Do giant oilfield discoveries make countries poor at taxation?*

Abraham Lartey †

This version: August, 2022
Current version

Abstract

It has often been argued that countries that produce natural resources mobilize less non-resource tax revenues than other countries. In this paper, I exploit the exogenous variation in the timing of giant oilfield discoveries to estimate the causal impact of natural resources on taxation. The timing of giant oilfield discoveries is arguably exogenous and thus renders them appealing to empirically examine this argument. This allows me to examine the performance of non-resource tax revenue effort before and immediately after discovery as well as the period corresponding to the inflow of revenues from the production. I find that non-resource tax revenues tend to increase in the period following the discovery before the onset of production and after production commences. This effect is due to an increase in non-resource indirect tax revenues. Further analysis shows that both the total and indirect non-resource tax revenues, experience an increase in only low- middle income countries. This effect is largely driven by an increase in the consumption of goods and services.
Keywords: Natural Resources, Tax revenue, Giant oilfield discovery
JEL classification codes:E6, H2, Q33, Q38, O13.

*I would like to express my gratitude to Fidel Pérez Sebastián and Pedro Albarran Pérez for their guidance and support. I will also like to express my gratitude to seminar and conference participants at the Applied Seminar, Universidad de Alicante, 41st ASSET Conference in Athens and the 8th PhD-Student workshop on Industrial and Public Economics(WIPE), Reus for their useful comments and suggestions. Financial support from Generalitat Valencia (GRISOLIA/2018/110) is duly acknowledged.
†University College London, London, United Kingdom. Email: a.lartey@ucl.ac.uk

1 Introduction

The Sustainable Development Goals (SDGs) emphasize the need for increased domestic resource mobilization to financing the development challenges. This is particularly important for countries endowed with natural resources because of the unique nature of extractive natural resources which are exhaustible and with volatile revenues. This volatility might be transmitted to the economy in the absence of good fiscal framework leading to macroeconomic instability (Van der Ploeg and Poelhekke, 2010). Also, the climate change challenge and energy transition require countries that are reliant on revenues from the exploitation of fossil fuels, to pursue economic diversification strategies (World Bank, 2022): i.e., reducing the dependence on the resource sector which implies mobilizing more non- resource revenues.

However, it has been argued that natural resource producing countries tend to mobilize less non-resource tax revenues (see Figure 1) compared to countries that do not produce natural resources (Moore, 1998, 2007; Collier, 2006). As argued by (Besley and Persson, 2013, 2014), revenues from natural resources can serve as a disincentive to design or maintain efficient tax systems. Specifically, the discovery of a natural resource e.g. oil can reduce the incentive to invest in fiscal capacity due to the anticipation of future revenue inflows (Besley and Persson, 2013). Furthermore, the reliance on the natural resource sector can lead to exchange rate appreciation thereby crowding out other sectors as predicted by theoretical models of Dutch disease (Corden and Neary, 1982; Corden, 1984). This could potentially erode the non-resource tax base resulting in lower tax revenues for a given tax rate.

In this paper, I exploit the exogenous variation in the timing of giant oilfields discoveries to ascertain the causal impact of natural resources on the mobilization of non-resource tax revenues.[1] Specifically, the paper seeks to answer the following questions: Does the discovery of giant oilfields lead to a lower non-resource tax revenues? Which components of non-resource tax revenues are affected? The features of giant oilfields discoveries provide a unique source of a macro-relevant news shock (Arezki et al., 2017). This is because there is a lag between discovery and initial production for about four to six years. This provides news about future output and future inflow of revenues and therefore allows one to directly examine the performance of non-resource revenue mobilization before and immediately after discovery, as

[1] I use the term oilfields discoveries to indicate both the discovery of oil and gas fields. Recent studies that have used this identification strategy includes Lei and Michaels (2014); Arezki et al. (2017); Harding et al. (2020) and Perez-Sebastian et al. (2021) to study the impact of natural resources on internal armed conflicts, real exchange rate appreciation, macroeconomic performance and protectionism respectively.

well as the period corresponding to the inflow of revenues from the production of petroleum. Also, the timing of giant resource discoveries is arguably exogenous and unpredictable due to the uncertain nature of oil exploration.

I combine a new dataset on giant oilfields discoveries (Horn, 2011; Arezki et al., 2017) with tax revenue data covering 178 countries for the period 1989 to 2012. The tax revenue data is from the publicly available International Centre for Tax and Development Government Revenue Dataset (ICTD-GRD). The ICTD-GRD has an improved coverage and makes a consistent distinction between resource and non-resource tax revenues.[2]

I find a positive effect of giant oilfields discoveries on non-resource tax revenues before and after production commences. This is mainly due to an increase in non-resource indirect tax revenues. Specifically, non-resource indirect tax revenues increase before and after the actual production following a giant discovery in low-middle income countries but no significant effect is found for high income countries. I also find that the impact on non-resource indirect tax revenues is largely driven by an increase in the tax base (increase in the consumption of goods and services) but not in the tax rate.

This paper makes two main contributions to the existing literature. First, I show that the abundance of natural resources has a positive effect on non-resource tax revenues. The result from existing empirical studies generally indicates a crowding out effect and fraught with endogeneity problems. Bornhorst et al. (2009) investigated if there is evidence of an offset between hydrocarbon revenues and other domestic revenues in a panel of 30 hydrocarbon producing countries. The authors found that a one percentage point increase in hydrocarbon revenues lowers domestic revenues by about 0.2 percentage points. Building on the work of the previous authors, Crivelli and Gupta (2014) investigated the impact of resource revenues on different components of non-resource revenues in 35 resource-rich countries. Their results indicate that resource revenues crowd out taxes on goods and services more for the VAT while the impact on corporate and trade taxes is smaller. Thomas and Trevino (2013) also found that non-resource revenue is negatively influenced by higher resource revenues for Sub Saharan Africa. This crowding out effect could be due to the way this issue has been investigated where non-resource revenues (as % of GDP) is normally regressed on resource revenues (% of GDP). This is problematic because resource revenue to GDP ratio can change

[2]The main source of tax revenue data used by researchers is the IMF Government Finance Statistics (GFS). The GFS has less coverage for developing countries. A detailed description of the ICTD-GRD and its advantages over the GFS can be found in Prichard et al. (2014).

due to an increase in the production of natural resources. Non- resource revenue to GDP ratio can appear to be crowded out due to this increased income thereby biasing the results. Ossowski and Gonzáles (2012) tried to overcome this endogeneity issue by regressing non-resource revenue (% of non-resource GDP) on resource revenue (% of resource GDP) using a sample of 15 Latin American and Caribbean countries and found a crowding out effect. However, expressing resource revenue as a ratio of resource GDP takes away the importance of resource revenue in the economy and rather focus on the effective tax rate in the resource sector since the revenues accrued from the sector depends on the fiscal regime. Second, I also show that the news of a natural resource discovery could have potential anticipatory effect on non-resource tax revenues even before production commences. This contrasts with the existing empirical studies which assumed that only the exploitation i.e., the revenue generated from production of natural resources tend to impact domestic resource mobilization thereby overlooking important short run implications.

The remainder of the paper is structured as follows. The next section presents the data and the empirical strategy. Section 3 presents and discusses the empirical results. Section **??** reports the results of the heterogeneous analysis. Section 4 conducts several tests to check the robustness of the results. Finally, the paper concludes in section 5.

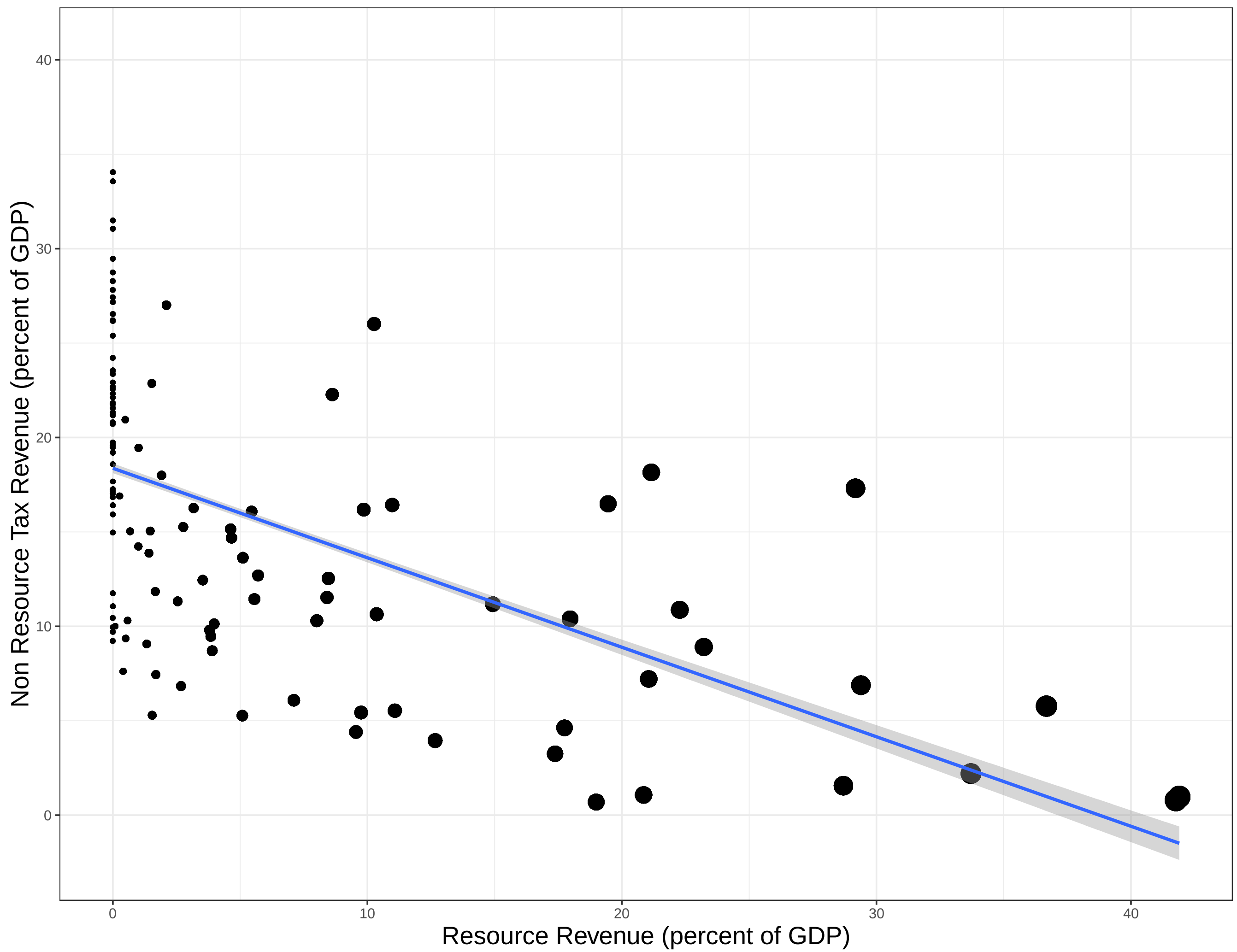


Figure 1: **Average resource revenues vs Average non-resource taxes (1990-2016)**

Notes: The figure shows scatter plot of the averages of resource revenues and non-resource tax revenues across different regions. The dots represent each country, and the size of the dots represents the magnitude of average resource revenues.

2 Data and Empirical Strategy

In this section, I provide a description of the data, empirical model and the identification strategy. In section 2.1, I describe the data used in the study. This is followed by the justification of the identification strategy in section 2.2. The empirical model is specified in section 2.3.

2.1 Data

The study make use of two main sources of data in addition to various control variables which are taken from several sources.

Non-Resource Tax Revenues Data. The first data is on tax revenues and comes from the International Centre for Tax and Development (ICTD) Government Revenue Dataset (GRD). The version of the data I use, covers the period 1980 to 2016. The ICTD GRD has an improved coverage for most developed and developing countries starting from 1989. There is also a consistent distinction between tax and non-tax revenues using standard system of classification.[3] This gives it an advantage over other sources of government revenues such as the IMF and the World Bank. Furthermore, tax revenues are decomposed into resource and non-resource tax revenues. This is important to adequately capture the impact of natural resources on non-resource tax revenues and its various components. Non-resource tax revenue is defined as tax revenues excluding social contributions and resource tax revenues.

Giant oil and gas discoveries. The second dataset is on giant oil and gas discoveries for a set of countries. This data comes from Horn (2011) and contains information on the timing, the location, and the estimated ultimately recoverable reserve (URR) of oil and gas (in oil equivalent) from 1868 to 2012. A discovery is considered a giant discovery if it has an estimated URR of at least 500 million barrel of oil equivalent (MMBOE). Since the data on my dependent variables start from 1989 and the data on giant oilfield discoveries end in 2012, I restrict the sample to cover the period, 1989 to 2012. Giant discoveries were made in 48 countries over the period under consideration. I plot the distribution of the number of giant oilfields discovery over time in Figure 2. One can see there is large variation in the number of discoveries made over time. Figure 3 shows the distribution of the number of discoveries across regions. One can see that there have been significant number of giant oilfields discovered in the various

[3]A detailed description of the ICTD-GRD can be found in Prichard et al. (2014).

regions.

Control and additional variables. There are several key determinants of tax revenues documented in the literature. These include population growth, real GDP per capita, the structure of the economy, trade balance, level of democracy, corruption, and financial openness. It is expected that countries with a higher per capita income, higher degree of financial openness, higher trade balance and a rapid population growth to be associated with higher tax collection. Also, more democratic countries are also expected to have a higher tax revenues. On the other hand, it is expected that less tax revenues will be mobilized if most of the economic activity is concentrated in the agricultural sector which is largely informal and difficult to tax. I proxy the structure of the economy agriculture value added (% of GDP). To measure the degree of financial openness, I use the IMF financial development index (Svirydzenka, 2016) which is a broad-based measure that considers the depth, access, and efficiency of financial institutions. The more financially opened a country is, the higher the value of the index. Trade openness is defined as the non-resource exports plus non-resource imports to GDP ratio. To measure corruption, I use the corruption risk score from the International Country Risk Guide (ICRG) which is an assessment of corruption within the political system. This corruption risk index has a minimum score of zero and maximum score of six. A higher score indicates a lower corruption risk. The data on population growth, real GDP per capita, the structure of the economy and trade balance are sourced from the World Bank Development Indicators (WDI). Table 1 presents the summary statistics of the main variables. A detailed definition of all the variables and their sources are provided in Table A.2 in Appendix A.

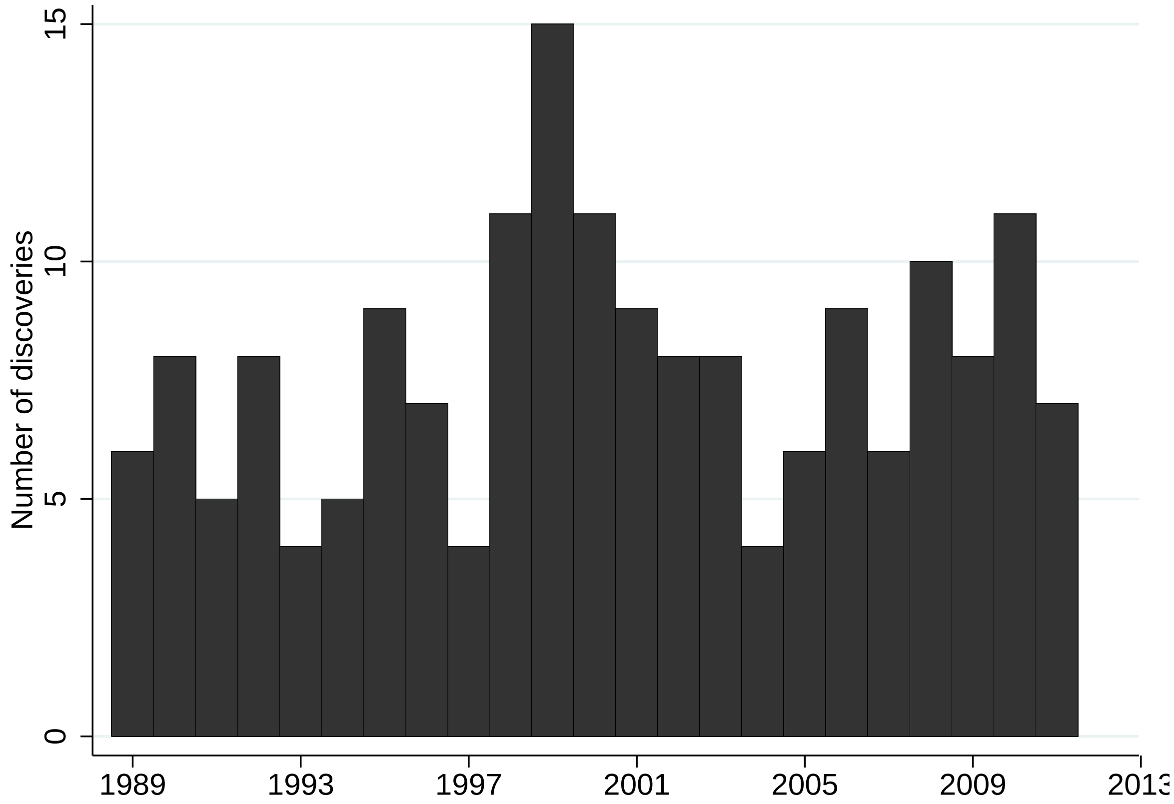


Figure 2: **Number of discoveries across time (1989-2012)**

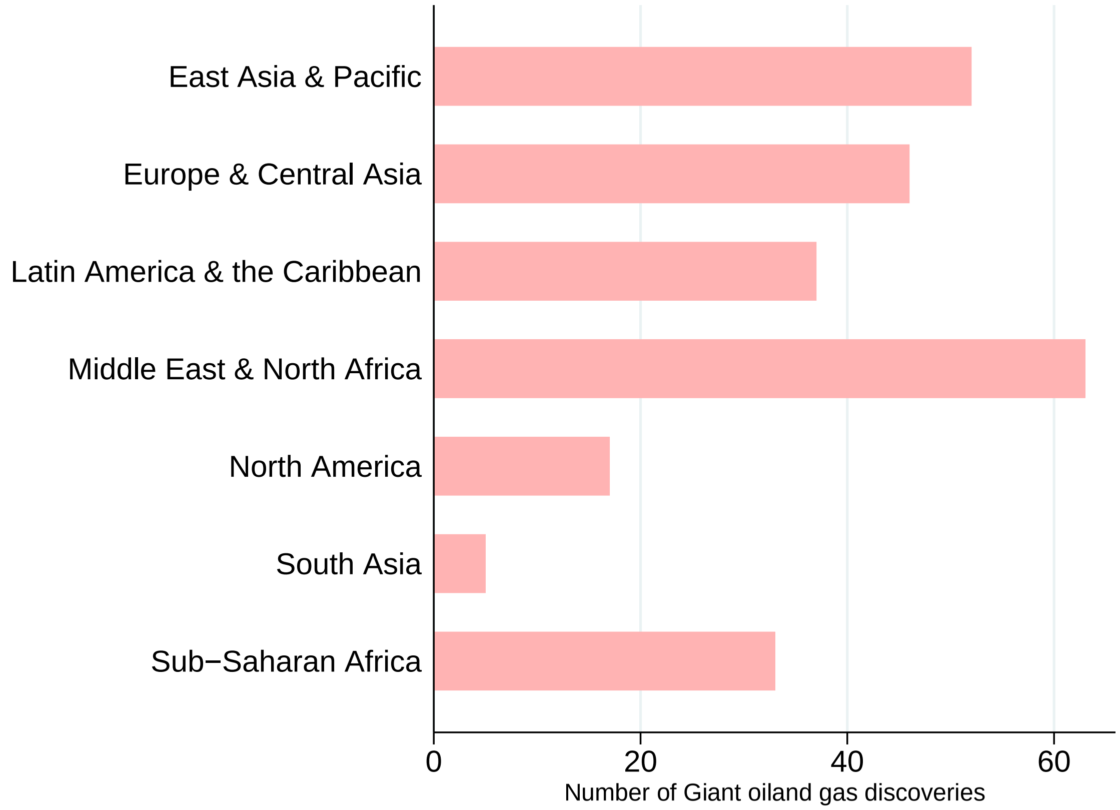


Figure 3: **Number of discoveries over regions (1989-2012)**

Notes: The country regional groupings are based on the classification of the World Bank.

Table 1: Summary Statistics (1989 to 2012)

	Mean	Min	Max
Giant oilfield discovery	0.04	0.00	1.00
Non-Resource Tax Revenues			
Non-Resource Tax Revenues (% of GDP)	16.00	0.09	53.87
Direct Tax Revenues (% of GDP)	6.22	0.00	25.02
Indirect Tax Revenues (% of GDP)	10.00	0.02	42.97
Taxes on International Trade (% of GDP)	2.68	-1.57	36.12
Other indirect taxes (% of GDP)	0.40	-2.11	7.91
Other variables			
Real GDP per capita	11140.79	115.79	111968.35
Population Growth	1.58	-6.18	16.33
Trade openness (millions)	-1295.20	-837288.00	359886.09
Non-Resource openness (millions)	1.54	0.00	2568.79
Level of democracy	3.00	-10.00	10.00
Corruption	2.96	0.00	6.00
Financial development	0.27	0.00	1.00

2.2 Identification Strategy

The identification strategy of the paper is based on the exogeneity of the timing of giant oil discoveries. Thus, the timing of these discoveries is considered to be independent of country specific characteristics. In this sub-section, I follow others in the literature to argue that this is the case.

Upstream petroleum exploitation can be broken down into four main stages: Exploration & Appraisal, Development, Production and Decommissioning. Exploration basically means to search for oil. Host governments can explore for oil on their own through a state agency or grant licenses to International Oil Companies (IOCs). Various factors influence the decision to undertake an exploration activity such as the geology, institutions, political and economic stability among others.[4] In addition, exploration of oil and gas (and other extractive natural resources) is affected by technological innovation and by the relative knowledge of geological features of a particular field (such as the structure of the oil field, the depth, location, etc.). This therefore makes exploration an uncertain activity. Only 2% of the total number of exploratory wells lead to a giant discovery and the relationship between the intensity of exploration drilling activity and making a giant discovery is rather weak (Toews and Vezina, 2016). Once some fields have been identified as potentially containing viable oil/gas resources, they are examined in more detail to establish their commercial viability. This is known as appraisal. Exploration activities can last for 1 to 5 years.

The development phase sets in when a commercial discovery is made in the previous stage, otherwise the operations are terminated. The development phase involves putting in the necessary infrastructure after a commercial discovery have been made to bring out the oil. Production occurs when the oil and gas is extracted for distribution or processing. The time lag from discovery to production can take on average 4 to 6 years. Production activities can last for at least 20 years on average. Finally, once it is no longer economically feasible to produce from the field, the site is returned to its original state as close as possible. This is known as decommissioning.

From the foregoing, it can be argued that the characteristic (s) of a country is likely to determine whether an exploration activity takes place or not but less likely to determine whether a giant discovery will be made. It is difficult to predict the timing of such a discovery from a country's (and companies') point of view and thus can be viewed as an unanticipated

[4]These licenses can be granted through direct negotiations or through an open and competitive bidding process.

news shock (see Arezki et al., 2017). One might argue that oil discoveries are somewhat predictable because some countries have larger oil endowments or they had discoveries in the past. Lei and Michaels (2014) showed that giant oilfield discoveries made in the recent past can predict whether a discovery can be made in the subsequent years or not.[5] This is also likely to work in the opposite direction: past discoveries can increase the cost of discovery thereby rendering future discoveries less likely (Arezki et al., 2017). Even in both cases, the exact timing of the giant oil discovery is less likely to be predictable. However, I account for the number of years with at least one giant discovery in the recent past in the empirical analysis.

2.3 Empirical Model

To examine the impact of giant oil and gas discoveries on domestic revenue mobilization from 1989 to 2012, the following econometric model is specified;

$$Y_{it} = \alpha_i + \theta_t + \sum_{j=-2}^{10} \beta_j Discovery_{it-j} + \gamma_1 pastdiscovery_{it} + \delta X_{it-1} + \epsilon_{it} \tag{2.1}$$

Where Y_{it} is tax revenues from the non-resource sector (expressed as a percentage of GDP) in country i at time t, α_i is a country dummy to account for country fixed effects, θ_t is a year dummy that accounts for time varying common shocks, $Discovery_{it}$ is a dummy variable equal to 1 if a country i makes at least one giant oil and gas discovery at time t. As discussed earlier, there is a significant lag between discovery and production which suggests that the effect might not be immediate but may take time to materialize and last for a period. I exploit these dynamics by including ten lags of the discovery variable. Also, there should not be any effect of giant oilfields discovery on non-resource revenues prior to the discovery at time t, conditional on discoveries made in the past, for the identification strategy to be valid. I include two leads of the discovery variable to check for this. This implies that β_j comprises 13 parameters because I have 13 Discovery dummies. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery.[6] As indicated in the previous section this accounts for the influence of discoveries made in the recent past on subsequent discoveries in the future.

[5]The authors showed that the likelihood of a giant discovery in year t increases from about one percent when there was no giant oilfield discoveries in the past 10 years to 87 percent if at least one giant oilfield discovery is made every year in the past 10 years.

[6]I compute this variable for each year and country in the study using the data from Horn (2011).

X_{it-1} is a vector of control variables. In the benchmark results, I control for the level of development (real GDP per capita and population growth) and the structure of the economy. The other control variables described in 2.1 were omitted because some countries had several missing observations. I however, include them in the robustness tests in section 4. To limit endogeneity issues, the control variables enter the regression with a lag.

To investigate the general pattern of what happens in the period after discovery and the production periods as opposed to what happened at a particular point in time, I estimate a version of model (2.1);

$$Y_{it} = \alpha_i + \theta_t + \beta_1 preproduction_{it} + \beta_2 production_{it} + \beta_3 pastdiscovery_{it} + \delta X_{it-1} + \epsilon_{it} \tag{2.2}$$

Where preproduction is a dummy variable which takes the value of 1 if there was at least one giant oilfield discovery between t-4 and t. Notice that the average time from discovery to production is 5 years. preproduction therefore captures what happens to non-resource tax revenues during the average five-year period it takes to start production. production is a dummy variable which takes the value of 1 if at least one giant oilfield discovery was made between t-10 and t-5. This variable on the other hand captures what happens on average 5 years after actual production commences i.e., between 5 to 10 years after at least one giant discovery was made.

3 Empirical Results

The main aim of this section is to document the causal impact of giant oil and gas discoveries on non-resource tax revenues. I present and discuss the results from models 2.1 and 2.2 with total non-resource tax revenues as the dependent variable in section 3.1. In section 3.2, I decompose total non-resource tax revenues into direct and indirect components. I conclude the section by documenting the main channels through which the non-resource tax revenues are affected by giant oil and gas discoveries in section 3.3.

3.1 Effect of oil discovery on Non-Resource Tax Revenues

In Figure 4, I graphically display the empirical estimates of the β coefficients from equation (2.1) with non-resource tax revenues as the dependent variable. The value of the beta coefficients is displayed on the y-axis while the periods (2 years before and 10 years after discovery) are shown in the x-axis. The first thing to notice is that the coefficients of the first two leads of the discovery variable are not statistically different from zero. This suggest that there is no anticipatory effect i.e., there is no statistically significant effect of giant oilfields discovery on non-resource revenue mobilization prior to the time of discovery.

Secondly, all the beta coefficients are positive. Specifically, non-resource tax revenues increased in the first year after discovery, declined in the second year, and stabilized in the 3rd and 4th years before increasing again from the fifth year onwards. The positive response for the first year after discovery is statistically different from zero while those of the other years are not statistically different from zero.

To ascertain what happens in general during the period following discovery and when production sets in opposed to what happened in a particular point in time, I report estimates of equation 2.2 in Table 2. For the sake of brevity, I only focus on column 3 which controls for the number of years with at least one giant discoveries in the recent past, real GDP per capita, population growth and the structure of the economy as done in Figure 4.

The results point towards a positive effect of a giant discovery on total non-resource tax revenues during both the average preproduction and production periods. These effects are also statistically different from zero. Non-resource tax revenues increase on average by 0.85 percentage points in the preproduction era, while it increase by about 0.67 percentage points during the average production era. Also, the discoveries made in the recent past has a positive and statistically significant effect on non-resource tax revenues.

The results as shown above contradict most of the results from the existing literature which normally points to a crowding out effect. As explained earlier, this crowding out effect could be due to the methodology employed in these studies. However, the data I use in the present study comes from the ICTD-GRD which is different from the data used in the existing studies. Could the contradiction in the result be due to the difference in the data and not the methodological difference? To show that this is not likely to be the case, I replicate the result of the existing studies using their methodology with the data from ICTD-GRD. The results are presented in Table A.1 in the Appendix. As can be seen, the results largely point out to a crowding out effect. Specifically, an additional percentage point increase in resource revenues (as % of GDP) decreases non-resource tax revenues (as % of GDP) and indirect non-resource revenues by 0.13 and 0.12 percentage points respectively. This is statistically significant. These estimates compare (although lower than) with those found in the existing literature. For example, while Bornhorst et al. (2009) found an offset of 0.2 percentage points, Crivelli and Gupta (2014) found an offset of 0.3 percentage points.

The foregoing therefore suggests that the differences in data sources is likely not a reason for the differences in the result obtained.

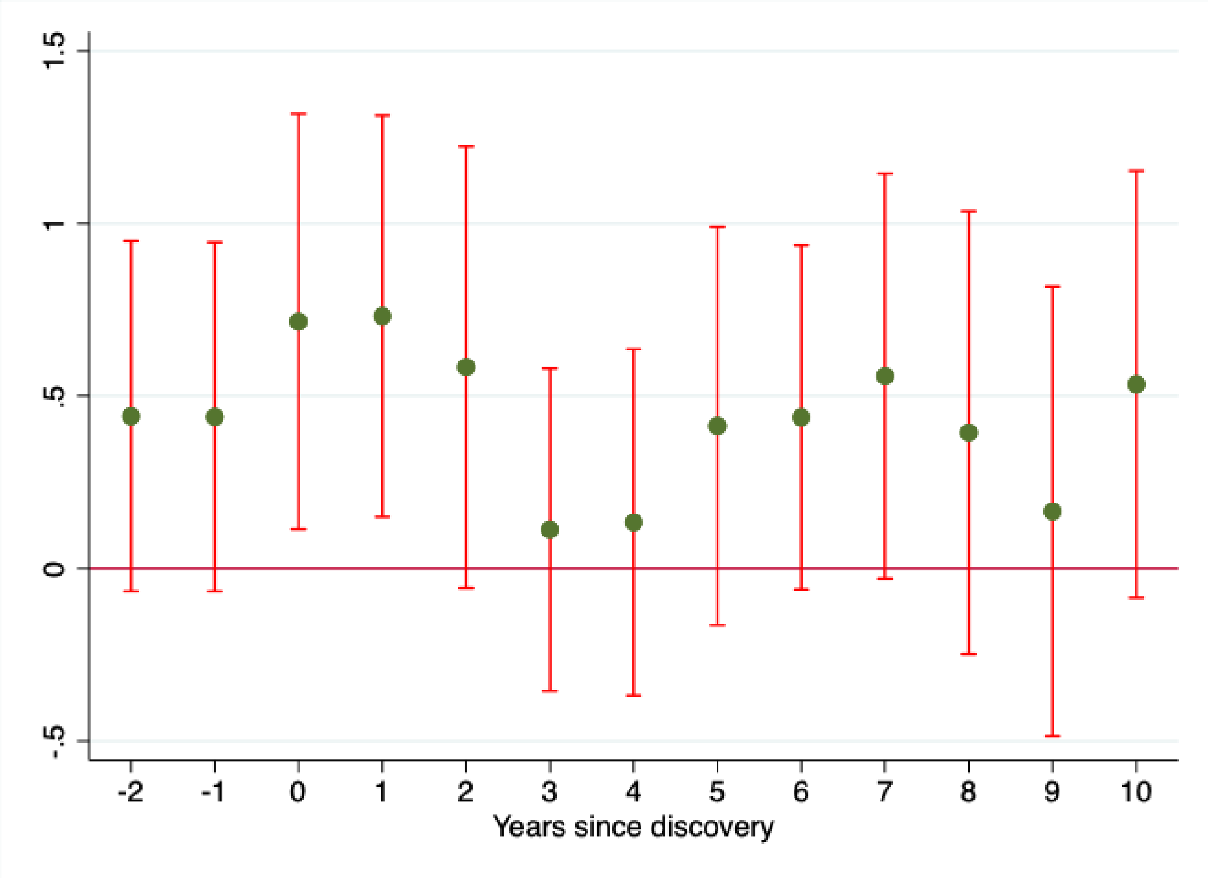


Figure 4: **Intertemporal effect of oil and gas discoveries on non-resource taxes**

Notes: The figure shows the effect of giant oil discoveries on non-resource taxes over time by including 2 leads and 10 lags (from t-2 to t+10) of the discovery variable. The dots show the point estimates, and the bars indicate 95% confidence intervals. The standard errors are clustered at the country level. All the regressions include country and year fixed effects. The regressions also control for the number of past years between t-20 and t-11 with at least one giant discovery, real GDP per capita, population growth and the structure of the economy.

Table 2: Effect of oil and gas discoveries on non-resource tax revenues

	1	2	3
preproduction	0.740***	0.893***	0.854***
	(0.211)	(0.212)	(0.216)
production	0.491***	0.700***	0.669***
	(0.189)	(0.200)	(0.207)
pastdiscovery		0.280***	0.267***
		(0.0723)	(0.0754)
Observation	3018	3018	3018
Adjusted R-squared	0.918	0.919	0.921
Year FE	✓	✓	✓
Country FE	✓	✓	✓
Other control variables			✓

*Notes: The dependent variable are non-resource tax revenues (% of GDP). All regressions include country and year fixed effects. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery. Other control variables include real GDP per capita, population growth and the structure of the economy. Robust standard errors, clustered at the country level, are reported in parentheses. *, ** and *** denotes significance at 10%, 5%, and 1% respectively.*

3.2 Effect of oil discovery on Non-Resource Direct and indirect Tax Revenues

It is widely known that the petroleum sector is an enclave and capital-intensive sector. This implies that relatively low number of jobs will be created directly per unit of capital invested. On the other hand, an oil discovery and its subsequent exploitation can lead to the import of materials and capital goods that can increase custom duties or spillover effects on local activity that can increase VAT. This suggests that different tax revenues might respond differently to a giant oil discovery. In this section, I decompose non-resource tax revenues into direct and indirect components. The ICTD-GRD data allows for such decomposition.

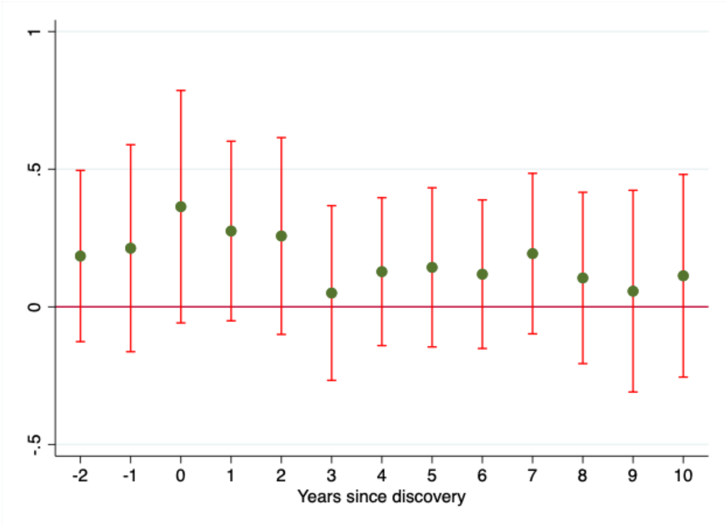
I first looked at the behavior of non-resource direct and indirect tax revenues over time by estimating model (2.1). The results are shown in Figure 5. For direct non-resource tax revenues, all the beta coefficients are not statistically different from zero although positive for all the years. On the other hand, the indirect component of non-resource tax revenues increased a year after a discovery, declined from the second up to the fourth year after discovery, and began to increase in the fifth year after discovery. Only the beta coefficients of the first and tenth year after discovery are different zero.

In addition, I examined the behavior of these two variables during the preproduction and production periods by estimating equation (2.2). The results are presented in Table 3. Using the preferred specification (column 3), the results indicates a positive and a statistically insignificant effect of giant discoveries on direct component of non-resources tax revenues in both the preproduction and production periods. The indirect component of non-resource revenues, on the other hand tends to increase during the preproduction period in all three specifications (see columns 4 to 6). During the preproduction and production periods, indirect non-resource tax revenues increase by 0.57 and 0.65 percentage points respectively, conditional on the structure of the economy, the level development of the economy and the number of years with at least one giant discoveries in the recent past (column 6). Taken, together the results suggest that giant oil and gas discoveries have a significant positive effect on total non-resource tax revenues which is driven by a significant positive effect on indirect non-resource tax revenues.

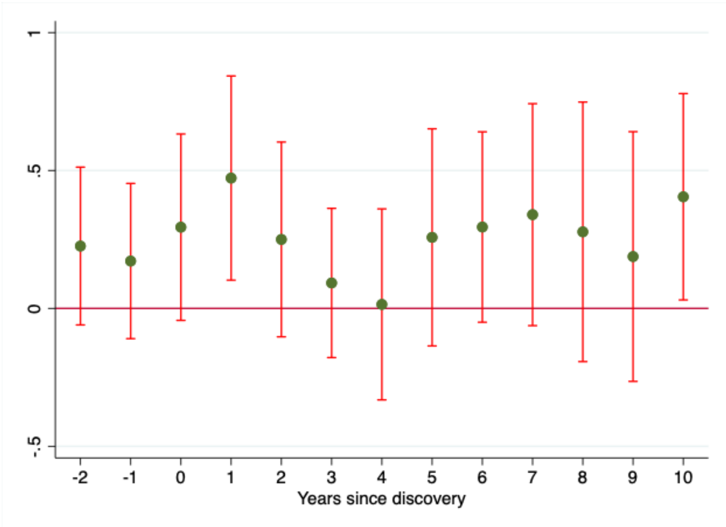
I further investigated the dynamics of the components of indirect taxes (tax revenues from goods & services, trade and other indirect tax revenues) to ascertain which of these components is driving the effect. The results are shown in Table 4. It can be seen that tax revenues from the consumption of goods and services experienced an increase in both the preproduction and production era. Trade tax revenues also experienced an increase during the average preproduction period.

Moreover, to further understand the results, I divide my sample into two. The first sample is made up of countries classified as high income countries by the World Bank. The second is made up of low and middle income countries also based on the classification by the World Bank. The main motivation for this division is that the same size of a giant oil discovery will have differential effect on a developed economy like USA compared to a developing economy like Ghana. The results are presented in Table (5). The results indicate that non-resource revenues decline in both the preproduction and production periods in high income countries. These effects are however not different from zero. Also, no statistically significant effect is found for direct and indirect non-resource tax revenues in the case of high income countries.

The general pattern of the results hold in the case of low and middle income countries. Specifically, both the total non-resource tax revenues and indirect non-resource tax revenues experience an increase in the preproduction and production era. Also, there is a statistically significant positive effect on direct tax revenues in this group of countries. The results therefore suggest that the effect of giant oil discoveries is significant on low and middle income countries as compared to high income countries.



(a) Direct Tax Revenues



(b) Indirect Tax Revenues

Figure 5: **Intertemporal effect of oil and gas discoveries on non-resource direct and indirect tax revenues**

Notes: The figure shows the effect of giant oil discoveries on non-resource direct(panel a) and indirect tax revenues (panel b) over time by including 2 leads and 10 lags (from t-2 to t+10) of the discovery variable. The dots show the point estimates, and the bars indicate 95% confidence intervals. The standard errors are clustered at the country level. All the regressions include country and year fixed effects. The regressions also control for the number of past years between t-20 and t-11 with at least one giant discovery, real GDP per capita, population growth and the structure of the economy.

Table 3: Effect of oil and gas discoveries on direct and indirect non-resource tax revenues

	Non-resource direct tax revenues			Non-resource indirect tax revenues		
	1	2	3	4	5	6
preproduction	0.127	0.140	0.157	0.498***	0.628***	0.572***
	(0.126)	(0.129)	(0.131)	(0.141)	(0.140)	(0.140)
production	-0.0603	-0.0421	0.0131	0.560***	0.738***	0.652***
	(0.105)	(0.109)	(0.114)	(0.141)	(0.147)	(0.150)
pastdiscovery		0.0245	0.0392		0.238***	0.210***
		(0.0420)	(0.0439)		(0.0438)	(0.0463)
Observation	3018	3018	3018	3018	3018	3018
Adjusted R-squared	0.929	0.929	0.930	0.882	0.883	0.885
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Other control variables			✓			✓

*Notes: The dependent variables are non-resource direct tax revenues (% of GDP) columns 1 to 3 and non resource indirect tax revenues(% of GDP) columns 4 to 6. All regressions include country and year fixed effects. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery. Other control variables include real GDP per capita, population growth and the structure of the economy. Robust standard errors, clustered at the country level, are reported in parentheses. * denotes significance at 10%, ** denotes 5%, and *** denotes 1%.*

Table 4: Effect of oil and gas discoveries on the components of indirect non-resource tax revenues

	1	2	3
preproduction	0.418***	0.243***	0.0547
	(0.125)	(0.0693)	(0.0403)
production	0.378***	-0.0188	-0.0421
	(0.130)	(0.0822)	(0.0534)
pastdiscovery	0.163***	0.00168	-0.0367*
	(0.0460)	(0.0232)	(0.0204)
Observation	2902	2903	2724
Adjusted R-squared	0.873	0.844	0.587
Year FE	✓	✓	✓
Country FE	✓	✓	✓
Other control variables	✓	✓	✓

*Notes: The dependent variables are tax revenues from goods & services (column 1), trade tax revenues (column 2) and other indirect tax revenues (column 3). They are all measured in percent of GDP. All regressions include country and year fixed effects. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery. Other control variables include real GDP per capita, population growth and the structure of the economy. Robust standard errors, clustered at the country level, are reported in parentheses. * denotes significance at 10%, ** denotes 5%, and *** denotes 1%.*

Table 5: **High Income vs Low-Middle Income countries**

	High Income			Low and Middle Income		
	nrtax	direct	indirect	nrtax	direct	indirect
preproduction	-0.269	-0.278	-0.00770	1.365***	0.385***	0.815***
	(0.401)	(0.253)	(0.221)	(0.231)	(0.138)	(0.172)
production	-0.326	-0.0670	-0.239	1.104***	0.141	0.959***
	(0.418)	(0.230)	(0.267)	(0.221)	(0.128)	(0.169)
pastdiscovery	-0.0268	-0.107	0.0958*	0.387***	0.112**	0.243***
	(0.106)	(0.0799)	(0.0530)	(0.0937)	(0.0475)	(0.0642)
Observation	937	937	937	2081	2081	2081
Adjusted R-squared	0.953	0.958	0.935	0.868	0.808	0.864
Year FE	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓
Other control variables	✓	✓	✓	✓	✓	✓

*Notes: The table presents the estimates of equation (2.2) for high income countries and low-middle income countries. nrtax is non-resource tax revenues; direct is non-resource direct tax revenues and indirect is non-resource indirect revenues. All regressions include country and year fixed effects. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery. Other control variables include real GDP per capita, population growth and the structure of the economy. Robust standard errors, clustered at the country level, are reported in parentheses. *, ** and *** denotes significance at 10%, 5%, and 1% respectively.*

3.3 Mechanisms

What could possibly account for the positive effect on revenues generated from taxing the consumption goods and services? The tax rate or the tax base? In order to answer this question, I compute the response of consumption (the tax base) and the consumption tax rate to a giant oil discovery shock by estimating model (2.2). The consumption tax is computed as:

$$\begin{aligned} Consumption \quad tax \quad rate &= \frac{tax \quad revenue \quad on \quad goods \quad \& \quad services}{Consumption} \\ &= \frac{Consumption \quad tax \quad rate * Consumption}{Consumption} \end{aligned} \tag{3.1}$$

Consumption is defined as real final consumption expenditures. The results are presented in Table 6. As can be seen the effect on the tax rate in both the preproduction and production periods are not different from zero. Consumption (the tax base for goods & services) on the other hand tend to increase in both the preproduction and production periods. The result suggests the positive effect of oil discovery on tax revenues from goods and services is largely due to an increase in the tax base (consumption) but not a change in tax policy (tax rate).

There are several potential ways through which an oil discovery could widen indirect tax base in both the preproduction and production periods. In the pre-production period (i.e the period between discovery and production), the necessary infrastructure is put in place to extract the oil. The economy can potentially benefit from imports of materials and capital goods that increase customs duties and VAT revenues. Also, the government could use the oil in the ground as collateral to borrow to smooth consumption. Furthermore, due to demand shocks at the local level in particular, goods and services can be demanded locally due to backward linkages leading to agglomeration effects and potentially increasing the VAT base (Aragón et al., 2015). In a recent study, Cavalcanti et al. (2019) found a positive spillover effects of oil discoveries on the local communities in Brazil due to a higher local demand for non-tradable services.

Table 6: **Effect of oil and gas discoveries on consumption and average tax rate**

	Consumption tax rate	Consumption (logs)
preproduction	-0.0102	0.0293**
	(0.0137)	(0.0124)
production	-0.00901	0.0242**
	(0.0141)	(0.0122)
pastdiscovery	-0.0251***	0.00553
	(0.00812)	(0.00398)
Observation	2626	2626
Adjusted R-squared	0.511	0.998
Year FE	✓	✓
Country FE	✓	✓
Other control variables	✓	✓

*Notes: The dependent variables are consumption tax rate (as defined in equation (3.1)) and the log of real consumption. Consumption is defined as the final consumption expenditures in constant local currency units. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery. All regressions include country and year fixed effects. The regressions also control for real GDP per capita and population growth. Robust standard errors, clustered at the country level, are reported in parentheses. *, ** and *** denotes significance at 10%, 5%, and 1% respectively.*

4 Robustness checks

The results so far suggest that the discovery of giant oilfields discovery does have a positive and statistically significant effect on total non-resource tax revenues in both the preproduction and production periods. In this section, I conduct several checks to test the robustness of this result. All the results are reported in Appendix B.

4.1 Additional control variables

First, as explained in section 2.3, several variables that can potentially affect tax revenues were omitted from the model due to missing data for some countries. These variables are trade balance, level of democracy, corruption, and financial openness. I include these variables as additional controls in model 2.2. The results are shown in Table B.1. As can be seen from the table, the number of observations reduce drastically. However, the result that giant oilfield discovery has a positive and statistically significant effect on total non-resource tax revenues still holds.

4.2 Different measures of giant oilfield discoveries

Secondly, I make use of different measures for giant oilfield discoveries. I first test if the result is sensitive to the location of the discovery. Discoveries can be made offshore (which is the typical case) or onshore. I also exclude from the sample discoveries that were made in the past three years and discoveries that were made subsequently to each other. This enables me to test if the potential endogeneity of discoveries that were made after the initial ones could affect the results. The results are reported in Table B.2. The results indicates that giant oilfield discovery has a positive and statistically significant effect on total non-resource tax revenues in both the preproduction and production era for all the different measures for giant oilfield discoveries.

4.3 Alternate dependent variables

Finally, I make use of three alternative measures of the dependent variable: the growth rate of non-resource tax revenues, non-resource tax revenue per capita and non-resource tax revenues to total tax revenues, to examine if the baseline results are sensitive to alternative measures of the dependent variables. As can be seen from Table B.3, a positive and statistically significant is found for all three alternative measures in both the pre-production and production periods.

5 Conclusions

In this paper, I investigate the effect of natural resources on tax revenue mobilization by exploiting the exogenous variation in the timing of giant oilfields discovery. This approach allows me to deal with the endogeneity of natural resources as the timing of the discovery of giant oilfields is arguably exogenous. In addition, this approach allows me to directly examine the performance of non-resource tax revenue mobilization before and immediately after discovery as well as the period corresponding to the inflow of revenues from the production of oilfields.

I do find that non-resource tax revenues experience an increase in both the pre-production and production periods. This result is robust to alternative measures of giant oil discoveries. This is due to an increase in indirect tax revenues in both the preproduction and the production era. Further analysis reveals that the positive effect on indirect non-resource tax revenues is due to an increase in the tax revenue mobilized from the consumption of goods and services. This is mainly driven by an increase in consumption of goods and services.

Taken together the result suggest that the crowding effect found in the literature, could be due to the way the issue has been investigated, where non-resource revenues (as % of GDP) is normally regressed on resource revenues (% of GDP). The results therefore suggest that the abundance of natural resources may not be a reason why countries mobilize less tax revenues.

References

Aragón, F. M., Chuhan-Pole, P., and Land, B. C. (2015). *The local economic impacts of resource abundance: What have we learned?* The World Bank.

Arezki, R., Ramey, V. A., and Sheng, L. (2017). News shocks in open economies: Evidence from giant oil discoveries. *The quarterly journal of economics*, 132(1):103–155.

Besley, T. and Persson, T. (2013). Taxation and development. In *Handbook of public economics*, volume 5, pages 51–110. Elsevier.

Besley, T. and Persson, T. (2014). Why do developing countries tax so little? *Journal of Economic Perspectives*, 28(4):99–120.

Bornhorst, F., Gupta, S., and Thornton, J. (2009). Natural resource endowments and the domestic revenue effort. *European Journal of Political Economy*, 25(4):439–446.

Cavalcanti, T., Da Mata, D., and Toscani, F. (2019). Winning the oil lottery: The impact of natural resource extraction on growth. *Journal of Economic Growth*, 24(1):79–115.

Collier, P. (2006). Is aid oil? an analysis of whether africa can absorb more aid. *World development*, 34(9):1482–1497.

Corden, W. M. (1984). Booming sector and dutch disease economics: survey and consolidation. *oxford economic Papers*, 36(3):359–380.

Corden, W. M. and Neary, J. P. (1982). Booming sector and de-industrialisation in a small open economy. *The economic journal*, 92(368):825–848.

Crivelli, E. and Gupta, S. (2014). Resource blessing, revenue curse? domestic revenue effort in resource-rich countries. *European Journal of Political Economy*, 35:88–101.

Harding, T., Stefanski, R., and Toews, G. (2020). Boom goes the price: Giant resource discoveries and real exchange rate appreciation. *The Economic Journal*, 130(630):1715–1728.

Horn, M. K. (2011). Giant oil and gas fields of the world. https://worldmap.harvard.edu/data/geonode:giant_oil_and_gas_fields_of_the_world_co_yxz.

Lei, Y.-H. and Michaels, G. (2014). Do giant oilfield discoveries fuel internal armed conflicts? *Journal of Development Economics*, 110:139–157.

Moore, M. (1998). Death without taxes: Democracy, state capacity, and aid dependence in the fourth world. *The democratic developmental state: Politics and institutional design*, 86.

Moore, M. (2007). How does taxation affect the quality of governance. *Tax Notes International*, 47(1):79–98.

Ossowski, R. and Gonzáles, A. (2012). Manna from heaven: The impact of nonrenewable resource revenues on other revenues of resource exporters in latin america and the caribbean. Technical report, IDB Working Paper Series.

Perez-Sebastian, F., Raveh, O., and van der Ploeg, F. (2021). Oil discoveries and protectionism: Role of news effects. *Journal of Environmental Economics and Management*, 107:102425.

Prichard, W., Cobham, A., and Goodall, A. (2014). The ictd government revenue dataset.

Svirydzenka, K. (2016). *Introducing a new broad-based index of financial development*. International Monetary Fund.

Thomas, M. A. H. and Trevino, M. J. P. (2013). *Resource dependence and fiscal effort in Sub-Saharan Africa*. Number 13-188. International Monetary Fund.

Toews, G. and Vezina, P.-L. (2016). Drilling activity and giant oil and gas discoveries. Technical report, MIMEO.

Van der Ploeg, F. and Poelhekke, S. (2010). The pungent smell of "red herrings": Subsoil assets, rents, volatility and the resource curse. *Journal of Environmental Economics and Management*, 60(1):44–55.

World Bank (2022). *Global Economic Prospects*. . Washington, DC: World Bank. doi: 10.1596/978-1-4648-1758-8. License: Creative Commons Attribution CC BY 3.0 IGO.

Appendix A: Additional Tables

Table A.1: Replication of existing studies

	(1)	(2)	(3)
resource revenue	-0.128***	-0.031	-0.119***
	(0.048)	(0.027)	(0.044)
logdppercapita	-0.340	-0.576*	0.122
	(0.346)	(0.298)	(0.210)
nropeness	0.000**	0.000*	-0.000
	(0.000)	(0.000)	(0.000)
agriculture	-0.089**	-0.041	-0.060
	(0.043)	(0.030)	(0.039)
corruption	0.473**	0.238*	0.145
	(0.206)	(0.143)	(0.134)
Constant	16.388***	4.741***	11.843***
	(2.221)	(1.599)	(1.325)
Observations	1,076	1,025	1,035
R-squared	0.115	0.155	0.095
Number of Countries	76	72	72
Control Variables	Yes	Yes	Yes

Notes: The table presents the regression of non-resource tax revenues on resource revenues from 1989 to 2012. The dependent variable in column 1 is non-resource tax revenues, non-resource direct tax revenues in column 2 and non-resource indirect revenues in column 3. logdppercapita is the log of real GDP per capita; nropeness is the non-resource exports plus non-resource imports to GDP ratio; Agriculture is the agriculture value added (% of GDP) and corruption is a 6-point corruption risk score from the International Country Risk Guide (ICRG). All regressions include country and year fixed effects. Robust standard errors clustered at the country level, are reported in parentheses. *, ** and *** denotes significance at 10%, 5%, and 1% respectively.

Table A.2: Variables definition and source

Variable	Definition	Source
Giant oilfield discovery	A dummy variable that indicates whether a country made at least one giant oil or gas field discoveries during the period under consideration	Horn (2011)
Non-Resource Tax Revenues		
Non-Resource Tax Revenues (% of GDP)	Non-resource tax excluding social contributions expressed as percentage of GDP	ICTD-GRD
Direct Tax Revenues (% of GDP)	Direct taxes excluding social contributions and resource revenue	ICTD-GRD
Indirect Tax Revenues (% of GDP)	Non-Resource Component of Indirect Tax expressed as percentage of GDP	ICTD-GRD
Taxes on International Trade (% of GDP)	Total taxes on international trade and transactions expressed as percentage of GDP	ICTD-GRD
Other indirect taxes (% of GDP)		ICTD-GRD
Other variables		
Real GDP per capita	GDP per capita divided by midyear population in constant local currency unit	WDI
Population Growth	Annual population growth rate	WDI
Trade openness (millions)	Difference between exports and imports of goods in millions	WDI
Non-Resource openness (millions)	Difference between non-resource export and non-resource imports to GDP ratio (in millions)	WDI
Level of democracy	Proxied with polity score. The polity score is computed by subtracting the $p_{autocracy}$ score from the $p_{democracy}$ score; the resulting unified polity scale ranges from +10 (strongly democratic) to -10 (strongly autocratic)	Center for systemic Peace
Corruption	A score based on the assessment of corruption within the political system.This corruption risk index has a minimum score of zero and maximum score of six	ICRG
Financial development	A broad-based measure that considers the depth, access, and efficiency of financial institutions which ranges from zero to one. Higher values means higher degree of financial openness	Svirydzenka (2016)

Data Sources: ICTD GRD: International Center for Tax and Development Government Revenue Dataset, WDI: World Bank Development Indicators and ICRG: International Country Risk Guide

Appendix B: Robustness Check Results

Table B.1: Effect of giant oil discoveries on non resource tax revenues: Additional control variables

	nrtax	direct	indirect
preproduction	0.641***	0.113	0.480***
	(0.215)	(0.133)	(0.138)
production	0.622***	-0.0281	0.579***
	(0.208)	(0.116)	(0.150)
pastdiscovery	0.133**	-0.00574	0.130***
	(0.0662)	(0.0422)	(0.0400)
Observation	2154	2154	2154
R-squared	0.933	0.942	0.889
Year FE	✓	✓	✓
Country FE	✓	✓	✓
Other control variables	✓	✓	✓

*Notes: The table presents the regression of non-resource tax revenues on resource revenues from 1989 to 2012 with additional control variables. The dependent variable in column 1 is non-resource tax revenues, non-resource direct tax revenues in column 2 and non-resource indirect revenues in column 3. The additional control variables are trade balance, level of democracy, corruption and financial openness. nrtax is non-resource tax revenues; direct is non-resource direct tax revenues and indirect is non-resource indirect revenues. All regressions include country and year fixed effects. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery. Other control variables include real GDP per capita, population growth and the structure of the economy. Robust standard errors, clustered at the country level, are reported in parentheses. *, ** and *** denotes significance at 10%, 5%, and 1% respectively.*

Table B.2: Robustness Tests using alternative measures of giant oilfield discoveries

	Onshore discoveries	Offshore discoveries	No discovery in the past 3 years	Non-sequential discoveries
preproduction	0.908***	0.923***	0.699***	0.830***
	(0.250)	(0.256)	(0.200)	(0.195)
production	1.000***	1.071***	0.738***	0.564***
	(0.300)	(0.309)	(0.204)	(0.200)
pastdiscovery	0.276***	0.273***	0.269***	0.294***
	(0.0722)	(0.0755)	(0.0757)	(0.0759)
Observation	3018	3018	3018	3018
Adjusted R-squared	0.919	0.921	0.921	0.921
Year FE	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Other control variables	✓	✓	✓	✓

Notes: The table shows the effect of different measures of giant oil discoveries on non-resource tax revenues. All regressions include country and year fixed effects. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery. Other control variables include real GDP per capita, population growth and the structure of the economy. Robust standard errors, clustered at the country level, are reported in parentheses., ** and *** denotes significance at 10%, 5%, and 1% respectively.*

Table B.3: Robustness Tests using alternative measures of the dependent variable

	1	2	3
preproduction	8.570*	0.0653*	0.0684***
	(4.979)	(0.0379)	(0.00857)
production	16.21***	0.0522*	0.0752***
	(5.994)	(0.0293)	(0.00850)
pastdiscovery	4.908***	-0.00210	0.00883***
	(1.472)	(0.0119)	(0.00254)
Observation	3018	3018	3018
Adjusted R-squared	0.678	0.0361	0.107
Year FE	✓	✓	✓
Country FE	✓	✓	✓
Other control variables	✓	✓	✓

*Notes: The table shows the effect of giant oil and gas discoveries on different measures of the dependent variable. Non-resource tax revenues per capita in column 1, growth rate of non-resource tax revenues in column 2 and tax structure (non-resource tax revenues to total tax revenues) in column 3. All regressions include country and year fixed effects. pastdiscovery is the number of past years between t-20 and t-11 with at least one giant discovery. Other control variables include real GDP per capita, population growth and the structure of the economy. Robust standard errors, clustered at the country level, are reported in parentheses. *, ** and *** denotes significance at 10%, 5%, and 1% respectively.*

Appendix C: List of countries in the sample[7]

Afghanistan, Albania, **Algeria**, **Angola**, Anguilla, Antigua and Barbuda, **Argentina**, Armenia, Aruba, **Australia**, Austria, **Azerbaijan**, Bahamas, The, Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bhutan, **Bolivia**, Bosnia and Herzegovina, Botswana, **Brazil**, Brunei Darussalam, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, **Canada**, Cape Verde, Central African Republic, Chad, Chile, **China**, **Colombia**, Comoros, Congo, Dem. Rep., **Congo, Rep.**, Costa Rica, Cote d'Ivoire, Croatia, Cuba, **Cyprus**, Czech Republic, **Denmark**, Djibouti, Dominica, Dominican Republic, Ecuador, **Egypt,**, El Salvador, **Equatorial Guinea**, Eritrea, Estonia, **Ethiopia**, Fiji, Finland, France, Gabon, Gambia, The, Georgia, **Ghana**, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong SAR, China, Hungary, Iceland, **India**, **Indonesia**, **Iran, Islamic Rep.**, **Iraq**, Ireland, **Israel**, Italy, Jamaica, Japan, Jordan, **Kazakhstan**, Kenya, Kiribati, Kosovo, **Kuwait**, Kyrgyz Republic, Lao PDR, Latvia, Lebanon, Lesotho, **Liberia**, **Libya**, Lithuania, Luxembourg, Macao SAR, China, Macedonia, FYR, Madagascar, Malawi, **Malaysia**, Maldives, Mali, Malta, Marshall Islands, Mauritania, Mauritius, Mexico, Micronesia, Fed. Sts., Moldova, Mongolia, Montenegro, Montserrat, Morocco, **Mozambique**, **Myanmar**, Namibia, Nauru, Republic of, Nepal, Netherlands, New Zealand, Nicaragua, Niger, **Nigeria**, **Norway** ,**Oman**, **Pakistan**, Palau, Panama, Papua New Guinea, Paraguay, **Peru**, **Philippines**, Poland, Portugal, **Qatar**, Romania, Rwanda, Samoa, San Marino, Sao Tome and Principe, **Saudi Arabia**, Senegal, Serbia, Seychelles, **Sierra Leone** , Singapore, Slovak Republic, Slovenia, Solomon Islands, Somalia, South Africa, South Sudan, Spain, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, **Sudan** , Suriname, Swaziland, Sweden, Switzerland, Syrian Arab Republic, Tajikistan, **Tanzania**, **Thailand**, Timor-Leste, Togo, Tonga, **Trinidad and Tobago**, Tunisia, Turkey, **Turkmenistan**, Tuvalu, Uganda, Ukraine, United Arab Emirates, **United Kingdom**, **United States**, Uruguay, Uzbekistan, Vanuatu, **Venezuela**, **Vietnam**, West Bank and Gaza, Yemen, Rep., Zambia, Zimbabwe

[7] Countries with at least one giant oilfield discovery during the period under consideration are in bold.