	0.000	0.000	0.000	0.000	0.862	0.057	0.000	0.786	0.765
AR(2)	0.000	0.000	0.124	0.007	0.297	0.940	0.993	0.243	0.523
Hansen	0.030	0.055	0.014	0.036	0.330	0.447	0.389	0.318	0.106
Econometrically valid?	No	No	No	No	Yes	Yes	Yes	Yes	Yes
N	2228	1989	1737	1321	1165	991	2308	2061	1804
N_g	135.000	135.000	133.000	178.000	178.000	177.000	139.000	139.000	137.000

Notes: Columns (1), (4), (7) use lags t-1; columns (2), (5), (8) use lags t-3; columns (3), (6), (9) use lags t-5. All regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation.

Source: ICTD GRD (2014).

Table 5b: Robustness check using tottax, totnontax and different measures of governance

	(1) Sys-GMM two-step b/se ICRG_Dem λ=1	(2) Sys-GMM two-step b/se ICRG_Dem λ=3	(3) Sys-GMM two-step b/se ICRG_Dem λ=5	(4) Sys-GMM two-step b/se WGI_Acc λ=1	(5) Sys-GMM two-step b/se WGI_Acc λ=3	(6) Sys-GMM two-step b/se WGI_Acc λ=5	(7) Sys-GMM two-step b/se Fh_dem λ=1	(8) Sys-GMM two-step b/se Fh_dem λ=3	(9) Sys-GMM two-step b/se Fh_dem λ=5
Lλ.DV	0.9813***	0.7108***	0.4876***	0.9404***	0.5551***	0.3178**	0.8865***	0.6595***	0.5061***
	(0.0317)	(0.0719)	(0.0974)	(0.0386)	(0.1558)	(0.1249)	(0.0264)	(0.0654)	(0.0859)
Lλ.TotTax	-1.5626	-0.8780	1.0686	-0.5508	0.5688	2.7895	-0.0216	3.5516**	2.0160
	(1.0976)	(2.0444)	(3.0846)	(0.3933)	(0.8148)	(1.7370)	(0.7805)	(1.7013)	(2.6972)
Lλ.totnotax	-0.8375*	-1.2664	-2.0345	-0.1348	-1.5489**	-1.1812*	-0.6017**	-0.7055	-2.5607**
	(0.4506)	(0.9706)	(1.3170)	(0.1795)	(0.6170)	(0.6224)	(0.2924)	(0.7480)	(1.0167)
Lλ.lgdp	-0.0077	0.0647	0.0803	0.0135	-0.0591	0.0401	0.0258	-0.0448	0.0197
	(0.1210)	(0.2065)	(0.3139)	(0.0280)	(0.0729)	(0.1458)	(0.0881)	(0.1825)	(0.2868)
AR(1)	0.000	0.000	0.000	0.000	0.548	0.102	0.000	0.588	0.819
AR(2)	0.000	0.000	0.134	0.003	0.793	0.956	0.956	0.299	0.556
Hansen	0.038	0.030	0.009	0.050	0.152	0.586	0.410	0.053	0.011
Econometrically Valid?	No	No	No	No	Yes	Yes	Yes	No	No
N	2228	1989	1737	1321	1165	991	2308	2061	1804
N_g	135.000	135.000	133.000	178.000	178.000	177.000	139.000	139.000	137.000

Notes: Columns (1), (4), (7) use lags t-1; columns (2), (5), (8) use lags t-3; columns (3), (6), (9) use lags t-5. All regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation.

Source: ICTD GRD (2014).

By contrast, we find that the CCE-MG results (not reported in order to conserve space) are generally not robust to the alternative governance measures, with only one significant result: a negative and significant effect of *totnontax* on the Freedom House measure of democracy (*fh_dem*). While we view the polity measure as preferable, owing to its greater objectivity and completeness, this again is a useful indication of the potential sensitivity of the results.

Finally, we re-run our logit estimations when employing an alternative binary measure of regime type (*regime_polity_60*), constructed by coding countries as democracies if they achieve a scaled polity score of 60 out of 100 or above.[40] The results reported in Table 5c, which employ the random effects specification, are broadly in line with those reported so far, though modestly stronger and more consistent. For autocracies, the *tax_rel* variable is positive and significant for one-, three- and five-year lags, while the *totnontax* variable is always negative and significant. Meanwhile, *tax_rel*, *tottax* and *totnontax* are each positive in democracies, achieving significance for some, but not all, lag lengths. The significance of the tax variables is consistent with previous results, while the positive significance of *totnontax* stands in contrast to the results with *regime* as the dependent variable, and is a further indication of the sensitivity of the results related to the impact of non-tax revenue in democracies.

[40] This is equivalent to a score of 2 or above (on a scale -10 to 10) on the polity2 variable as coded in the *Polity IV* dataset.

Table 5c Effect of tax reliance, total tax and non-tax revenue on polity_60 using a random effects logit model and different lags

	(1) Logit b/se λ=1	(2) Logit b/se λ=1	(3) Logit b/se λ=3	(4) Logit b/se λ=3	(5) Logit b/se λ=5	(6) Logit b/se λ=5
Lλ.regime_polity_60	-3.4330	-1.9006	-9.6984***	-5.8873**	-12.1386***	-5.8773**
	(2.4541)	(1.9198)	(2.8713)	(2.4150)	(3.4365)	(2.9226)
Lλ.Tax_Rel	1.3228*		3.0893**		3.1148**	
	(0.7883)		(1.1998)		(1.4051)	
Lλ.polity60xtaxrel	0.2630		0.4985		3.4687**	
	(1.3299)		(1.4103)		(1.5338)	
Lλ.TotTax		-0.7704		2.3706		5.1719
		(2.6391)		(4.4601)		(5.8315)
Lλ.TotNonTax		-6.2807**		-10.9171***		-8.1569*
		(3.1762)		(3.8958)		(4.1731)
Lλ.polity60xtottax		6.3995		17.7123***		19.0934***
		(4.5516)		(6.1669)		(7.3766)
Lλ.polity60xtotnontax		6.3582		11.3208**		2.7649
		(4.0669)		(4.8610)		(5.0596)
Aggregate effects in democracies						
$\beta 1 + \beta 2$	1.5858		3.5878***		6.5835***	
	(1.0715)		(1.3910)		(1.5206)	
$\beta 1 + \beta 3$		5.6291		20.0828***		24.2653***
		(3.7123)		(6.0793)		(6.9622)
$\beta 2 + \beta 4$		0.0774		0.4037		-5.3920
		(2.5364)		(3.8963)		(4.3862)
Random effects model descriptors						
σ_μ	.0060889	.0039629	2.701151	2.848901	4.108202	4.30622
ρ	.0000113	4.77e-06	.6892271	.7115695	.8368703	.8493195
N	2571	2571	2305	2305	2018	2018
N_g	157.000	157.000	157.000	157.000	156.000	156.000

Notes: Columns (1), (2) use lags t-1; columns (3), (4) use lags t-3; columns (5), (6) use lags t-5. All regressions include a standard set of control variables, which are excluded in order to conserve space: lgdp, civil_war and growthpc.* p<0.10, ** p<0.05, *** p<0.01. $\beta 1 + \beta 2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta 1 + \beta 3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta 2 + \beta 4$ captures the joint significance of *totnontax* and *regime x totnontax*.

Source: ICTD GRD (2014).

5.3 Sensitivity to alternative time periods

Our core analysis relies on data from 1990 onwards, during which data is both more complete and reliable.[41] However, as a robustness check we have experimented with extending the analysis year-by-year back to 1980, and the results are generally robust in all cases.

When employing the Sys-GMM estimators (reported in Tables A1a and A1b in the Appendix) we find a positive coefficient on *taxrel* for every starting year back to 1982, though the results

[41] In 1990 the number of complete country-year observations that can be employed in the analysis is 89 – already significantly smaller than the peak of 140 covered in 1998, though the latter includes the availability of data for a large number of post-Soviet states following the fall of the Soviet Union. The number of annual observations declines to 76 in 1985, and to 54 in 1980. Notably, the results are also robust to employing a shorter time series beginning in the mid-1990s, when the data is most complete.

fall slightly short of significance when data is extended back to 1980 and 1981. The negative coefficient on *totnontax* remains significant as early as 1983, but is again slightly short of significance for earlier start dates.

The pattern is similar for the CCE-MG results, reported in Tables A1c and A1d: there is a positive coefficient on *tax_rel* for the majority of start dates, while the positive coefficient on *tottax* is robust for every start date. The *totnontax* variable is somewhat less robust, remaining significant for start dates back to 1987, but losing significance (though retaining its sign and magnitude) for earlier years. This is all consistent with our earlier results, while the modest fall in significance with longer time series may reflect either the importance of having relatively complete data or the strengthening of the resource curse over time, as suggested by Andersen and Ross (2014).[42]

When we pursue the same exercise employing the logit model, reported in Tables A1e and A1f, we again find that the results are robust to extending the sample back in time. All the results become somewhat more significant and larger in magnitude as the time series is extended, which is likely explained by the benefits of increasing the number of regime transitions in the analysis. Notably, this increased significance with a longer time series applies equally to the *tottax* variable, with the results signalling that increased tax revenue reduces the likelihood of transition from autocracy to democracy, but has a positive impact in reinforcing democracy in already democratic states. However, as in the earlier results, this pattern disappears in the fixed effects results (not reported here), in which *tottax* is positive but insignificant in both autocracies and democracies.

5.4 Sensitivity to sample selection

We likewise explore the sensitivity of our results to excluding particular groups of countries. For each of our estimation methods, we begin by excluding the eleven members of OPEC from the analysis, as well as adding recently-withdrawn Ecuador and Venezuela to the group of excluded countries. We subsequently exclude, in turn, countries from the Middle East and North Africa (MENA) region and Eastern Europe and Central Asia (EECA), both as defined by the World Bank. Both are home to significant resource producers, while the latter has undergone major governance changes since the collapse of the Soviet Union. Finally, we also re-estimate our results when excluding OECD countries, and when excluding countries with populations below either 250,000 or 1,000,000.

Essentially all the results, reported in Tables A2a through A2i in the Appendix, are fully robust to these changes in the sample. In general the results are very modestly weaker when we exclude major groups of oil producers, with a handful of coefficients falling just below the cut-off for statistical significance at conventional levels. However, these changes are both small and in the minority. Meanwhile, the results are almost entirely unchanged when the OECD and small countries are excluded, if anything becoming modestly more significant. This robustness to the exclusion of OECD countries is particularly important, as it addresses concerns that our findings may be driven merely by the correspondence of high levels of taxation, low levels of non-tax revenue and high levels of democracy in OECD states.

[42] They argue for a key turning point after 1980, following the two global oil crises and the nationalisation of oil supplies in many countries, but it is plausible that the resource curse may also have grown progressively stronger through the 1980s.

5.5 Reliance on central government data only

A key innovation of the ICTD *Government Revenue Dataset* is to combine data for general government for more decentralised states with central government data for more centralised countries. As explained earlier, this is an attempt to account for the potentially severe underestimation of revenue collection in federal states when relying only on central government data, and to expand data coverage. To ensure robustness we also test our results using exclusively central government data, reported in Table 6. The results are generally robust to this alternative, but the more significant message is that the core results across all three estimation models are more significant, larger in magnitude and more in line with expectations when employing the joint central and general government dataset. This appears to reinforce further the analytical value of relying on higher quality and more complete data.

Table 6 Effect of tax reliance, total tax and total non-tax on regime using Sys-GMM, CCE-MG and RE logit, central government only

	(1) Sys-GMM CenGen λ=1	(2) CCE-GM w/o trend CenGen λ=0	(3) Sys-GMM Central λ=1	(4) CCE-GM w/o trend Central λ=0	(5) Sys-GMM CenGen λ=1	(6) CCE-GM w/o trend CenGen λ=0	(7) Sys-GMM Central λ=1	(8) CCE-GM w/o trend Central λ=0	(9) Logit CenGen λ=1	(10) Logit CenGen λ=1	(11) Logit Central λ=1	(12) Logit Central λ=1
Lλ.Tax_Rel	18.0684***	14.3193**	12.2450**	17.8429					1.6350		1.6217	
	(5.3422)	(7.2816)	(5.2651)	(12.3272)					(1.1342)		(1.0944)	
Lλ.regimextax_rel									3.9285*		2.9702	
									(2.2339)		(2.2010)	
Lλ.TotTax					-11.7086	73.7190**	-26.2176*	166.7257		-6.2755*		-5.0117
					(15.5717)	(30.8115)	(14.3801)	(117.9682)		(3.7480)		(3.6794)
Lλ.TotNonTax					-26.0380***	-137.2167**	-27.5514**	-80.0266		-12.5968**		-11.8297**
					(8.3181)	(57.2205)	(10.8562)	(91.8027)		(5.3654)		(5.4811)
Lλ.regimextottax										16.7893**		9.0804
										(8.1803)		(7.4974)
Lλ.regimextotnontax										8.6664		-6.3171
										(7.4712)		(9.7615)
Aggregate effects in democracies												
$\beta1 + \beta2$									5.5635***		4.5919**	
									(1.9740)		(1.9135)	
$\beta1 + \beta3$										10.5138		4.0686
										(7.2802)		(6.5666)
$\beta2 + \beta4$										-3.9303		-18.1467**
										(5.1777)		(8.1166)
Sys-GMM diagnostic statistics												
AR(1)	0.000		0.000		0.000		0.000					
AR(2)	0.648		0.660		0.650		0.668					
Hansen	0.376		0.697		0.291		0.269					
Random effects model descriptors												
σ_μ									.0009771	.001475	.0007979	.0014336
ρ									2.90e-07	6.61e-07	1.93e-07	6.25e-07
N	2579	2682	2502	2582	2579	2674	2502	2578	2175	2305	2255	2255
N_g	155.000	154.000	148.000	146.000	155.000	152.000	148.000	145.000	150.000	157.000	150.000	150.000

Notes: Dataset merging central and general government data employed in columns (1), (2), (5), (6), (9) and (10). Dataset using only central government data employed in columns (3), (4), (7), (8), (11) and (12). Sys-GMM and CCE-MG regressions include a control for *lgdp*, not reported in order to conserve space. Logit regressions include a standard set of controls, not reported here to conserve space: *lgdp*, *civil_war* and *growthpc*. All independent variables – *polity*, *tax_rel*, *tottax* and *totnontax* - are contemporaneous for CCE-MG estimates, and lagged t-1 for GMM and Logit estimates. All GMM regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. $\beta1 + \beta2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta1 + \beta3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta2 + \beta4$ captures the joint significance of *totnontax* and *regime x totnontax*. Source: ICTD GRD (2014).

5.6 Applying ECM methods

Finally, we conclude our robustness tests by subjecting our new data to the ECM regressions favoured by the widely-cited work of Haber and Menaldo (2011). ECM regressions provide estimates of both the short-run effect of a change in the independent variable on the dependent variable, and the long-run equilibrium relationship between the two. The dependent variable in an ECM is the change or difference of the variable of interest, so in our case it is the change in *Democracy*, denoted

$$\Delta_n Democracy_t = Democracy_t - Democracy_{□-n}.$$

This gives us the following specification:

$$\begin{aligned}\Delta_n Democracy_{i,t} &= \alpha Democracy_{i,t-n} + \beta \Delta\, Revenue_{i,t-n} + \gamma\left(Democracy_{i,t-n} - \delta Revenue_{i,t-n}\right) \\ &+ X'_{i,t-n}\rho + u_{i,t} \qquad [5]\end{aligned}$$

where *X* is a vector of control variables, *Democracy* is our continuous measure based on the polity dataset, and *Revenue* represents our three distinct revenue variables. The time horizon of the short-run effect is defined by *n*. As explained in the text, the econometric drawback of this method relative to the CCE-MG estimator is that it does not allow for heterogeneous coefficients or for cross-section dependence. We run these regressions for the sake of completeness.

To do so, we first replicate Haber and Menaldo's (2011) results, to ensure that we are using precisely the same specification. We then substitute our comparatively complete panel in place of their data on fiscal reliance, which covers only eighteen resource-dependent states. The results are reported in Table 7. The differenced independent variables represent the short-run effects, while the lagged dependent and independent variables are used to calculate the long-run effect, which is also reported in the table. We find that the *tax_rel* variable is positive, but not significant either in the short run or the long run. However, when we break it into its component parts we find that non-tax revenue (*totnontax*) has a negative and statistically significant effect on accountability in the long run, though with a smaller magnitude than the earlier ECM and Sys-GMM results, while tax revenue (*tottax*) is insignificant. Surprisingly, the significance is reversed in the short run, with non-tax revenue insignificant and tax revenue significant and positive. Overall, the results remain consistent with the existence of a political resource curse, but also again point to the sensitivity of specific findings to changes in specification.

Table 7 Replication of Haber and Menaldo's ECM

	(1) ECM b/se	(2) ECM b/se	(3) ECM b/se	(4) ECM b/se
L.polity_s	-0.2786***	-0.2806***	-0.2778***	-0.2800***
	(0.0266)	(0.0269)	(0.0265)	(0.0268)
L.Fiscal Rel	2.5288	2.9024		
	(2.4422)	(2.7259)		
D.Fiscal Rel (short-run effect)	5.1847	5.5581		
	(4.1515)	(4.5820)		
Fiscal_Rel long-run effect	9.0771	10.3429		
	(9.3425)	(10.4501)		
L.tottax			-8.3551	-9.6287
			(7.6739)	(6.9788)
L.totnontax			-7.0160***	-10.0829***
			(2.6306)	(3.3987)
D.tottax (short-run effect)			33.7072***	35.4939***
			(11.2430)	(11.9656)
D.totnontax (short-run effect)			-2.1321	-3.1885
			(4.4614)	(4.9855)
Tottax long-run effect			-30.0720	-34.3897
			(29.8489)	(26.8660)
Totnontax long-run effect			-25.2522***	-36.0118***
			(9.7649)	(13.2528)
N	2648	2539	2648	2539
N_g	158.000	153.000	158.000	153.000
r2	0.199	0.209	0.203	0.214

Notes: ECM regressions follow Haber and Menaldo (2011). Interpolation is used for missing data to ensure there are no gaps in time series. Regressions use Driscoll–Kraay standard errors. Robust t-statistics are in brackets. LRM standard errors estimated using the delta method: −1(b(Fiscal Reliance t − 1)/b(Polity t − 1)). Residuals were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend).* p<0.10, ** p<0.05, *** p<0.01.

Source: ICTD GRD (2014).

6 Discussion

Our central purpose has been to re-examine existing tests of the political resource curse hypothesis using theoretically more appropriate independent variables which focus on the extent of government revenue from tax and non-tax sources. By drawing on the new ICTD GRD, and employing several complementary econometric approaches from the literature, this paper provides more theoretically and empirically robust tests of the existence and extent of the political resource curse.

Overall, we find strong evidence across all the estimation strategies for the existence of a political resource curse. Across the vast majority of the tests we find *both* a statistically significant positive relationship between tax reliance and democracy *and* a statistically significant negative relationship between total non-tax revenue and democracy. While one or the other variable occasionally falls slightly short of standard thresholds of statistical significance, at least one of the two is significant in virtually every model specification and robustness test across each of the Sys-GMM, CCE-MG, logit and ECM estimations, with the lone exception coming when we add highly incomplete data from 1980 and 1981. Consistent with prior arguments by Ross (2004) and Andersen and Ross (2014), the strength of these results increases as the length of the lag on the revenue variables increases. The broad results are robust to alternative control variables, alternative measures of democracy, the

inclusion of additional lags, the extension of the time series, systematically excluding particular groups of countries or employing only central government data.

The magnitude of the effects is in every case very substantial, but also intuitively plausible. The Sys-GMM and CCE-MG results indicate that over the long term a country with non-tax revenue at the level of Nigeria (average of 22 per cent of GDP) is expected to have a 31 to 43 percentage point lower level of democracy than an otherwise similar country with non-tax revenue at the level of Senegal (average of 1.5 per cent of GDP). They further suggest that the level of democracy is expected to be 60 per cent to 80 per cent lower – that is, almost at the bottom of the polity scale – in cases of extreme resource wealth of 40 per cent of GDP or more (e.g. Angola and Bahrain).

The results of the logit regressions are expressed in terms of the likelihood of transitioning to or from democracy, but tell a comparable story. The likelihood of a transition to democracy is almost three times larger (4.2 per cent vs. 1.6 per cent) in a country with very low levels of non-tax revenue, as compared to a country with even 10 per cent of GDP coming from non-tax sources. While the estimates become more imprecise at higher values of non-tax revenue, the results suggest that democratic transitions become almost 90 per cent less likely when non-tax revenue shifts from 2 per cent of GDP to 20 per cent of GDP, while becoming almost impossible at levels of non-tax revenue on par with Angola, Kuwait, Saudi Arabia or Bahrain. The implication is that recent discoveries of oil in several low-income countries pose a significant threat to democracy.

While the results overall offer strong support for the existence of a political resource curse, the results also reveal significant sensitivity to changes in econometric estimation strategy and model specification. Intuitively, this sensitivity is not surprising: it is difficult to capture complex political relationships using cross-country econometric methods, as illustrated most vividly in evidence of the sensitivity of results linking income and democracy (Acemoglu et al. 2008). Highlighting this sensitivity has correspondingly been an explicit goal of this research, and underpins our reliance on multiple econometric approaches.

However, the sensitivity of the results also holds potential lessons. The sensitivity of the results is most vividly apparent in the weak results when employing the Diff-GMM estimator (Table 1a) and, to a lesser extent, the ECM estimator (Table 7). In seeking to explain these results a recent review article by Ross (forthcoming) is instructive, as he argues that theories of the political resource curse are largely theories about long-term effects in levels (i.e. over the long term countries with high levels of resource wealth are likely to be less democratic), rather than about year-on-year changes (i.e. increases in resource wealth cause relatively immediate declines in democracy, and vice versa). Consistent with this contention, both the ECM and Diff-GMM estimators rely on first differences and yield comparatively weak results, whereas we find much stronger results employing Sys-GMM, CCE-MG and logit models that rely on changes in levels of the explanatory variables.

The more cross-cutting sensitivity of the results relates to the total tax variable, which tends towards being positive but ranges from positive significance to negative significance, while most often being insignificant, depending on the model and specification. This is true even as the *tax_rel* and *totnontax* variables remain relatively consistent. The simplest conclusion to be drawn is that tax revenue has little consistent impact on levels of democracy. However, an equally plausible conclusion based on the data is that the relationship between taxation and democracy is complex and non-linear.

The first source of support for such a reading of the evidence lies in the more positive and sometimes significant relationship between *tottax* and democracy using the CCE-MG estimator. The obvious econometric explanation for the stronger result on *tottax,* as compared to the Sys-GMM estimates, lies in the fact that the CCE-MG model is not dynamic,

and does not attempt to control for reverse causation. As a result, it captures the bi-directional relationship between *tottax* and *democracy*, while it is not constrained by assumptions about the dynamic structure of any such relationship. Both are likely important.

Most basically, research elsewhere has suggested a positive impact of democracy on tax revenue, which should inflate the coefficient on *tottax* in the CCE-MG model (Timmons 2010). Less obviously, recent research suggests that the temporal structure of the relationship between taxation and democracy may differ from that implied by the model specification employed here, and in the literature more broadly. As initially described by Ross (2004), the use of lagged variables in the core specification in part reflects a belief that, if taxation is a cause of expanded democracy, it is 'sensible to [assume] that a change in the independent variables (taxes) should precede a change in the dependent variable [democracy]' by the length of the lag (Ross 2004: 238). However, this may be a poor description of the empirical reality.

On one hand, any impact of taxation on democracy may occur over the long term, through subtle changes in levels of political engagement among increasingly burdened taxpayers. Indeed, consistent with this view, the *tottax* variable becomes positive, though still slightly short of positive significance, when a longer lag is employed in the Sys-GMM results. As importantly, and as recognised by Ross (2004), even if taxation is the cause of increased democracy, changes in democracy may in practice precede changes in taxation – or taxation may decline rather than increase prior to increases in democracy. This will be true, for example, if a government chooses to first expand democracy in an effort to facilitate subsequent tax collection. It will similarly be true if expanded democracy is a response to declining tax collection, spurred by successful resistance to taxation by taxpayers (Prichard 2010). Put more simply, the mixed results for *tottax* are at least consistent with there being a poor fit between the dynamic structure of the model and the actual dynamics of the causal processes linking taxation and democracy.

The second indication of a potentially more complex relationship between tax and democracy comes from the logit regressions. As a starting point, the fixed effects version of the results, though relying on a relatively small sample, suggests a positive impact of tax collection on the likelihood of democracy. By contrast, the random effects results reveal no effect of total tax collection on the likelihood of democratisation in autocracies. However, the random effects results indicate a positive, and often significant, impact of tax collection on the likelihood of existing democracies remaining democratic. Taken together, the evidence suggests an ambiguous relationship between taxation and democracy in autocracies, but, tentatively, a positive impact of taxation on democracy in existing democracies.

Such a conclusion is consistent with theory. Any relationship between taxation and democracy rests on the ability of taxpayers to engage in collective action, to resist unpopular government taxation and to bargain for reciprocal concessions from governments. There are clear reasons to expect these processes to be more likely in democracies, where collective action is eased and there are institutionalised spaces for bargaining between citizens and governments. While strong conclusions are impossible given the unevenness of the results, the logit results echo emerging arguments elsewhere that tax-democracy links are not guaranteed, and that context is likely to condition particular outcomes (Prichard 2010).

That said, the logit results also highlight a further risk in interpreting the results on *tottax*, which has been overlooked in earlier results. Simply, while *tottax* captures the burden of taxation on citizens and businesses, it may also be acting as a proxy for broader state capacity.[43] If this is the case then we would expect *tottax* to have a positive impact on democracy in democracies, as it would be indicative of strong state capacity and stability.

[43] Indeed, there is a long tradition of studies employing tax collection as a indicator of state capacity (Lieberman 2002).

Likewise, we would expect *tottax* to have an insignificant or negative impact in autocracies, as it would be a signal of the relative strength and institutionalisation of individual autocratic governments. The bottom line is, again, that while our results on the *tottax* variable are comparatively mixed, this may be best understood as a reflection of the substantial complexity of the relationship. However, these possibilities are merely speculative given the sensitivity of the results, and this appears to be an area in which complementary research strategies are particularly needed.

7 Conclusions

The premise of this paper has been straightforward: despite expanding econometric tests of the political resource curse, and growing econometric complexity, these tests have been undermined by their reliance on theoretically imperfect explanatory variables, which have in turn been driven by the stark weaknesses of cross-country government revenue data. As with many areas of development research, the most useful strategy for generating improved results thus lies not in more complex methods, or more complex hypotheses, but in improving the quality of underlying data, testing more theoretically precise propositions, and subjecting the results to a variety of complementary econometric tests.

In doing so we find what we believe to be the most robust existing evidence of a political resource curse. We demonstrate, across a range of methods, that the magnitude of these effects is large, making the composition of government revenue a highly influential determinant of governance outcomes across countries. However, while the broad pattern of results is compelling, we also seek explicitly to highlight the sensitivity of these results at the margin. This sensitivity is unsurprising, given the obvious challenges of capturing complex political relationships through cross-country econometric methods. However, the sensitivity also holds an important message: while cross-country data offers valuable insights into the broad pattern of relationships, more nuanced claims based on aggregate data are at best indicative and should be treated with caution. More in-depth understanding correspondingly demands that cross-country results be paired with evidence from alternative, and complementary, research strategies.

Appendix

Table A1a Effect of tax reliance on polity using Sys-GMM and different time periods

	(1) Sys-GMM 1990 b/se	(2) Sys-GMM 1989 b/se	(3) Sys-GMM 1988 b/se	(4) Sys-GMM 1987 b/se	(5) Sys-GMM 1986 b/se	(6) Sys-GMM 1985 b/se	(7) Sys-GMM 1984 b/se	(8) Sys-GMM 1983 b/se	(9) Sys-GMM 1982 b/se	(10) Sys-GMM 1981 b/se	(11) Sys-GMM 1980 b/se
L.Polity_s	0.8604***	0.8896***	0.8964***	0.9004***	0.9065***	0.9219***	0.9276***	0.9271***	0.9297***	0.9376***	0.9373***
	(0.0351)	(0.0290)	(0.0220)	(0.0217)	(0.0219)	(0.0191)	(0.0176)	(0.0175)	(0.0171)	(0.0152)	(0.0154)
L.Tax_Rel	16.9375***	15.2916***	14.2645***	12.0857**	11.5332***	10.7277**	11.1559**	10.5812*	11.2088***	7.0901	4.1392
	(5.7279)	(4.4339)	(4.3247)	(4.7667)	(4.3362)	(4.8447)	(5.3443)	(5.4614)	(4.2611)	(5.6027)	(3.8638)
L.lgdp	-0.4163	0.5821	0.4882	0.9248	0.4146	1.0053	-0.2884	0.2135	0.7370	0.7816	0.4201
	(0.7412)	(0.6164)	(1.0474)	(1.0032)	(1.0649)	(0.8691)	(1.2686)	(1.0178)	(0.9313)	(0.7552)	(0.7349)
AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.637	0.315	0.314	0.308	0.339	0.414	0.383	0.280	0.278	0.273	0.261
Hansen	0.295	0.542	0.754	0.358	0.104	0.337	0.604	0.112	0.198	0.589	0.505
N	2579	2662	2747	2827	2904	2980	3047	3112	3176	3234	3286
N_g	155.000	155.000	155.000	155.000	155.000	155.000	155.000	155.000	155.000	155.000	155.000

Notes: All regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation.

Source: ICTD GRD (2014).

Table A1b Effect of total tax and non-tax revenue on polity using Sys-GMM and different time periods

	(1) Sys-GMM 1990 b/se	(2) Sys-GMM 1989 b/se	(3) Sys-GMM 1988 b/se	(4) Sys-GMM 1987 b/se	(5) Sys-GMM 1986 b/se	(6) Sys-GMM 1985 b/se	(7) Sys-GMM 1984 b/se	(8) Sys-GMM 1983 b/se	(9) Sys-GMM 1982 b/se	(10) Sys-GMM 1981 b/se	(11) Sys-GMM 1980 b/se
L.Polity_s	0.8499***	0.8497***	0.8630***	0.8838***	0.8921***	0.8981***	0.9066***	0.9090***	0.9124***	0.9167***	0.9225***
	(0.0292)	(0.0298)	(0.0254)	(0.0205)	(0.0217)	(0.0211)	(0.0183)	(0.0185)	(0.0174)	(0.0182)	(0.0196)
L.TotTax	-16.3798	-10.2521	-6.8171	-7.5960	-13.5660	-2.6619	-8.2840	-6.8054	-3.1735	-9.9313	-13.0366
	(12.5232)	(13.0362)	(11.2002)	(9.7204)	(11.6538)	(12.1827)	(10.2273)	(10.6147)	(10.6738)	(11.3914)	(12.8338)
L.totnotax	-21.0339***	-15.0837**	-16.1040**	-13.8625**	-14.2562*	-10.6663	-11.6652	-11.8753*	-7.9345	-10.2843	-9.4276
	(7.6095)	(7.2206)	(7.8240)	(6.7499)	(8.1947)	(7.3076)	(7.2745)	(6.3943)	(5.2557)	(7.9388)	(9.1250)
L.lgdp	-0.4988	0.2019	-0.0575	0.5831	0.9205	1.0370	0.3972	0.7382	0.5957	0.7879	0.7930
	(0.8171)	(0.8823)	(1.0595)	(0.8255)	(1.1198)	(1.0051)	(1.2532)	(0.9780)	(0.8923)	(1.0283)	(0.8083)
AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.645	0.342	0.307	0.319	0.348	0.394	0.372	0.267	0.265	0.273	0.270
Hansen	0.264	0.216	0.136	0.517	0.077	0.210	0.377	0.491	0.518	0.102	0.191
N	2579	2662	2747	2827	2904	2980	3047	3112	3176	3234	3286
N_g	155.000	155.000	155.000	155.000	155.000	155.000	155.000	155.000	155.000	155.000	155.000

Notes: All regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation.

Source: ICTD GRD (2014).

Table A1c Effect of tax reliance on polity using CCE-MG and different time periods

	(1) CCE-MG 1990 b/se	(2) CCE-MG 1989 b/se	(3) CCE-MG 1988 b/se	(4) CCE-MG 1987 b/se	(5) CCE-MG 1986 b/se	(6) CCE-MG 1985 b/se	(7) CCE-MG 1984 b/se	(8) CCE-MG 1983 b/se	(9) CCE-MG 1982 b/se	(10) CCE-MG 1981 b/se	(11) CCE-MG 1980 b/se
Tax_Rel	14.3193**	13.5314*	11.3142	8.5969	9.2579	10.9711*	12.6872**	13.2134**	13.1117*	11.2743*	9.2019
	(7.2816)	(7.0519)	(6.9403)	(6.4065)	(6.0321)	(5.8037)	(6.1401)	(6.4742)	(6.8197)	(6.4473)	(6.1377)
lgdp	9.6114*	10.3129**	11.1428**	11.8789**	11.7053**	12.1287**	11.7679**	11.1703**	9.9456*	8.6399*	9.5346**
	(5.1366)	(5.1814)	(5.0632)	(5.1639)	(5.2673)	(5.2596)	(5.3140)	(5.2471)	(5.1533)	(4.8592)	(4.7225)
N	2682	2768	2853	2933	3010	3086	3153	3218	3282	3340	3392
N_g	154.000	154.000	154.000	154.000	154.000	154.000	154.000	154.000	154.000	154.000	154.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. Residuals of the CCE-MG regressions were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend).

Source: ICTD GRD (2014).

Table A1d Effect of total tax and non-tax revenue on polity using CCE-MG and different time periods

	(1) CCE-MG 1990 b/se	(2) CCE-MG 1989 b/se	(3) CCE-MG 1988 b/se	(4) CCE-MG 1987 b/se	(5) CCE-MG 1986 b/se	(6) CCE-MG 1985 b/se	(7) CCE-MG 1984 b/se	(8) CCE-MG 1983 b/se	(9) CCE-MG 1982 b/se	(10) CCE-MG 1981 b/se	(11) CCE-MG 1980 b/se
TotTax	73.7190**	81.0666***	79.8498***	63.1868**	57.0215**	60.1626**	64.8290**	55.8214**	52.9791**	44.7606*	43.8241*
	(30.8115)	(28.5882)	(26.6159)	(25.6068)	(24.3854)	(25.0955)	(25.4202)	(25.4878)	(25.5259)	(25.3493)	(25.4999)
totnotax	-137.2164**	-129.8983**	-111.6711**	-92.7440*	-83.1549	-81.0482	-76.1173	-78.8728	-80.8322	-78.2299	-70.3315
	(57.2204)	(55.7169)	(54.5496)	(52.5431)	(52.1920)	(51.9147)	(52.2718)	(49.8624)	(49.5681)	(50.2689)	(49.5853)
lgdp	10.7081*	7.0621	8.1764	10.2727*	12.0286**	11.9873**	11.2816**	11.3928**	9.9653**	9.6076**	10.3746**
	(6.4542)	(5.5665)	(5.3372)	(5.2728)	(5.0786)	(5.0486)	(5.0455)	(4.9173)	(4.8519)	(4.7264)	(4.6133)
N	2674	2760	2845	2925	3002	3078	3145	3210	3274	3332	3384
N_g	152.000	152.000	152.000	152.000	152.000	152.000	152.000	152.000	152.000	152.000	152.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. Residuals of the CCE-MG regressions were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend).

Source: ICTD GRD (2014).

Table A1e Effect of tax reliance on regime using random effects logit and different time periods

	(1) Logit 1990 b/se	(2) Logit 1989 b/se	(3) Logit 1988 b/se	(4) Logit 1987 b/se	(5) Logit 1986 b/se	(6) Logit 1985 b/se	(7) Logit 1984 b/se	(8) Logit 1983 b/se	(9) Logit 1982 b/se	(10) Logit 1981 b/se	(11) Logit 1980 b/se
L.regime	-5.8295*	-4.7503	-5.4258*	-4.6072	-4.2589	-4.3057	-4.4396	-4.0695	-4.0695	-4.0695	-4.0688
	(3.4134)	(3.2288)	(3.1917)	(3.1318)	(3.1083)	(3.0804)	(3.0976)	(3.0943)	(3.0943)	(3.0943)	(3.1179)
L.Tax_Rel	1.4456	1.4285	1.3852	1.5599*	1.5749*	1.6176*	1.7938**	1.9286**	1.9286**	1.9286**	1.9338**
	(0.9997)	(0.9552)	(0.9232)	(0.9121)	(0.9143)	(0.8986)	(0.8987)	(0.8980)	(0.8980)	(0.8980)	(0.8977)
L.regimextaxrel	1.5535	1.0018	1.2196	0.9869	0.9862	0.9363	0.5851	0.4587	0.4587	0.4587	0.4429
	(1.8058)	(1.7950)	(1.7441)	(1.7445)	(1.7533)	(1.7489)	(1.7589)	(1.7580)	(1.7580)	(1.7580)	(1.7560)
L.lgdp	-0.2798	-0.2658	-0.2271	-0.1607	-0.1481	-0.1569	-0.1168	-0.0710	-0.0710	-0.0710	-0.0632
	(0.2024)	(0.1923)	(0.1860)	(0.1800)	(0.1797)	(0.1753)	(0.1720)	(0.1691)	(0.1691)	(0.1691)	(0.1688)
L.regimexlgdp	1.6316***	1.5256***	1.5873***	1.5060***	1.4724***	1.4859***	1.5345***	1.5006***	1.5006***	1.5006***	1.5087***
	(0.3644)	(0.3371)	(0.3367)	(0.3291)	(0.3253)	(0.3224)	(0.3233)	(0.3224)	(0.3224)	(0.3224)	(0.3261)
Aggregate effect in democracies											
$\beta 1 + \beta 2$	2.9991**	2.4302	2.6047*	2.5467*	2.5611*	2.5538*	2.3789	2.3873	2.3873	2.3873	2.3767
	(1.4984)	(1.5156)	(1.4757)	(1.4838)	(1.4926)	(1.4973)	1.5094	(1.5086)	(1.5086)	(1.5086)	(1.5071)
Random effects model descriptors											
σ_μ	.0010165	.0022275	.0020144	.0021315	.0011413	.001588	.0019006	.0018612	.0018612	.0018612	.0024203
ρ	3.14e-07	1.51e-06	1.23e-06	1.38e-06	3.96e-07	7.67e-07	1.10e-06	1.05e-06	1.05e-06	1.05e-06	1.78e-06
N	2675	2771	2866	2955	3042	3127	3203	3276	3276	3276	3328
N_g	178.000	178.000	178.000	178.000	178.000	178.000	178.000	178.000	178.000	178.000	178.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. $\beta 1 + \beta 2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1).

Source: ICTD GRD (2014).

Table A1f Effects of total tax and total non-tax on regime using random effects logit and different time periods

	(1) Logit 1990 b/se	(2) Logit 1989 b/se	(3) Logit 1988 b/se	(4) Logit 1987 b/se	(5) Logit 1986 b/se	(6) Logit 1985 b/se	(7) Logit 1984 b/se	(8) Logit 1983 b/se	(9) Logit 1982 b/se	(10) Logit 1981 b/se	(11) Logit 1980 b/se
L.regime	-1.5338	-0.7722	-1.2884	-0.6758	-0.3395	-0.2936	-0.6770	-0.4559	-0.4559	-0.4559	-0.4669
	(2.7011)	(2.5709)	(2.5134)	(2.4754)	(2.4505)	(2.4288)	(2.4123)	(2.4020)	(2.4020)	(2.4020)	(2.4196)
L.TotTax	-5.2028	-5.1182	-4.8903	-4.9707	-5.0682*	-6.1969**	-6.4802**	-5.8654**	-5.8654**	-5.8654**	-5.9108**
	(3.5678)	(3.3197)	(3.1534)	(3.0598)	(3.0670)	(3.0806)	(3.0451)	(2.9536)	(2.9536)	(2.9536)	(2.9601)
L.TotNonTax	-10.3526**	-10.4235**	-10.1350**	-11.0946***	-11.2291***	-12.1784***	-13.1596***	-13.4167***	-13.4167***	-13.4167***	-13.4987***
	(4.6311)	(4.4202)	(4.1837)	(4.2187)	(4.2315)	(4.2497)	(4.3307)	(4.3382)	(4.3382)	(4.3382)	(4.3439)
L.regimextottax	15.4740**	18.5316**	18.7130**	18.6609***	18.9726***	20.2246***	20.3023***	19.7911***	19.7911***	19.7911***	19.7624***
	(7.5928)	(7.4198)	(7.2735)	(7.1952)	(7.1787)	(7.1663)	(7.0444)	(7.0122)	(7.0122)	(7.0122)	(7.0173)
L.regimextotnontax	7.7160	9.2527	8.8476	9.8079	9.9148	10.8771*	12.0088*	12.2649*	12.2649*	12.2649*	12.2755*
	(6.4349)	(6.6052)	(6.3560)	(6.3947)	(6.4129)	(6.4366)	(6.5225)	(6.5057)	(6.5057)	(6.5057)	(6.4797)
L.lgdp	0.0484	0.0590	0.0846	0.1621	0.1774	0.2109	0.2657	0.2965	0.2965	0.2965	0.3065
	(0.2466)	(0.2310)	(0.2195)	(0.2131)	(0.2126)	(0.2072)	(0.2036)	(0.2009)	(0.2009)	(0.2009)	(0.2004)
L.regimexlgdp	0.9008**	0.7254*	0.7873**	0.7044*	0.6653*	0.6348*	0.6731*	0.6527*	0.6527*	0.6527*	0.6611*
	(0.4269)	(0.4040)	(0.3946)	(0.3892)	(0.3855)	(0.3816)	(0.3799)	(0.3775)	(0.3775)	(0.3775)	(0.3792)
Aggregate effects in democracies											
$\beta 1 + \beta 3$	10.2711	13.4134**	13.8227**	13.6902**	13.9043**	14.0277**	13.8221**	13.9256**	13.9256**	13.9256**	13.8516**
	(6.7310)	(6.6672)	(6.5872)	(6.5406)	(6.5183)	(6.5024)	(6.3833)	(6.3867)	(6.3867)	(6.3867)	(6.3874)
$\beta 2 + \beta 4$	-2.6366	-1.1707	-1.2874	-1.2867	-1.3142	-1.3013	-1.1508	-1.1517	-1.1517	-1.1517	-1.2231
	(4.4721)	(4.9074)	(4.7838)	(4.8043)	(4.8161)	(4.8333)	(4.8756)	(4.8441)	(4.8441)	(4.8441)	(4.8042)
Random effects model descriptors											
σ_μ	.0015408	.0010409	.0020486	.0025107	.0011372	.0016216	.001968	.0027575	.0027575	.0027575	.0024587
ρ	7.22e-07	3.29e-07	1.28e-06	1.92e-06	3.93e-07	7.99e-07	1.18e-06	2.31e-06	2.31e-06	2.31e-06	1.84e-06
N	2675	2771	2866	2955	3042	3127	3203	3276	3276	3276	3328
N_g	178.000	178.000	178.000	178.000	178.000	178.000	178.000	178.000	178.000	178.000	178.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. $\beta 1 + \beta 3$ captures the joint significance of the *tottax* and *regime x tottax* variables, and thus captures the impact of *tottax* on *regime* in democracies (e.g. when *regime* = 1). $\beta 2 + \beta 4$ similarly captures the joint significance of *totnontax* and *regime x totnontax*.

Source: ICTD GRD (2014).

Table A2a Effect of revenue variables on polity, Sys-GMM and CCE-MG, excluding OPEC

	(1) Sys-GMM No OPEC b/se λ=1	(2) CCE-MG No OPEC b/se λ=0	(3) Sys-GMM No OPEC b/se λ=1	(4) CCE-MG No OPEC b/se λ=0	(5) Sys-GMM No OPEC2 b/se λ=1	(6) CCE-MG No OPEC2 b/se λ=0	(7) Sys-GMM No OPEC2 b/se λ=1	(8) CCE-MG No OPEC2 b/se λ=0
L1.Polity_s	0.8483***		0.8493***		0.8466***		0.8459***	
	(0.0396)		(0.0311)		(0.0399)		(0.0306)	
L λ .Tax_Rel	14.1709***	10.7858			14.2054***	11.7948		
	(5.3115)	(7.5098)			(5.4848)	(7.5640)		
L λ .TotTax			-10.6448	29.4464			-10.9488	41.5375
			(13.8054)	(25.6993)			(14.0105)	(33.6230)
L λ .totnotax			-45.7684***	-129.1017**			-45.5146***	-132.9616**
			(12.7652)	(57.4292)			(12.8236)	(58.6307)
AR(1)	0.000		0.000		0.000		0.000	
AR(2)	0.653		0.676		0.664		0.686	
Hansen	0.314		0.197		0.356		0.230	
N	2381	2476	2381	2468	2348	2442	2348	2434
N_g	144.000	143.000	144.000	141.000	142.000	141.000	142.000	139.000

Notes: Regressions include a control for lgdp, not reported in order to conserve space. All independent variables - *polity*, *tax_rel*, *tottax* and *totnontax* - are contemporaneous for CCE-MG estimates, and lagged t-1 for GMM estimates. All GMM regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. Residuals of the CCE-MG regressions were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend).

Source: ICTD GRD (2014).

Table A2b Effect of revenue variables on polity, Sys-GMM and CCE-MG, excluding MENA

	(1) Sys-GMM No MENA b/se λ =1	(2) CCE-MG No MENA b/se λ =0	(3) Sys-GMM No MENA b/se λ =1	(4) CCE-MG No MENA b/se λ =0
L1.Polity_s	0.8511*** (0.0420)		0.8500*** (0.0374)	
L λ .Tax_Rel	16.0808*** (5.8897)	12.2552 (7.6781)		
L λ .TotTax			-14.7182 (17.5481)	81.9161** (34.0554)
L λ .totnotax			-20.3807* (11.2403)	-148.3973** (61.6052)
AR(1)	0.000		0.000	
AR(2)	0.603		0.635	
Hansen	0.225		0.116	
N	2387	2482	2387	2474
N_g	144.000	143.000	144.000	141.000

Notes: Regressions include a control for lgdp, not reported in order to conserve space. All independent variables - *polity*, *tax_rel*, *tottax* and *totnontax* - are contemporaneous for CCE-MG estimates, and lagged t-1 for GMM estimates. All GMM regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. Residuals of the CCE-MG regressions were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend).

Source: ICTD GRD (2014).

Table A2c Effect of revenue variables on polity, Sys-GMM and CCE-MG, excluding EECA

	(1) Sys-GMM No CECA b/se λ =1	(2) CCE-MG No CECA b/se λ =0	(3) Sys-GMM No CECA b/se λ =1	(4) CCE-MG No CECA b/se λ =0
L1.Polity_s	0.8687*** (0.0317)		0.8434*** (0.0292)	
L λ .Tax_Rel	16.9828*** (5.3436)	16.6450* (9.9554)		
L λ .TotTax			-24.2689 (16.9798)	22.7024 (39.1279)
L λ .totnotax			-18.2430** (8.4942)	-83.4526 (70.8292)
AR(1)	0.000		0.000	
AR(2)	0.752		0.753	
Hansen	0.326		0.460	
N	2338	2428	2338	2428
N_g	133.000	133.000	133.000	133.000

Notes: Regressions include a control for lgdp, not reported in order to conserve space. All independent variables - *polity*, *tax_rel*, *tottax* and *totnontax* - are contemporaneous for CCE-MG estimates, and lagged t-1 for GMM estimates. All GMM regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. Residuals of the CCE-MG regressions were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend).

Source: ICTD GRD (2014).

Table A2d Effect of revenue variables on polity, Sys-GMM and CCE-MG, excluding OECD

	(1) Sys-GMM No OECD b/se λ =1	(2) CCE-MG No OECD b/se λ =0	(3) Sys-GMM No OECD b/se λ =1	(4) CCE-MG No OECD b/se λ =0
L1.Polity_s	0.8700*** (0.0394)		0.8553*** (0.0262)	
L λ .Tax_Rel	20.5571*** (5.3783)	29.2320*** (9.6204)		
L λ .TotTax			-9.0870 (15.6977)	48.3573 (40.5882)
L λ .totnotax			-26.3057*** (9.0503)	-229.8694*** (77.5813)
AR(1)	0.000		0.000	
AR(2)	0.652		0.665	
Hansen	0.318		0.483	
N	2090	2181	2090	2173
N_g	127.000	126.000	127.000	124.000

Notes: Regressions include a control for lgdp, not reported in order to conserve space. All independent variables - *polity*, *tax_rel*, *tottax* and *totnontax* - are contemporaneous for CCE-MG estimates, and lagged t-1 for GMM estimates. All GMM regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. Residuals of the CCE-MG regressions were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level (with and without trend). The Pesaran (2007) panel unit root test fails to reject non-stationarity when a trend is included; without trend, it fails to reject in regression 2 but does reject in regression 4.

Source: ICTD GRD (2014).

Table A2e Effect of revenue variables on polity, Sys-GMM and CCE-MG, excluding small countries

	(1) Sys-GMM No <250000 b/se λ=1	(2) CCE-MG No <250000 b/se λ=0	(3) Sys-GMM No <250000 b/se λ=1	(4) CCE-MG No <250000 b/se λ=0	(5) Sys-GMM No <1000000 b/se λ=1	(6) CCE-MG No <1000000 b/se λ=0	(7) Sys-GMM No <1000000 b/se λ=1	(8) CCE-MG No <1000000 b/se λ=0
L1.Polity_s	0.8741***		0.8612***		0.8576***		0.8438***	
	(0.0359)		(0.0280)		(0.0336)		(0.0295)	
L λ .Tax_Rel	18.0684***	14.3193**			21.2386***	12.5025*		
	(5.3422)	(7.2816)			(5.6642)	(7.3414)		
L λ .TotTax			-11.7086	73.7190**			-21.2869	83.1988**
			(15.5717)	(30.8115)			(15.4780)	(33.7848)
L λ .totnotax			-26.0380***	-137.2164**			-30.8082***	-157.6456**
			(8.3181)	(57.2204)			(9.7561)	(67.2014)
AR(1)	0.000		0.000		0.000		0.000	
AR(2)	0.648		0.650		0.596		0.596	
Hansen	0.376		0.291		0.363		0.527	
N	2579	2682	2579	2674	2361	2455	2361	2443
N_g	155.000	154.000	155.000	152.000	145.000	143.000	145.000	140.000

Notes: Regressions include a control for lgdp, not reported in order to conserve space. All independent variables - *polity*, *tax_rel*, *tottax* and *totnontax* - are contemporaneous for CCE-MG estimates, and lagged t-1 for GMM estimates. All GMM regression results based on the two-step system-GMM procedure. Windmeijer standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The p-values reported for the Hansen statistic are for the null hypothesis that instruments are valid. The values reported for AR(2) are p-values for tests of second order autocorrelation. Residuals of the CCE-MG regressions were tested for non-stationarity. In all specifications the Maddala and Wu (1999) panel unit root test rejects non-stationarity at the 1% level, while the Pesaran (2007) panel unit root test fails to reject non-stationarity (in both cases, with and without trend).

Source: ICTD GRD (2014).

Table A2f Effect of revenue variables on regime, using logit, excluding OPEC

	(1) Logit No OPEC b/se	(2) Logit No OPEC b/se	(3) Logit No OPEC b/se With controls	(4) Logit No OPEC b/se With controls	(5) Logit No OPEC2 b/se	(6) Logit No OPEC2 b/se	(7) Logit No OPEC2 b/se With controls	(8) Logit No OPEC2 b/se With controls
L.regime	-7.2744* (3.8054)	-0.9191 (2.7934)	-7.7908* (4.2103)	-1.0720 (3.1391)	-7.5696* (4.0059)	-1.9989 (2.9379)	-8.4578* (4.4283)	-2.4226 (3.3167)
L.Tax_Rel	1.3143 (1.1005)		1.4803 (1.1717)		1.3945 (1.1403)		1.5931 (1.2206)	
L.regimextaxrel	3.2727 (2.1138)		4.6463* (2.4115)		2.8918 (2.2546)		4.4091* (2.5784)	
L.TotTax		-5.5983 (3.6842)		-6.3885* (3.8472)		-5.0279 (3.7156)		-5.8185 (3.8972)
L.TotNonTax		-9.8828* (5.0990)		-13.0402** (6.1668)		-9.6241* (5.2685)		-13.0944** (6.4451)
L.regimextottax		18.6729** (8.1128)		20.1703** (8.4754)		15.1396* (8.3462)		17.1873* (8.7702)
L.regimextotnontax		4.2387 (6.6050)		-4.6449 (10.8004)		4.1585 (6.8609)		-5.5789 (11.1864)
Aggregate effects in democracies								
$\beta 1 + \beta 2$	4.5870** (1.8038)		6.1266*** (2.1060)		4.2863** (1.9445)		6.0021*** (2.2702)	
$\beta 1 + \beta 3$		13.0745* (7.2565)		13.7817* (7.5704)		10.1117 (7.5069)		11.36885 (7.8881)
$\beta 2 + \beta 4$		-5.6440 (4.2175)		-17.6851** (8.8605)		-5.4655 (4.4114)		-18.6733** (9.1392)
Random effects model descriptors								
σ_μ	.0015209	.0011511	.0009669	.0014001	.0008446	.0015209	.0009849	.0014303
ρ	7.03e-07	4.03e-07	2.84e-07	5.96e-07	2.17e-07	7.03e-07	2.95e-07	6.22e-07
N	2497	2497	2153	2153	2468	2468	2124	2124
N_g	167.000	167.000	146.000	146.000	165.000	165.000	144.000	144.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. Regressions in columns (3), (4), (7) and (8) include a standard set of controls, not reported here to conserve space: *lgdp*, *civil_war* and *growthpc*. $\beta 1 + \beta 2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta 1 + \beta 3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta 2 + \beta 4$ captures the joint significance of *totnontax* and *regime x totnontax*. Source: ICTD GRD (2014).

Table A2g Effect of revenue variables on regime, using logit, excluding MENA

	(1) Logit no MENA b/se	(2) Logit no MENA b/se	(3) Logit no MENA b/se With controls	(4) Logit no MENA b/se With controls
L.regime	-7.2744* (3.8054)	-0.9191 (2.7934)	-7.7908* (4.2103)	-1.0720 (3.1391)
L.Tax_Rel	1.3143 (1.1005)		1.4803 (1.1717)	
L.regimextaxrel	3.2727 (2.1138)		4.6463* (2.4115)	
L.TotTax		-5.5983 (3.6842)		-6.3885* (3.8472)
L.TotNonTax		-9.8828* (5.0990)		-13.0402** (6.1668)
L.regimextottax		18.6729** (8.1128)		20.1703** (8.4754)
L.regimextotnontax		4.2387 (6.6050)		-4.6449 (10.8004)
Aggregate effects in democracies				
$\beta 1 + \beta 2$	2.8625* (1.4941)		3.3315** (1.5256)	
$\beta 1 + \beta 3$		9.7673 (6.7117)		10.2372 (7.1954)
$\beta 2 + \beta 4$		-2.3959 (4.4581)		-3.6501 (5.1376)
Random effects model descriptors				
σ_μ	.0010094	.0015342	.0010014	.0014558
ρ	3.10e-07	7.15e-07	3.05e-07	6.44e-07
N	2497	2497	2153	2153
N_g	167.000	167.000	146.000	146.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. Regressions in columns (3) and (4) include a standard set of controls, not reported here to conserve space: *lgdp*, *civil_war* and *growthpc*. $\beta 1 + \beta 2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta 1 + \beta 3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta 2 + \beta 4$ captures the joint significance of *totnontax* and *regime x totnontax*.

Source: ICTD GRD (2014).

Table A2h Effect of revenue variables on regime, using logit, excluding OECD

	(1) Logit no OECD b/se	(2) Logit no OECD b/se	(3) Logit no OECD b/se With controls	(4) Logit no OECD b/se With controls
L.regime	-7.2744* (3.8054)	-0.9191 (2.7934)	-7.7908* (4.2103)	-1.0720 (3.1391)
L.Tax_Rel	1.3143 (1.1005)		1.4803 (1.1717)	
L.regimextaxrel	3.2727 (2.1138)		4.6463* (2.4115)	
L.TotTax		-5.5983 (3.6842)		-6.3885* (3.8472)
L.TotNonTax		-9.8828* (5.0990)		-13.0402** (6.1668)
L.regimextottax		18.6729** (8.1128)		20.1703** (8.4754)
L.regimextotnontax		4.2387 (6.6050)		-4.6449 (10.8004)
Aggregate effects in democracies				
$\beta 1 + \beta 2$	2.7528* (1.5275)		3.1577** (1.5875)	
$\beta 1 + \beta 3$		8.9744 (6.9314)		9.0302 (7.5290)
$\beta 2 + \beta 4$		-2.4845 (4.4798)		-3.8494 (5.2523)
Random effects model descriptors				
σ_μ	.0015918	.00152	.0009886	.0014122
ρ	7.70e-07	7.02e-07	2.97e-07	6.06e-07
N	2497	2497	2153	2153
N_g	167.000	167.000	146.000	146.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. Regressions in columns (3) and (4) include a standard set of controls, not reported here to conserve space: *lgdp*, *civil_war* and *growthpc*. $\beta 1 + \beta 2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta 1 + \beta 3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta 2 + \beta 4$ captures the joint significance of *totnontax* and *regime x totnontax*.

Source: ICTD GRD (2014).

Table A2i Effect of revenue variables on regime, using logit, excluding small countries

	(1) Logit No <250000 b/se	(2) Logit No <250000 b/se	(3) Logit No <250000 b/se With controls	(4) Logit No <250000 b/se With controls	(5) Logit No <1000000 b/se	(6) Logit No <1000000 b/se	(7) Logit No <1000000 b/se With controls	(8) Logit No <1000000 b/se With controls
L.regime	-7.2744*	-0.9191	-7.7908*	-1.0720	-7.5696*	-1.9989	-8.4578*	-2.4226
	(3.8054)	(2.7934)	(4.2103)	(3.1391)	(4.0059)	(2.9379)	(4.4283)	(3.3167)
L.Tax_Rel	1.3143		1.4803		1.3945		1.5931	
	(1.1005)		(1.1717)		(1.1403)		(1.2206)	
L.regimextaxrel	3.2727		4.6463*		2.8918		4.4091*	
	(2.1138)		(2.4115)		(2.2546)		(2.5784)	
L.TotTax		-5.5983		-6.3885*		-5.0279		-5.8185
		(3.6842)		(3.8472)		(3.7156)		(3.8972)
L.TotNonTax		-9.8828*		-13.0402**		-9.6241*		-13.0944**
		(5.0990)		(6.1668)		(5.2685)		(6.4451)
L.regimextottax		18.6729**		20.1703**		15.1396*		17.1873*
		(8.1128)		(8.4754)		(8.3462)		(8.7702)
L.regimextotnontax		4.2387		-4.6449		4.1585		-5.5789
		(6.6050)		(10.8004)		(6.8609)		(11.1864)
Aggregate effects in democracies								
$\beta1 + \beta2$	3.3003**		3.4981**		3.6935**		4.0201**	
	(1.4913)		(1.5300)		1.5291		(1.5856)	
$\beta1 + \beta3$		10.4555		10.9750		16.2215**		18.2476**
		(6.9786)		(7.1910)		(8.2932)		(8.9674)
$\beta2 + \beta4$		-3.5371		-3.8580		-2.7869		-3.0934
		(4.9678)		(5.1795)		(5.1249)		(5.3670)
Random effects model descriptors								
σ_μ	.0021619	.0015575	.0020911	.001483	.0024077	.0016506	.0020194	.0015594
ρ	1.42e-06	7.37e-07	1.33e-06	6.69e-07	1.76e-06	8.28e-07	1.24e-06	7.39e-07
N	2497	2497	2153	2153	2468	2468	2124	2124
N_g	167.000	167.000	146.000	146.000	165.000	165.000	144.000	144.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. Regressions in columns (3), (4), (7) and (8) include a standard set of controls, not reported here to conserve space: *lgdp*, *civil_war* and *growthpc*. $\beta1 + \beta2$ captures the joint significance of the *tax_rel* and *regime x taxrel* variables, and thus captures the impact of *tax_rel* on *regime* in democracies (e.g. when *regime* = 1). $\beta1 + \beta3$ similarly captures the joint significance of *tottax* and *regime x tottax,* while $\beta2 + \beta4$ captures the joint significance of *totnontax* and *regime x totnontax*. Source: ICTD GRD (2014).

References

Acemoglu, D., Johnson, S., Robinson, J. and Yared, P. (2008) 'Income and Democracy', *American Economic Review, 98*(3): 808-842

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

Andersen, J. and Aslaksen, S. (2013) 'Oil and Political Survival', *Journal of Development Studies* 100: 89-106

—— and Ross, M. (2014) 'The Big Oil Change: A Closer Look at the Haber-Menaldo Analysis', *Comparative Political Studies* 47: 993-1021

Aslaksen, S. (2010) 'Oil and Democracy: More than a Cross-Country Correlation?', *Journal of Peace Research* 47(4): 421-431

Baskaran, T. and Bigsten, A. (2013) 'Fiscal Capacity and the Quality of Government in Sub-Saharan Africa', *World Development* 45(C): 92-107

Bird, R. (2011) 'Subnational Taxation in Developing Countries: A Review of the Literature,' *Journal of International Commerce, Economics and Policy* 2(1): 139-161

Birdsall, N. (2007) *Do No Harm: Aid, Weak Institutions, and the Missing Middle in Africa*, Center for Global Development Working Paper 113, Washington DC: Center for Global Development

Bond, S. (2002) *Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice*, Working Paper 09/02, London: Institute for Fiscal Studies

Bornhorst, F., Gupta, S., and Thornton, J. (2008) *Natural Resource Endowments, Governance and the Domestic Revenue Effort: Evidence from a Panel of Countries*, IMF Working Paper WP/08/170, Washington DC: International Monetary Fund

Brunnschweiler, C. and Bulte, E. (2008) 'The resource curse revisited and revised: A tale of paradoxes and red herrings', *Journal of Environmental Economics and Management* 55(3): 248-264

Cavalcanti, T., Mohaddes, K. and Raissi, M. (2011) 'Growth, Development and Natural Resources: New Evidence using a Heterogeneous Panel Analysis', *The Quarterly Review of Economics and Finance* 51(4): 305-18

Cheibub, J., Gandhi, J. and Vreeland, J. (2010) 'Democracy and Dictatorship Revisited', *Public Choice* 143(1-2): 67-101

Collier, P. (2006) 'Is Aid Oil? An Analysis of Whether Africa Can Absorb More Aid', *World Development* 34(9): 1482-1497

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

Eberhardt, M. and Teal, F. (2013) 'Structural Change and Cross-Country Growth Empirics', *World Bank Economic Review* 27(2): 229-271

Fayad, G., Bates, R. and Hoeffler, A. (2011) *Income and Democracy: Lipset's Law Inverted*, OxCarre Research Paper 61

Freedom House, *Freedom in the World 2013*, retrieved 8 November 2013, <http://www.freedomhouse.org>

Gupta, A. (2007) *Determinants of Tax Revenue Efforts in Developing Countries*, IMF Working Paper 07/184, Washington DC: International Monetary Fund

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

Hausman, J. (1978) 'Specification Tests in Econometrics', *Econometrica* 46(6): 1251-1271

—— and McFadden, D. (1984) 'Specification Tests for the Multinomial Logit Model', *Econometrica* 52(5): 1219-1240

Herb, M. (2005) 'No Representation Without Taxation? Rents, Development and Democracy', *Comparative Politics* 37(3): 297-316

Holly, S., Pesaran, M.H. and Yamagata, T. (2010) 'A Spatio-Temporal Model of House Prices in the US,' *Journal of Econometrics* 158(1): 160-73

Humphreys, M. (2005) 'Natural Resources, Conflict, and Conflict Resolution', *The Journal of Conflict Resolution* 49(4): 508-537

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

IMF, *Article IV Staff Reports*, <http://www.imf.org/external/ns/cs.aspx?id=51>

IMF, *Government Finance Statistics*, <http://elibrary-data.imf.org/FindDataReports.aspx?d=33061&e=170809>

IMF, *World Economic Outlook*, <http://www.imf.org/external/pubs/ft/weo/2014/01/weodata/index.aspx>

Jensen, N. and Wantchekon, L. (2004) 'Resource Wealth and Political Regimes in Africa', *Comparative Political Studies* 37(7): 816-841

Jerven, M. (2013) 'Comparability of GDP Estimates in sub-Saharan Africa: The Effect of Revisions in Sources and Methods since Structural Adjustment', *Review of Income and Wealth,* 59(S1): S16-S36

Levi, M. (1988) *Of Rule and Revenue*, Berkeley: University of California Press

Lieberman, E. (2002) 'Taxation Data as Indicators of State-Society Relations: Possibilities and Pitfalls in Cross-National Research', *Studies in Comparative International Development* 36(4): 89-115

Maddala, G. and Wu, S. (1999) 'A Comparative Study of Unit Root Tests with Panel Data and a Simple New Test', *Oxford Bulletin of Economics and Statistics* 61(S1): 631–652

Marshall, M., Jaggers, K., and Gurr, T. (2012) *Political regime characteristics and transitions, 1800-2008*, *Polity IV Project*, Fairfax County: Center for Systemic Peace

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

Moore, B. (1966) *The Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*, Boston: Beacon

Moore, M. (1998) 'Death Without Taxes: Democracy, State Capacity and Aid Dependence in the Fourth World', in M. Robinson and G. White (eds), *The Democratic Developmental State: Politics and Institutional Design* Oxford: Oxford University Press

Morrison, K. (2009) 'Oil, Nontax Revenue and the Redistributional Foundations of Regime Stability', *International Organization* 63(1): 107-138

Morrissey, O., Prichard, W. and Torrance, S. (2014) *Aid and Taxation: Exploring the Relationship Using New Data*, ICTD Working Paper 21, Brighton: International Centre for Tax and Development

Nickell, S. (1981) 'Biases in Dynamic Models with Fixed Effects', *Econometrica* 49: 1417-1426

Pesaran, M.H. (2007) 'A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence', *Journal of Applied Econometrics* 22(2): 265-312

—— (2006) 'Estimation and Inference in Large Heterogeneous Panels With a Multifactor Error Structure', *Econometrica* 74(4): 967-1012

—— and Smith, R. (1995) 'Estimating Long-Run Relationships From Dynamic Heterogeneous Panels', *Journal of Econometrics* 68(1): 79-113

Prichard, W. (2010) *Taxation, Responsiveness and Accountability in Sub-Saharan Africa,* (DPhil), Brighton: University of Sussex

—— Cobham, A. and Goodall, A. (2014) *The ICTD Government Revenue Dataset*, ICTD Working Paper 19, Brighton: International Centre for Tax and Development

PRS Group, *International Country Risk Guide*, <https://www.prsgroup.com/about-us/our-two-methodologies/icrg>

Przeworski, A., Alvarez, M., Cheibub, J. and Limongi, F. (2000) *Democracy and Development: Political Institutions and Well-Being in the World 1950-1990*, Cambridge: Cambridge University Press

Ramsay, K. (2011) 'Revisiting the Resource Curse: Natural Disasters, the Price of Oil and Democracy', *International Organization* 65(3): 507-529

Roodman, D. (2009) 'A Note on the Theme of Too Many Instruments', *Oxford Bulletin of Economics and Statistics 71*(1): 135-158

—— (2006) *How to Do xtabond2: An Introduction to 'Difference' and 'System' GMM in Stata,* Center for Global Development Working Paper Number 103, Washington DC: Center for Global Development

Ross, M. (forthcoming) 'What Have We Learned About the Resource Curse?', *Annual Review of Political Science*

—— (2012) *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*, Princeton: Princeton University Press

—— (2004) 'Does Taxation Lead to Representation?', *British Journal of Political Science* 34: 229-249

—— (2001) 'Does Oil Hinder Democracy?', *World Politics* 53: 325-361

Smith, B. (2004) 'Oil Wealth and Regime Survival in the Developing World, 1960-99', *American Journal of Political Science* 48(2): 232-246

Timmons, J. (2010) 'Taxation and Representation in Recent History', *The Journal of Politics* 72(1): 191-208

Tsui, K. (2010) 'More Oil, Less Democracy? Theory and Evidence from Crude Oil Discoveries', *Economic Journal* 121(551): 89-115

UCDP/PRIO, *Armed Conflict Dataset*, <http://www.pcr.uu.se/research/ucdp/datasets/ucdp_prio_armed_conflict_dataset/>

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

van der Ploeg, F. (2011) 'Natural Resources: Curse or Blessing?', *Journal of Economic Literature* 49(2): 366-420

Vijverberg, W. (2011) *Testing for IIA with the Hausman-McFadden Test,* IZA Discussion Paper No.5826, Bonn: Institute for the Study of Labor Law

Wantchekon, L. (2002) 'Why Do Resource Dependent Countries Have Authoritarian Governments?', *Journal of African Finance and Economic Development* 2: 57-77

Wiens, D., Poast, P. and Clark, W. (forthcoming) 'The Political Resource Curse: An Empirical Re-Evaluation', *Political Research Quarterly*, Doi: 10.1177/1065912914543836

World Bank, *World Development Indicators*, <http://data.worldbank.org/data-catalog/world-development-indicators>

World Bank, *Worldwide Governance Indicators*, <http://info.worldbank.org/governance/wgi/index.aspx#home>

Wright, J., Frantz, E. and Geddes, B. (2013) 'Oil and Autocratic Regime Survival', *British Journal of Political Science*: 1-20

International Centre for Tax and Development
at the Institute of Development Studies
Brighton BN1 9RE, UK
T: +44 (0)1273 606261
F: +44 (0)1273 621202
E: info@ictd.ac
www.ictd.ac