Norwegian School of Economics

Bergen, fall, 2015

NHH

Oil exploration and institutions

Do oil companies prefer to drill in democracies?

Guro Lystad

Supervisor: Torfinn Harding

Master thesis, Master of Science in Economics and Business Administration, major in Economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

In this master thesis, I investigate the impact institutional quality has on exploration and drilling activities in the oil and gas industry. I use ordinary least square regressions (OLS) and two stage least square regressions (2SLS) to perform a cross-sectional analysis on a sample of 86 countries. The analysis examines the impact institutional quality, measured by the Polity IV index, has on three different variables; new field wildcat wells (NFW), the number of days spent drilling in a country and the success rate from new field wildcats.

The findings indicate a positive and significant relationship between institutional quality and drilling activity when measured by NFW wells drilled and the number of drilling days. This supports the theory that oil companies are risk adverse and prefer to drill in countries where there is low political risk and stable institutions. When looking at the discovery rate, it turns out that the probability of finding oil is negatively correlated with institutional quality. A possible explanation could be that oil companies only chose unstable countries when the likelihood of finding oil is very high. The strong consistency in the results, even to large changes in the dataset, suggests that institutional quality is a determinant of oil exploration.

Preface

This master thesis was written as a part of a double master's degree in Economics and Sustainable Development at the Norwegian School of Economics and HEC Paris. Writing a thesis while doing a double degree has not been easy, and I am very content to finally submit this paper after a process that has taken nearly two years altogether.

With a background from development studies and foreign aid assistance, I started working on this thesis with a certain set of expectations about the extraction of resources in developing countries. But, as I completed an internship at an international oil company, and saw the results of my analyses, I had to face some of my old prejudice against the oil industry. In the end, the fact that oil companies behave rationally and try to avoid unnecessary risk might not be that surprising, but to me, it was quite unexpected.

I would like to take the opportunity to thank the people at Statoil ASA in the department for Exploration and Valuation, where I did my summer internship in 2014. Not only did they provide me with access to global data bases such as Wood Mackenzie and IHS, they also arranged for interviews with the risk analysis team in Stavanger and provided valuable insight and knowledge about the oil and gas industry. I am also grateful to my current employer, THEMA Consulting Group, who accommodated me with office space and free lunch all through the summer of 2015 when writing the thesis. Other people worth mentioning are Samantha Stephens, Anders Lund Eriksrud and Marit Austgulen for their input and comments. Last, but not least, I wish to thank my supervisor, Torfinn Harding, for suggesting an interesting topic and for enthusiastic guidance and valuable discussions through the entire process.

Oslo, December 2015

Guro Lystad

Contents

Table of Contents

Figures

Tables

1. Introduction

In a decade where the oil price has surged and the discovery rate of conventional reserves have fallen (Crooks, 2015), oil companies have been searching all over the globe, looking for new drilling opportunities. The quest for oil and gas has resulted in exploration activities in costly and controversial areas, such as the Amazon rainforest or the Arctic (Vaughan, 2014) (Staalesen, 2011).

The perception amongst large petroleum actors and advisors, such as IHS and BP, seems to be that oil companies will have no objections against drilling in any country, as long as there is profit to make (Alon, et al., 2006). Even in Statoil, the policy is that all countries can (theoretically) be considered for oil and gas exploration, as long as there are no sanctions imposed through the Norwegian government (Klausen, et al., 2014).

When I asked my colleagues at Statoil whether they believed that countries with good institutional qualities and democracies had a higher drilling activity than countries with low institutional scores, most responded that they would presume rather the opposite.

The reaction is not uncommon. Most people would mention autocratic OPEC members in the Middle East, such as Saudi Arabia and the United Arab Emirates or corrupt regimes in Africa like Nigeria and Angola, when asked about listing oil-rich countries. This perception is supported by the theory about the "resource curse", which suggests a negative relationship between resource wealth and institutional quality (Norman, 2009).

Recent research, however, suggests a strong correlation between institutional quality and oil exploration. Bohn and Deacon (2000) discover a negative relationship between expropriation risk and oil exploration, while Cust and Harding (2014) find that when presented with equal geological conditions, investors prefer countries with higher institutional quality.

In this thesis I seek to explore what impact institutional quality has on oil exploration and drilling. The research complements the work of Cust and Harding by exploring the same hypothesis, but with a different method and dataset. While they investigate the difference in drilling activity between the borders of two countries, I use aggregated data from IHS Global

Insight and Wood Mackenzie to perform a cross-sectional analysis to investigate whether institutional quality may be an explanatory factor to drilling activity between 1970 and 2010. I also examine the relationship between oil discoveries and institutional quality to see whether the drilling activity correlates with a higher success rate of finding oil.

The findings from the regressions display a significant and positive relationship between the institutional quality in a country and the number of new field wildcat wells drilled. The results from an ordinary OLS regression indicates that, everything else held constant, an improvement in the institutional score from Polity IV with one point, could increase the number of NFW wells drilled with 7%. If we compare two countries, such as Brazil and Norway, the regression estimates suggest that drilling activity would be almost 50% higher in Norway due to the differences in institutional quality.

The same relationship occurs when investigating the correlation between the number of days spent drilling in a country and institutional quality. Increasing the institutional quality with one point suggests an increase of 9.5% in the number of days spent drilling in a country.

The resource curse theory suggest that oil exploration has a negative impact on institutional quality. To correct for potentially negative bias, I run a two-stage least square regression, using estimates of the Polity IV score before the drilling period, as an instrument. The results suggest an even stronger impact from institutional quality when excluding the unsystematic variance. An increase in institutional quality with one point, everything else held constant, would increase the drilling of NFW wells with 13%. The omitted variable bias in the ordinary OLS regressions seem to have a negative impact on the drilling activity, which supports the existence of a resource curse.

To see whether the higher drilling activity could be explained by better geological conditions and thus a bigger likelihood of discovering oil, I use the success rate as a dependent variable. It turns out that there is a significant and negative correlation between institutional quality and the success rate. When the institutional quality score increases with one point, the success rate drops with one percentage point. A reason for this is probably that oil companies have already exhausted the most accessible resources in low-risk countries with good institutions, while they only enter high-risk countries if the likelihood of finding oil is very high.

The results stay the same when run on a sample of high-income countries. When using a low-income sample, the institutional variable does not show signs of significance in any of the regressions. In low-income countries, the institutional system tends to be more fragile and change more often than in high-income countries (Collier & Hoeffler, 2009). Institutional quality or the form of government alone, may not be enough to provide a low-risk environment for oil investments in low-income countries.

To check the robustness of the results, I exchange the dependent variable of NFW wells drilled, and success rate, with similar variables from a different dataset. In spite of large regional differences, the results stay robust when looking at drilling activity. An increase in institutional quality by one point increases the number of wells drilled with 11% in the OLS regression. However, the negative correlation between success rate and institutions, does not hold anymore. The results are also robust to replacing the institutional variable and to excluding outliers. Overall, the strong consistency in the results, even to large changes in the dataset, suggests that institutional quality is a determinant of oil exploration.

The first section of this thesis presents the theoretical foundation of the analysis. This includes a literature review as well as an overview of the basics of oil exploration that investigates the rational behavior of an oil company and what kind of political risks as well as risk-mitigating measures affect an investment decision. The second part presents the dataset and variables used, and explains the methodology of the analysis and the measures taken to comply with the underlying assumptions of an OLS analysis. Finally, the main results are presented and analyzed, followed by other results and robustness tests with a conclusion at the end.

2. Theory

2.1 Literature review

The pressure of increasing the reserve-replacement ratio, an indicator used to evaluate company stocks that measures a company's ability to add new reserves to the portfolio, has made most oil producers ready to accept drilling in countries with a high-risk profile (Adams, 2015). The same notion can be found amongst large petroleum actors such as IHS and BP; oil companies will drill in almost any country in the world and rather mitigate the risks ex ante (Alon, et al., 2006).

Previous literature supports this view, claiming that in countries with weaker institutions and property rights, the extraction rate would be higher due to the risk of nationalization in the future (Long, 1975). Some economists suggest that international oil companies benefit from investing in fragile states, as weak institutions and low market regulations reduce government take and increase company profit. It could even function as a competitive advantage as they gain access to markets and resources where other companies might be deterred from investing. Jedrzej Frynas (1998) suggests that this has been the case for Shell, which gained a first-mover advantage in Nigeria. The companies handle the risk by employing private security firms to sustain law and order in the area in which they operate while the rest of the country suffers from civil unrest and lawlessness (Reno, 2001).

William Reno (2001) finds that the five Sub-Saharan states that received the most foreign direct investment in the 1990's, were also those with the highest risk ratings assigned by political risk analysts. Another common denominator amongst these five was a high share of mineral exports. Three of the countries, Angola, Nigeria and Equatorial Guinea, had oil as their primary export (Reno, 2001).

The idea of fragile, resource-dependent countries with corrupt leaders who hand over the country's oil wealth to international corporations, is closely linked to the theory of the resource curse.

Jeffrey Sachs and Andrew Warner (1995) first established the notion of the resource curse when they found that countries with a high ratio of exports of natural resources to GDP experienced slower economic growth compared to other resource-poor countries in the period from 1970 to 1990. They offer the explanation that resource-dependency led to de-industrialization in other sectors (the Dutch Disease) and made resource-oriented economies vulnerable to shocks in commodity prices (Sachs & Warner, 1995).

Another aspect of the resource curse theory, is that abundance of natural resources deteriorates institutional quality, which again reduces growth. Researchers have found a negative relationship between natural resources and rule of law (Norman, 2009), corruption (Leite & Weidmann, 1999) and government effectiveness (De Rosa & Iootty, 2012). A possible explanation for this relationship is that the wealth reduces the system of checks and balances between the government and the public. This can result in lavish spending, poor investment choices or a general negligence of other economic sectors that will reduce the contribution to economic growth (Norman, 2009).

Meanwhile, another part of the research suggests that countries associated with low risk and better institutional quality have been historically preferred for oil exploration (Jojarth, 2008). David and Wright (1997) claims that the increased exploration activity in the United States between 1850 and 1950 was not a result of geological prospects, but rather a favorable investment climate for new technologies, available markets and institutional and political structures.

Bohn and Deacon (2000) find that drilling and production was more extensive in countries associated with low ownership risk. Increased ownership risk is associated with lower democracy scores and reduced investment in the oil industry. Their discoveries are supported by Cust and Harding (2014) who detect that oil companies prefer to invest in countries with higher institutional quality when operating in countries with similar geological conditions.

The following section presents an overview of the process of discovering and extracting oil together with the typical risks an international oil company faces and the attempts it makes to mitigate the non-geological risks. Although the analysis in this thesis is based upon data of both national oil companies (NOCs) and international oil companies (IOCs), the theory is focused on investment decisions made by IOCs. This because the option of deciding between drilling opportunities in different countries does not apply in the same way for national oil companies, who mainly operate in their home countries.

2.2 Oil exploration and drilling

Ever since the first oil well was drilled in 1859 in Pennsylvania, people have tried to figure out where oil is located. While the early methods mostly involved searching the ground for oil leaking through the surface, the technology nowadays is highly advanced, not to mention costly, involving seismic studies and 3D mapping (Bret-Rouzaut & Favennec , 2007).

The process for locating and extracting oil is a long and costly process. It makes sense that investors are cautious before expanding into new territories. The following section goes briefly through the process of oil exploration and extraction to explain the different phases of oil production.

Figure 1: The life-cycle of oil exploration and production

Source: Cairn Energy (2015)

When deciding upon where to invest in oil and gas, oil companies have to reach some overall decisions on where they want to drill. They have to consider how much money they are willing to spend and how much risk they want to take on. This involves both geological and political risk. Even before a company sets foot in a country to search for oil, it has considered the potential risks concerning the stability of the region and the country. The company will also consider the general prospects of finding and producing oil. This includes the drilling costs and the quality of the oil, as well as the fiscal terms presented by the government. Several of these aspects are covered before an oil company decides to bid on an option to prospect. The option to prospect an area is normally given at auctions where companies compete against each other for licensing rounds. Those who win are given the right to explore an area for oil and gas resources (Collier & Venables, 2011).

When a company has acquired the right to search for oil, the next step is discovering the oil. This is a long process consisting of conducting several geological studies including seismic imaging for offshore exploration. The cost of these studies varies between 5 and 20 million USD per exploration site. The final stage of the exploration process is the drilling of an exploration well to determine whether there is oil and gas at the chosen site (BERA, 2005/2006). If the area has never been explored for oil and gas previously, the first wells are called new field wildcats. Drilling these wells involves a high degree of economic risk since it means exploring unknown geological territory (Murray, 1990). The drilling of an exploration well takes between 2-6 months depending on the geology on the site, the costs increase with the drilling period. The success of finding oil when drilling an exploration well is between 10-30% (Bret-Rouzaut & Favennec , 2007).

If the site during the exploration phase shows promising signs of containing oil and gas, the oil companies drill appraisal wells to get a better understanding of the size and characteristics of the discovery. The findings help the companies to determine whether the site is commercially viable and to decide on the strategy and investments needed to optimize the extraction of the fuels (Cairn Energy, 2015).

In the development phase, the oil companies plan the extraction process of the fuels. This includes investment analysis with risk adjustments where the company also has to decide on the amount of capital expenditures (capex) that will be spent in the oil field. There is usually a trade-off between the investments that have to be made and the oil recovery from the field: the more you invest, the larger oil recovery. Thus, the oil companies have to decide on the optimal solution for maximizing their profit. In this phase, the IOC will perform several analyses regarding the technological and economical risks associated with the project. The development plan has to be approved by local and central government in the operating country (Cairn Energy, 2015). During this phase the capital investments are very high, as the company drills the production and injection wells to extract the oil in addition to investments for storage, separation and treatment plants and transport such as a pipeline (Bret-Rouzaut & Favennec, 2007). The development phase is what makes the oil and gas industry more capital-intensive than other industries. The capital expenditure to revenue ratio (total capex/one year revenue) is 17% in oil and gas while in other industries it lies at 6-7% (Bret-Rouzaut & Favennec, 2007). This is also the reason why companies are concerned with the risk of nationalization. If a government decides to nationalize an oil field without compensating the company, large investment costs will be lost.

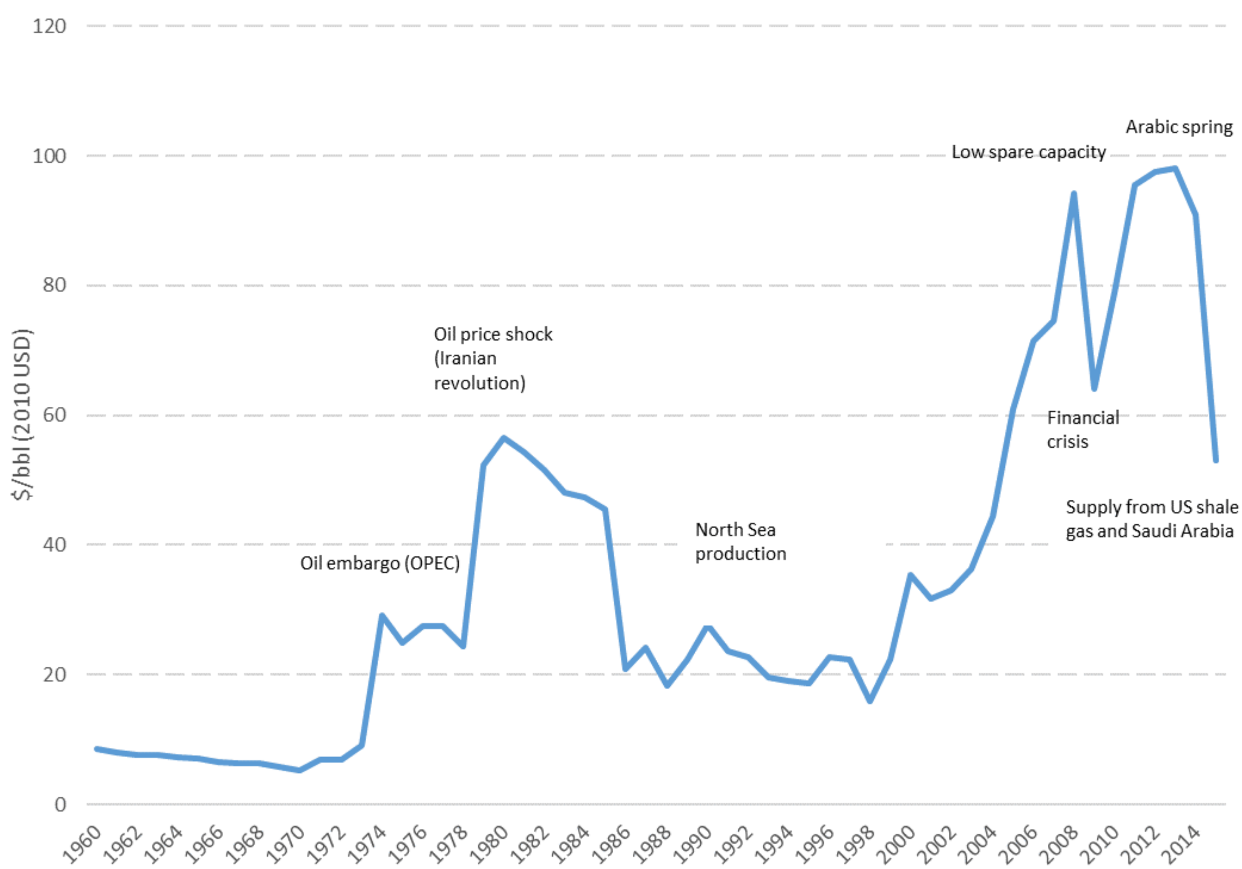
The production phase of an oil field can last for several decades, in this phase most of the costs have already occurred and only the operation costs remain. During production, even if the oil price drops, the extraction process will continue as long as the revenues exceed the operational costs including tax payments. The final stage of oil production is the decommissioning phase where the oil company has to clean up and close down the oil field (Cairn Energy, 2015).

2.3 The history of oil drilling

In the period between 1940 and 1960, oil exploration and production was driven by 7 large companies – the seven sisters- who have now merged into four companies: Exxon Mobil, Shell, BP and Chevron, controlled 83% of the world’s crude oil production in 1960 (Nøstbakken, 2013). When OPEC was founded in the 1960s, the national oil companies (NOCs) took over a large share of the production (Tsui, 2011). Nowadays, the landscape is more divided with NOCs in particular controlling around 90% of current oil reserves, while both large and small IOCs are competing over the remaining 10% (The Economist, 2013).

The quest for oil is mainly influenced by the oil price. When prices surge, oil companies are willing to spend more money searching for oil in areas where the costs are higher and where it would not be economically feasible to drill for oil with a lower oil price. The exploration costs per well increase because the companies have to access more complicated locations in addition to the high demand from other oil companies which increases the costs of renting equipment.

Figure 2: Development in the crude oil price 1960-2014



Source: The World Bank (2015), EIA (2015), Gauthier (2015)

The graph shows the historical development of the oil price during the last century. After the two oil price shocks in 1972 and 1978, Saudi Arabia ceased to function as a swing producer. This increased the production capacity and made the oil price fall. During the 1990s the oil price was at 20$ per barrel and both OPEC and IOCs started building down the production capacity. Before the financial crisis in 2008, the demand for oil was high, and capacity low, which made the prices surge again. This increased the exploration activity together with the costs (Gauthier, 2014).

In the fall of 2015, the oil price has been at 50$ per barrel. Oil companies have started reducing production and exploration due to the excess capacity, which this time provided both by Saudi Arabia and shale gas/oil in the USA (Gauthier, 2015).

2.4 What scares an oil company away?

Investments in the oil and gas sector are costly; they require a commitment of several decades and large upfront investments. It is rational behavior for an oil company to think twice before making investments in unstable environments, afflicted with wars or rent-seeking governments prone to nationalization (Jojarth, 2008). The investment framework for exploration and drilling sets the terms for the type of contract and license, fiscal frameworks, transferability of funds and risk management and several other factors that influence how attractive a country is for an IOC (Toft & Duero, 2011).

Several factors are included in the risk assessment of oil companies when considering new investment decisions. The most important one is probably the risk of expropriation. Expropriation or ownership risk means that the government could take full or partial control over the production process. The most severe form is full nationalization of the extraction process where the national oil company seizes equipment and production without remunerating the oil company for sunk costs. Venezuela in 2007 is an example of this type of nationalization (Reuters, 2011). Other kinds of expropriation can be license cancellations or restrictions, and price controls (Alon, et al., 2006). Expropriation is more likely to happen in the oil industry, since resource wealth in most cases judicially belongs to the country, a new government would feel more entitled to take control over natural resource production compared to other industries. Moreover, the substantial upfront investments give the companies an incentive to continue its operation as long as it generates a positive cash flow, even though the expected profit is reduced (Alon, et al., 2006) and (Manzano & Monaldi, 2008). It has been shown that expropriation reduces exploration activity due to reduced foreign investment and reduced efficiency in the national oil company (NOC) (Toft & Duero, 2011).

To reduce the risk of expropriation, the government must prove itself stable, reliable and make a credible commitment to uphold property rights (Collier & Venables, 2011). There is a strong correlation between institutional quality and reduced ownership risk. Autocratic regimes have a reduced cost of expropriation since the leader will not be held accountable for its actions and are thus more prone to expropriation risk (Stroebel & van Benthem, 2013). Parliamentary democracy is the most secure system to invest in, due to its system of checks and balances (Bohn & Deacon, 2000).

Civil war and terrorism are other important reasons why companies avoid certain countries. Countries such as South Sudan and Iraq are places where it is almost impossible to operate due to the conflict level. Also in Nigeria, terrorism and criminality have been of increasing concern for oil companies. Shell, for instance, estimates it will lose approximately 10% of yearly oil output to thieves in the Niger Delta (Alon, et al., 2006). The practice of dealing with this type of risk differs from one IOC to another. Statoil is very cautious about investing in conflict areas and not to compromise the security of their employees (Klausen, et al., 2014) while BP has invested largely in security measures such as private guards to ensure the safety of personnel and equipment at the operation site (Alon, et al., 2006).

Other economic and political factors that will influence oil investments, are regulations on transferability of funds, domestic oil price and corruption. The regulations on transferability of funds put restrictions on the transfer of income out of the host country. For international oil companies these types of protocols impose difficulties since they often need available capital for investments worldwide. The company is not subject to less taxation by moving out profits, but it will usually increase currency risk if the company has to keep its profits in local currency in the national banks (Klausen, et al., 2014). A domestic oil price scheme or requirements for supplying part of the production to local markets are generally considered as another form of taxation. Especially when the oil price rises, oil exporting countries tend to set the national price much lower, which creates a high opportunity cost for the companies (Manzano & Monaldi, 2008).

Corruption might be a surprising risk to add to the list. Paul Collier (2011) paints a picture of oil companies as corporations without ethics, ready to bribe politicians to strike the best deals. While the issue of corruption is present and serious in oil producing states, the cost of corruption can be high for IOCs. Companies such as Statoil are subject to Norwegian rules and regulations, accusations of corruption might decrease both its stock and brand value (Klausen, et al., 2014). In corrupt regimes, the investor is dependent of remaining in favor with the ruling group and for the group to stay in power, which can be costlier than stable fiscal systems (Bohn & Deacon, 2000).

2.5 What attracts an oil company?

Naturally, the most important reason for an oil company to drill anywhere would be geological factors and the prospects of finding oil. There are, however, some measures countries who want to escalate exploration can implement to increase their attractiveness.

Government take means all types of income collected by the state, which implies a cost for the producing company. This includes all sorts of taxes, splits of profit and restrictions on oil price and exports. In general, we separate between two types of fiscal systems; concessionary and contractual regimes. Concessionary systems are the norm in most OECD countries, they give the oil company control over the extraction process for a given time period and within a certain geographical area. The production is still subject to tax and royalties, but all risk, both upside and downside, lies with the company. The other type of regime is contractual regimes. These differ between production sharing contracts and service contracts. In production sharing regimes, the state holds the right to the oil, while the contractor receives a share of the revenues called profit oil, the contractor bears the exploration risk. In service contracts, the contractor receives a cash fee for extraction of the oil and gas, while the government takes ownership of all production (Johnston, 2007).

One important point is that there is also an upside to risk, especially in oil exploration. This goes especially for oil price risk, where profits have been vast in the last decade. Service contracts and some production sharing contracts do not allow the companies to capture this upside, but still expect them to take on all risk in exploration. This is one reason why many oil companies will avoid service contracts (Klausen, et al., 2014). Some contractual regimes also have large resource taxes that allows the state to capture an upside of windfall revenue, for instance Norway with a special petroleum tax of 51% in addition to corporate taxes of 27%. The difference between the Norwegian system and production sharing contracts, is that in Norway, the government takes responsibility for potential losses and exploration risk also, through generous depreciation schemes and refunding of negative taxes (Deloitte, 2014).

Although the tax rate in Norway is high, the taxation only applies when a field is starting to make profit, which means it will not affect any drilling decision. A well-designed fiscal system, like the Norwegian one, should avoid distortionary taxes that will reduce drilling activity. While royalties have been found to reduce both the development and lifespan of oil fields, income taxes barely distort drilling (Stroebel & van Benthem, 2013) and (Manzano & Monaldi, 2008).

The problem that arises with non-distortionary taxes, is that it takes several years of drilling activity before the government starts receiving tax income. A common problem in countries with weak institutions is that the leaders tend to have myopic preferences, which makes them favor early income streams since they might not be able to profit themselves from tax income generated a decade later. This results in actions such as high signature bonuses and corporate taxes that are not linked to production (Ross, 1999). Companies, however, wish to defer tax payments as far as possible to reduce the net present value of the investment.

In the early exploration phase there are several problems related to asymmetric information that will reduce efficient exploration. Oil corporations are reluctant to pay for geological surveys and exploration only to end up with the government expropriating the assets when oil discoveries are made. Companies prefer to buy the rights to a discovery and wait for other companies to prospect the area (Collier & Venables, 2011). A way to reduce the agency problem is for the government itself to pay for early exploration and geological surveys. Making this information public will allow more companies to participate in the auctions and will probably increase the price companies are willing to pay (Collier & Venables, 2011).

2.6 How do oil companies estimate country risk?

Until recently, the most common way to handle country risk has been to increase the required rate of return for operations in high-risk countries (Aven & Flørenæs, 2004). This approach creates some problems however, as it treats country risk as a symmetric risk captured through the beta-value in the CAPM model. By adding country risk to the discount rate, it increases the required rate of return exponentially over the investment period, which indirectly means that country risk increases exponentially over time (Klausen, et al., 2014). Country risk is treated as a cost of capital, while it is actually an asymmetric risk where serious incidents that affect production cannot be expected to occur annually. Lasserre (1985) finds that in cases where the cost of capital is higher than the value of the resource extracted, an increase in the discount rate can reduce the extraction period. To deal with this issue, Statoil has started adding country risk directly through the cash flow calculation as a part of the investment costs, instead of increasing the required rate of return (Aven & Flørenæs, 2004). It is plausible that oil companies have turned down investment prospects in developing countries due to a negative NPV based on a discount rate that has been estimated incorrectly.

2.7 Is it better to extract in democracies?

The most important factor for any oil company is a predictable and stable framework that will not change during the exploration and production process. An IOC will prefer countries with high tax rates, but institutional frameworks which protect the investment and ensure stability, over unsecure investments where the government take can change overnight (Radon, 2007). All the factors mentioned in the previous sections show a close relationship with institutional quality. They can however, also correlate with income. Risk-mitigating procedures such as geological surveys and progressive tax systems require public capital and administrative capacities that many poor countries lack (Manzano & Monaldi, 2008). Before World War II, the United States spent 1 million pounds annually on geological surveys, which can explain the high drilling activity in the same period (David & Wright, 1997). Most developing countries will never be able afford the same means of investment.

Another question that arises, is whether democracy itself has a risk mitigating impact on oil and gas investment.

Most of the papers previously cited have used either expropriation risk or constraints on government as measured by Polity IV or Freedom House as estimators of institutional quality. The risk of using expropriation or property rights as institutional indicators is that they do not differ between dictatorships that respect ownership rights and democracies where the leaders are obliged to follow the laws. Conversely, the Polity IV index that measures the constraints imposed on the chief executive, tends to vary with electoral outcomes and might not be a correct indicator of permanent institutions (Glaeser, et al., 2004).

As the resource curse literature discusses, institutional quality tends to deteriorate with resource dependence. Another aspect, presented by Paul Collier (2007), is that in countries with high resource abundance, autocracies tend to outperform democracies in terms of economic growth. He attributes this to the intense electoral competition that will increase public spending and reduce investment. This suggests that myopic preferences are not a phenomenon only occurring in autocracies as discussed earlier, but also in democracies, especially when the means spent originates from resource wealth rather than the voters tax money (Collier, 2007). In Ghana, the government has increased public spending to meet the high expectations from the public following its recent oil discoveries (Looney, 2014). There is a possibility for expropriation in democracies too, if the newly elected government decides that the investment terms provided by the previous government were too generous and decides to renegotiate the contracts. This suggests that democracy, as measured by the Polity IV project, might not be of the most relevant risk-mitigating factor for international oil companies.

2.8 Why is there a resource curse if oil companies prefer countries with high institutional quality?

One conclusion several researchers come to, is that the occurrence of the resource curse depends on the initial state of the institution before oil drilling occurs. Institutions described as extractive or “grabber-friendly” promote rent-seeking activities that does not contribute to economic growth

(Mehlum, et al., 2006). Extractive institutions do not have the same system of checks and balances in a well-functioning democracy that deters the executive powers from corruption and rent-seeking activities (Acemoglu & Robinson, 2012). Mehlum, et al, (2006) find that the resource curse is only present in countries where the institutions are defined as grabber-friendly, while (Haber & Menaldo, 2011) claim to find no evidence of increased authoritarianism in resource dependent countries. Collier and Hoeffler (2009) suggests that newer democracies in developing countries struggle to receive the benefits from resource revenues, since the candidates are turning to patronage politics, spending the oil revenue to get themselves elected. They argue that democracy alone, without checks and balances that hinder corruption and lavish spending of resource revenues, will not reduce the impact of the resource curse (Collier & Hoeffler, 2009).

It is thus possible that even though there is a correlation between natural resource dependence and degrading institutional quality, this does not necessarily reflect the drilling activity in a particular country. The exploration rate could still be correlated with institutional quality, while the discovery rate is more exogenous to these factors.

While reserves might be more abundant in countries with poor institutions, the assumed costs of operating in a high-risk country, could outweigh the benefits of finding larger reserves. As mentioned previously, oil companies have not used sufficient techniques estimating the actual cost of country risk. It is thus hard to find a specific level of country risk where the costs would outweigh the benefits of oil exploration. It is more likely that this factor is about perceived costs which might distort some investment. The lower the oil price, and the more readily available reserves in countries with good institutions, the lower likelihood of oil exploration in high-risk countries. The exploration activity and the risk an oil company is willing to take will usually correlate with the oil price and available reserves (Gauthier, 2015).

3. Presentation of the analysis and the dataset

This chapter presents the model set-up and the variables used in the analysis. The chapter touches upon some of the methodology, which will be discussed in more detail in chapter 4.

3.1 Previous research

The methodology used in the regressions mainly follows the work done by Bohn and Deacon (2000) and Cust and Harding (2014). Bohn and Deacon perform a cross-country analysis using panel data from 27 countries in the period from 1957-1988. Their dependent variable is a logarithmic version of the number of wells drilled per year, while the policy variable is a self-made ownership risk index (Bohn & Deacon, 2000). Cust and Harding use micro data from Wood Mackenzie on the location of oil wells on the border between two nations to determine whether drilling activity has been higher in the country with better institutional qualities assigned by Freedom House and Polity IV. They also investigate whether the discovery rate has been higher in countries with better institutions (Cust & Harding, 2014).

This analysis distinguishes itself from the previous work by looking at a larger dataset coming from both Global Insight (IHS) and Wood Mackenzie provided by Statoil ASA. It covers 86 countries and a time period of 40 years, which makes it more extensive than the study by Bohn and Deacon. The methodology differs too: while Bohn and Deacon use panel data with country fixed effects, I perform a cross-sectional analysis using an instrument for institutions in order to avoid possible endogeneity issues coming from the institutional variable. Cust and Harding address the identification problem by studying drilling close to country borders and estimate the impact of bilateral differences in institutional quality, while I run an overall regression looking at aggregate drilling data from each country while correcting for geological factors by adding control variables to the regressions.

3.2 Model set-up and equations

There are strong indications that institutional quality is affected by natural resource extraction. If I were to perform a panel data analysis, estimating whether the extraction rate changed with alterations in the institutional level, it would be impossible to determine which way the causality went.

To avoid this issue, I look at a cross-section of observations, trying to determine how the observations differ between countries, instead of looking at the development within a country over time. This is a common method when institutional quality is an independent variable, for example when estimating the correlation between institutions and wealth (Acemoglu, et al., 2005).

A cross-sectional analysis implicates that each variable only occurs as one observation for each country. This means that variables changing over time, such as the oil price, cannot be used in the regressions. In this analysis, three different methods have been used to present the data in a cross-sectional form. First, the dependent variables measuring the number of NFW wells drilled or the number of drilling days are presented in aggregated form, using the total sum of all observations in the time period from 1970 to 2010. The success rate and institutional variable, on the other hand, is presented as the average of observations in the time period between 1970 and 2010. Finally, some of the control variables are stationary variables that are not expected to change over time (at least not dramatically), such as landlockedness, ethnic divisions or surface area. Another advantage of using static variables is that they are exogenous to oil exploration which makes it easier to prove the causality of the dependent variable (Woolridge, 2009).

The hypothesis of the analysis is that oil exploration does not depend on geological factors alone, but also on the institutional quality of a country. In order to be able to prove a causal relationship between institutional quality and oil exploration, there is a need for correcting for the geological factors, which naturally have an impact on oil drilling. For this reason, several variables are added to the regression in order to reduce the unsystematic variance originating from geographical and geological effects in each country.

All regression models in this analysis use the same explanatory variables, while the dependent variable differs. The following equations (1), (2) and (3) represent the equations that are used in the main share of the analysis.

Regression equations:

$$NFW_i = \beta_0 + \beta_1 I_i + \beta_2 A_i + \beta_3 D_i + \beta_4 Y_i + \beta_5 E_i + \beta_6 O_i + \beta_7 R_i + \beta_8 L_i + z_i \beta_9 G_i + u_i \quad (1)$$

$$Days_i = \beta_0 + \beta_1 I_i + \beta_2 A_i + \beta_3 D_i + \beta_4 Y_i + \beta_5 E_i + \beta_6 O_i + \beta_7 R_i + \beta_8 L_i + z_i \beta_9 G_i + u_i \quad (2)$$

$$Success_i = \beta_0 + \beta_1 I_i + \beta_2 A_i + \beta_3 D_i + \beta_4 Y_i + \beta_5 E_i + \beta_6 O_i + \beta_7 R_i + \beta_8 L_i + z_i \beta_9 G_i + u_i \quad (3)$$

NFW wells drilled from 1970 to 2010, the number of days drilled from 1970 to 2010 and the success rate from NFW wells drilled between 1970 and 2010 are the dependent variables. I represents the estimate for the institutional variable measured by the Polity IV index. In the ordinary OLS regressions, the institutional variable is the averaged index between 1970 and 2010. In some of the regressions, the Polity IV score between 1960 and 1969 is used as an instrument to correct for endogeneity issues. The other variables in the regression are control variables in the following order; Area, Distance to Equator, Number of years with drilling activity, Ethnolinguistic fractionalization, OPEC member, Regional control variables, Landlockedness and a vector measuring the geological conditions. The following section presents the variables used in the analysis, where they originate from and how they have been treated.

3.3 Description of dataset and variables

The data used in the regression is a mixture of data from publicly available sources provided by institutions such as the World Bank and economists at Harvard and UCLA, and data available through subscription at Wood Mackenzie and Global Insight (IHS).

The control variables that are added are largely motivated by the regressions run by both Bohn and Deacon as well as Cust and Harding, while the first paper includes variables such as OPEC membership and land area, the former adds variables like ethnolinguistic fractionalization, landlocked status and the year drilling activity started. Some variables such as the quality of the oil measured by the API gravity and depth of the well has not been available and is thus excluded.

New field wildcat

The first dependent variable is estimating the numbers of new field wildcats drilled in the period between 1970 and 2010. A new field wildcat (NFW) is a test well located far from producing areas and on surfaces where there has not been any previous oil and gas production (Murray, 1990). The dataset has been extracted from IHS and its databank for oil and gas: "Petroleum Economics and Policy Solutions". IHS reports the number of new field wildcat discoveries together with the success rate, which allowed me to estimate the total number of NFW wells drilled in each country. The dependent variable contains the sum of all NFW wells drilled in each country from 1970 to 2010. The dataset also separates between onshore and offshore wells that permits for testing whether the impact of institutions is stronger for onshore drilling. The variable is logarithmically transformed, the reason behind this choice is explained in the chapter 4.

The number of NFW wells drilled is a good indicator of oil companies' interest in a country or area, since it depends to a lesser degree on geological history and previous production, compared to related indicators, such as production rates. For instance, it can be expected that countries that offer favorable conditions, such as subsidized geological surveys and well-designed fiscal regimes, will attract more companies eager to drill test wells in that area, compared to countries more reluctant to foreign direct investment.

Drilldays

Drilldays measures the number of days spent on drilling in each country in the period from 1970 to 2010. The data has been retrieved from Wood Mackenzie and consist of wells drilled between 1970 and 2010, the list comprises 95 068 observations. The number of days spent on each well has been summarized for each country, and the result is the variable drilldays. A logarithmic transformation is used in the regressions.

Success rate (IHS)

The success rate measures the discovery rate of all NFW wells drilled between 1970 and 2010. It derives from the same IHS dataset as NFW wells, but instead of measuring drilling activity, it measures the outcome of the drilling. The variable is estimated as the average discovery rate between 1970 and 2010 for each country and is expressed as a number between 0 and 1.

Number of wells (Wood Mackenzie) and success rate (Wood Mackenzie)

The final part of the analysis consists of a chapter where I test the robustness of the results by replacing the dependent variables coming from the IHS dataset in the main regression with similar variables originating from the dataset provided by Wood Mackenzie. The variable replacing NFW wells drilled, is the sum of the number of wells for each country, enlisted in the dataset, in the period from 1970 to 2010. The dataset includes both exploratory and appraisal wells and differs somewhat from the NFW variable in the IHS dataset. The success rate is the number of successful wells divided by the total number of wells drilled. The variables are estimated the same way as those they are replacing, i.e. as a logarithm for wells and percentage between 0 and 1 for success rate.

Polity IV

The variable used to measure institutional quality is from the Polity IV project, which has estimated institutional quality back to the 1800s. The scale of the indicator runs from -10 (hereditary monarchy) to +10 (consolidated democracy). The Polity IV index comprises mostly of ratings related directly to the definition of democracies such as; fair elections, suppression of opponents and constraints on the power of the executive leader (M. Marshall, 2014). The Polity IV index is broadly accepted as a measure of institutional quality (Cust & Harding, 2014). It is also the only index that covers the period from 1960-69, which is essential to reduce endogeneity issues.

The Polity IV variable has been normalized to the form 0 (full autocracy) to 1 (full democracy). I have used the polity2 variable which has been adjusted for time series analysis (M. Marshall, 2014). In the cases where two countries have been separated and then unified (Germany, Vietnam and Yemen), has the country part with the Polity IV value closest to the current regime been chosen (West Germany, North Vietnam and North Yemen).

The Polity IV score variable is represented in two versions in the analysis. The one used in the main regressions is the average of Polity IV scores between 1970 and 2010. The other version is an average from 1960-69 used as an instrument to avoid the possible endogeneity problem regarding the resource curse. The Polity IV index has been criticized for varying too much with electoral outcomes instead of presenting institutional quality (Glaeser, et al., 2004). Averaging the index over a longer time period reduces the impact of shorter change in governance and displays the long-term trend of institutional quality.

Freedom House

The variable for political rights from Freedom House functions as a substitute for the Polity IV index to test the robustness of the results. The Freedom House index measures the freedom in a country, but most of the underlying indicators are similar to those in the Polity IV index, i.e. free elections, strong opposition, free political choice etc. (Cust & Harding, 2014). It is thus widely used as an indicator on institutional quality. The Freedom House index first appeared out in 1973 so it cannot be used as an instrument in the 2SLS regressions. In the robustness tests I replace the variable estimating Polity IV 1970-2010 with an averaged Freedom House score from 1973-2010. The variable has been normalized to go from 0 to 1 in the same manner as Polity IV.

3.4 Other control variables

The regression analysis treats drilling decisions and discovered reserves as endogenous variables that can be affected by other factors such as institutional quality. This assumes, however, that geological factors such as where oil and gas can be found is also endogenous. The reason why most of the following control variables have been added to the regression is an attempt to control for the geological conditions and other factors that could have an impact on drilling activity to be able to better estimating the impact coming from institutional quality.

Area

A variable for surface area has been added to the regression to control for the size of the country. Normally, one would expect a correlation between the geological surface available for drilling and number of wells drilled. The data derives from the World Bank and is referring to country area in 2005. The variable has been transformed to a logarithm

Distance to equator

Since the regression is a cross-sectional analysis, it does not control for country specific effects like a time series analysis would. The distance to equator has been added together with regional variables to control for some of the geological aspects, which applies to different regions around the world.

Regional variables

Variables for seven different regions have been added to the analysis in an attempt to correct for regional differences. The regions included are: South Asia, East Asia and the Pacific, Sub-Saharan Africa, Middle East and North Africa, Latin America and the Caribbean, Western Europe and North America. The reference group for the regions is Central Europe and Central Asia which have been excluded to function as a baseline for the other regions.

Start year

Oil companies tend to drill where it has been drilled before (David & Wright, 1997). This makes sense since reserves have already been proven and the geological risk has been reduced. A variable reporting the first year of drilling activity has been added to correct for this factor. The years have been extracted from the Wood Mackenzie dataset of wells drilled. It is likely that full drilling activity started later than when the first well was drilled, but at least it indicates when geological interest commenced.

Ethnolinguistic fractionalization

The ethnolinguistic fractionalization measures the probability that two randomly selected individuals from a given country will not be from the same ethnolinguistic group. The variable was estimated in 1985 (Treisman, 2007). It is often used as an exogenous indicator of conflict and civil war (Easterly & Levine, 1997), factors that are associated with the resource curse. The addition of the variable will correct for these factors that might scare off oil companies but which are not directly linked to institutional quality.

Landlocked

Landlockedness is often a feature that reduces FDI and economic growth. The limited access to sea routes increases both costs and the complications for oil and gas transport (Faye et al, 2004). The variable is exogenous to both institutions and test drilling.

OPEC

The largest oil exporting countries might differ from smaller oil exporters. The bargaining power for a country with vast reserves is much higher than in countries with smaller reserves. Another factor that might affect the analysis, is the difference between national oil companies and international oil companies. Since the datasets from Wood Mackenzie and IHS do not differ between production by national and international oil companies, it is not possible to identify these differences. Some of the most powerful national oil companies originate in OPEC countries. Adding a dummy variable for OPEC members, given their membership status in 2014, will correct for some of the effects mentioned above.

Geological indicator

The geological indicator variable is a vector consisting of geographical indicators extracted from a dataset provided by John Gallon and Jeffrey Sachs from Harvard University. The variables measure the percentage of land area that is polar and non-desert, percentage of area covered by boreal forest, areas covered by tropical or subtropical forest and areas in dry and wet temperate zones. In addition, two variables measuring soil suitability for agriculture have been added. The vector is estimated based on regressions run for each dependent variable, where the parameters for the geological variables mentioned, has been used to estimate the relative importance of each variable included in the vector. This means that the coefficients of the vector differ slightly when changing the dependent variable, but it has not affected the outcome of any of the regressions. A table showing the different vectors estimated is added to the appendix.

A drawback of the data collected is that it measures geographical rather than geological factors. The type of forest covering a country does not necessarily reflect its natural resource wealth.

Geological data on country level, however, have turned out to be very hard to find. Neither IHS and Wood Mackenzie, nor public datasets, have provided any better indicators.

3.5 Graphical description of the dataset

The following section presents a graphical display of the dependent variables from the dataset in order to get a better understanding of the drilling activity, separated into regions.

Figure 3 and 4, show the number of NFW wells drilled, separated by region and by the decade the drilling occurred. In the Wood Mackenzie dataset, USA drilled far more wells than any other country, so North America has been separated into its own graph. Figure number 5 displays the total number of wells drilled in each decade.

An interesting observation from the graphs is that East Asia and the Pacific have drilled most wells (besides North America), largely due to China, Australia and Indonesia. The Middle East and North Africa has only drilled approximately 6000 wells, which is surprising considering it is an oil rich region. This creates a suspicion that the dataset might be subject to some level of measurement bias, where IOCs might be better at reporting practices than the NOCs operating in the Middle East, but it could also be due to the limitations OPEC countries put on themselves during the 1960s and 70s (Bohn & Deacon, 2000).

Figure 3: NWF wells drilled per region and decade

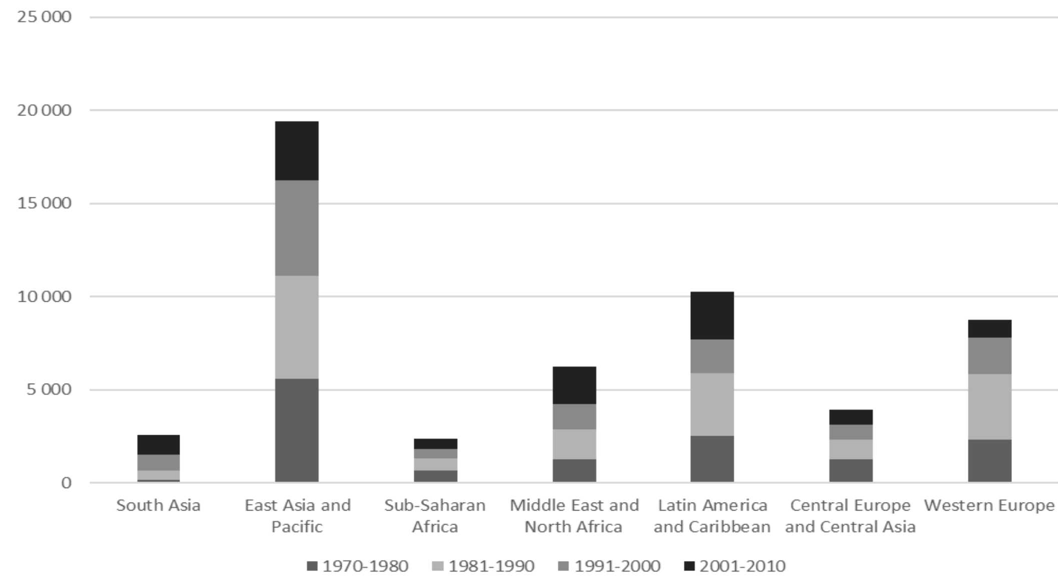
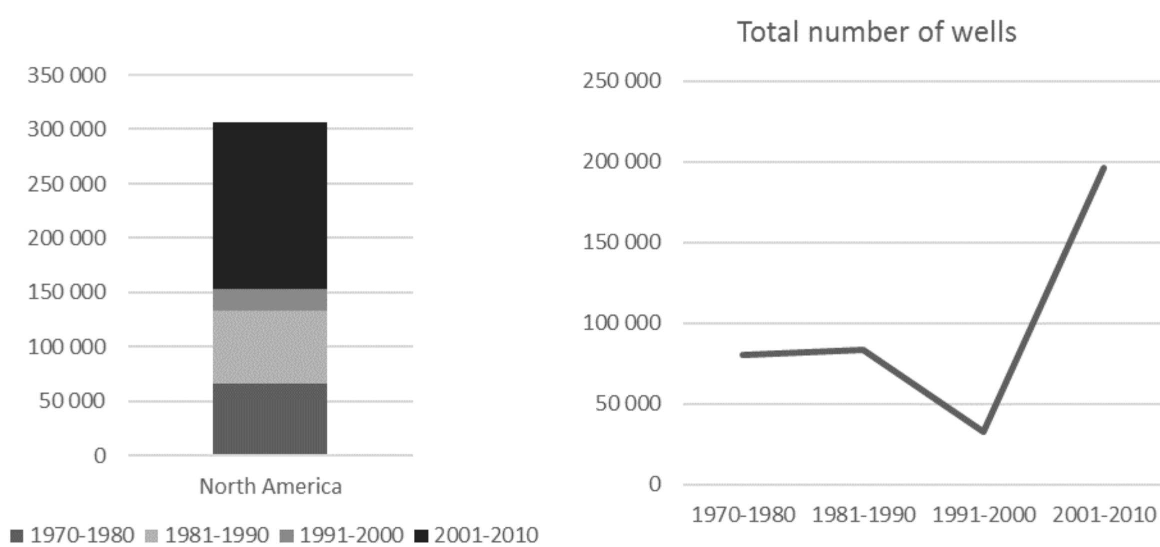


Figure 4: NFW wells drilled in North America (left) and NFW wells drilled from 1970-2010 (right)



When comparing figure 5 with the graph displaying the development in the oil price in the first chapter, there is a clear connection between the rise in the oil price and the number of wells drilled. The drilling intensity dropped when the oil price was at 20 USD per barrel in the 1990s, but increased when the capacity was reduced in 2000. This means that the wells drilled between 2000 and 2010 will be overrepresented in the dependent variable measuring NFW wells from 1970 to 2010.

Figure 5: Number of drilldays

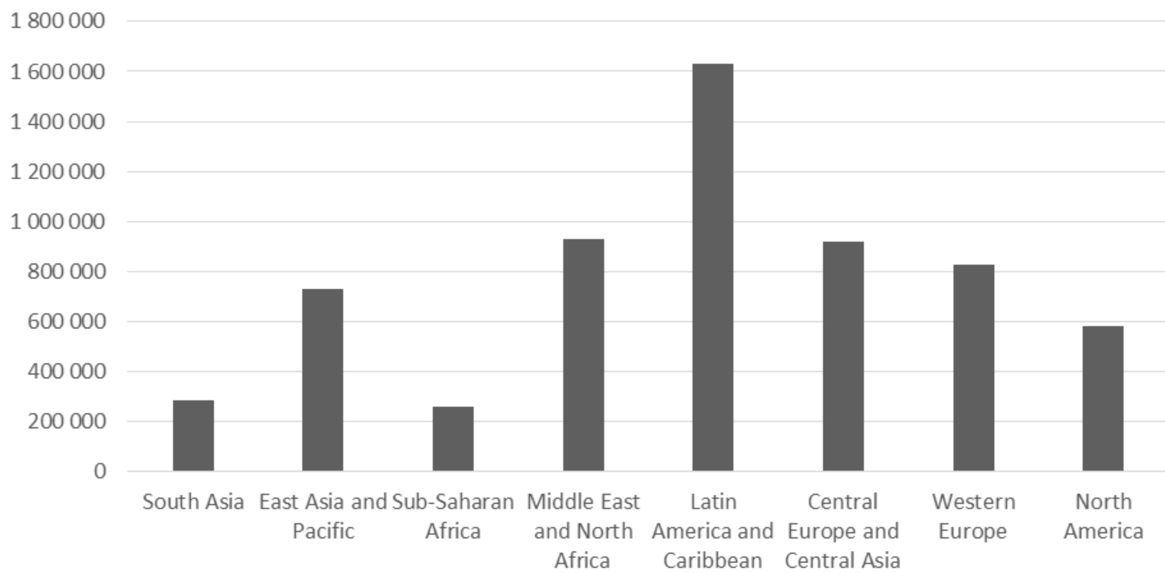
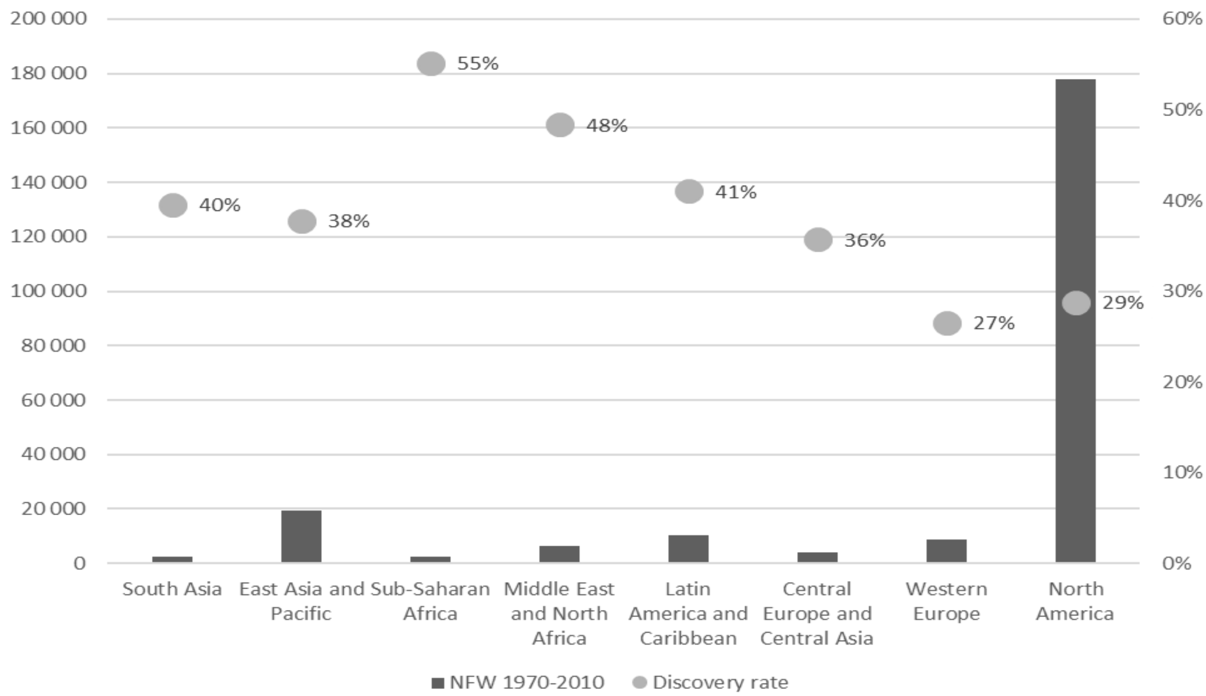


Figure 6 shows the dependent variable measuring the number of drilldays separated by region. In this case, the region with the most drilldays is Latin America and the Caribbean, mostly due to the activity in Argentina, Venezuela and Colombia. The variable measuring drilldays comes from the Wood Mackenzie dataset and there are smaller differences between the regions compared to the IHS case with NFW wells.

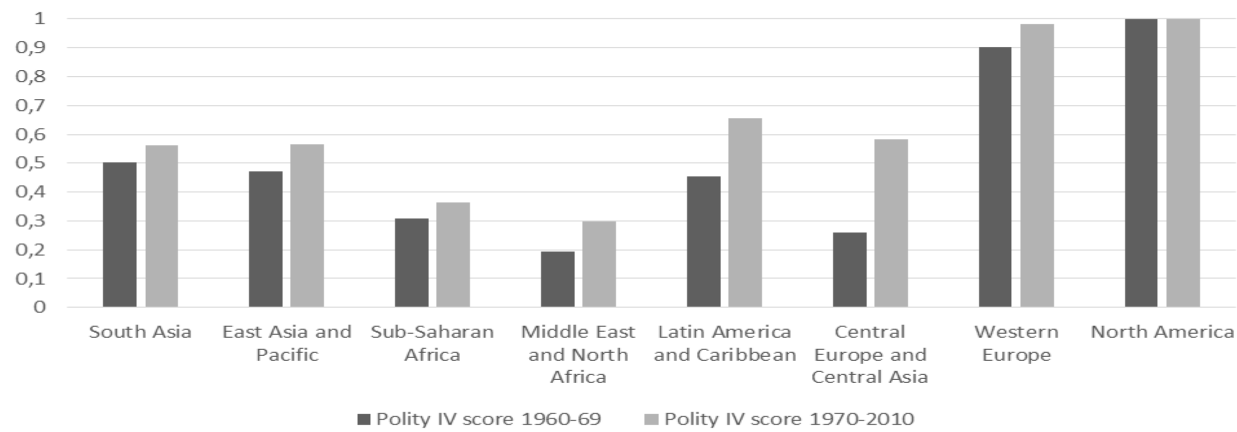
Figure 7 shows the overall number of NFW wells drilled per region on the left axis and the corresponding discovery rates on the right axis. The discovery rate ranges from 27% in Western Europe to 55% in Sub-Saharan Africa. This is much higher than the discovery rate suggested by Bret-Rouzaut & Favennec (2007) which was between 10 and 30%. A possible explanation could be that the dataset comprises wells back to the 1970s and that the discovery rate has dropped in recent years as more reserves have been explored. Some regions, like Sub-Saharan Africa, have an average discovery rate at 55 %, but with a very small number of NFW wells drilled. This indicates that those wells that have been drilled have been in areas where the likelihood of finding oil is very high, normally what we would refer to as "low-hanging fruit".

Figure 6: NFW wells drilled between 1970-2010 and the discovery rates



The final graph presents the development in institutional quality measured by the normalized version of the Polity IV index between 1960-69 and 1970-2010. It shows that institutional quality has improved in all regions but North America, where the region already had the highest level possible. The region with the lowest average score is the Middle East and North Africa, with Sub-Saharan Africa following closely.

Figure 7:Development in the Polity IV score per region



4. Methodology and model diagnostics

The following section goes through the methodology of the analysis and comments on whether the dataset meet the requirements and the measures that have been made to fulfill them.

4.1 Regression with ordinary least squares cross sectional analysis (OLS)

The first regressions in the analysis follow an ordinary least squares regression (OLS). The OLS model is a method for estimating a linear regression model, where the estimates for the parameters are acquired through a minimization of the sum of squared residuals (Woolridge, 2009). The residuals are the difference between the observations and the values of the dependent variable estimated by the OLS model.

An OLS regression has to fulfill certain requirements in order to be the best linear unbiased estimator (BLUE). These are:

Assumption 1
The regression has to be linear in parameters.

This entails that each one-unit increase in one of the explanatory variables, everything else held constant, will have the same impact on y, regardless of the initial value of the independent variable (Woolridge, 2009).

In the regressions using NFW wells and drilldays, I am using a log-linear approach on the following format:

$$Log(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_x X_x + u_i \quad (4)$$

The regressions with success rate is an ordinary level-level regression.

The parameters capture the systematic variance of the regression, while the error term captures the unsystematic variance. The purpose of the regression is to find whether some of the variance observed in the dependent variable Y, can be retrieved in the explanatory variables. If the systematic variance is large enough, and the unsystematic quite small, the independent variable X would be a significant explanatory factor in explaining the variance observed in the dependent variable Y. It is common to require 5% significance in statistical analysis, i.e., a 5% chance for the estimated explanatory variable to fall on the outside of an estimated confidence interval.

Assumption 2

The data must be a random sample of observations.

The dataset used in the regressions are showing all the countries where data were available. There is a possibility that countries with better institutions have better reporting practices in the oil industry than in countries with little or no bureaucracy. The datasets provided by IHS and Wood Mackenzie are known for reliability and are perceived as trustworthy by the oil and gas sector. In one case however, I encountered a possible problem of sample selection. In the dataset of wells provided by Wood Mackenzie, the data comprise 95 068 observations from 1970 to 2010. Only 399 observations of the dependent variable drilldays were missing. The variable that estimated the amount of dollars spent per well had 56 886 observations reported as missing. When looking closer at the dataset, it seemed that it was mostly developed countries, which often have strong institutional quality, that had reported well costs, while those with weaker institutions were missing. To avoid the possibility of a sample selection bias, costs per well was excluded as a dependent variable from the regressions.

Assumption 3

There is some variation in the samples in the explanatory variables.

This is a very weak assumption and holds very well in the datasets used in this paper.

Assumption 4

The error term has a zero conditional mean, i.e., $E(u|X) = 0$. For all values of X will the expected value of the error term be zero. This assumption applies to all the independent variables. The conditions for assumption 4 to hold are discussed in the following paragraphs.

Four reasons to endogeneity and a biased OLS estimator

Omitted variables and simultaneity

A reason to add more control variables is to reduce the variance in the error term and rather correct for it in the regression itself. By including other factors that simultaneously affect the dependent variable, the explained variance increases and makes the model more precise. It reduces the problems with spurious correlation, i.e., when a relationship between two variables are found, but there might be other unobserved factors that also happen to be correlated with one of the explanatory variables (Woolridge, 2009).

An important quality of all the explanatory variables is that they have to be exogenous. If one of the independent variables happens to be determined simultaneously with the dependent variable, it is hard to prove which variable decides the outcome of the other. This is the issue with using an institutional indicator from today, since research suggest that institutional quality deteriorates with oil drilling. The rest of the control variables chosen in the regressions are expected to be exogenous.

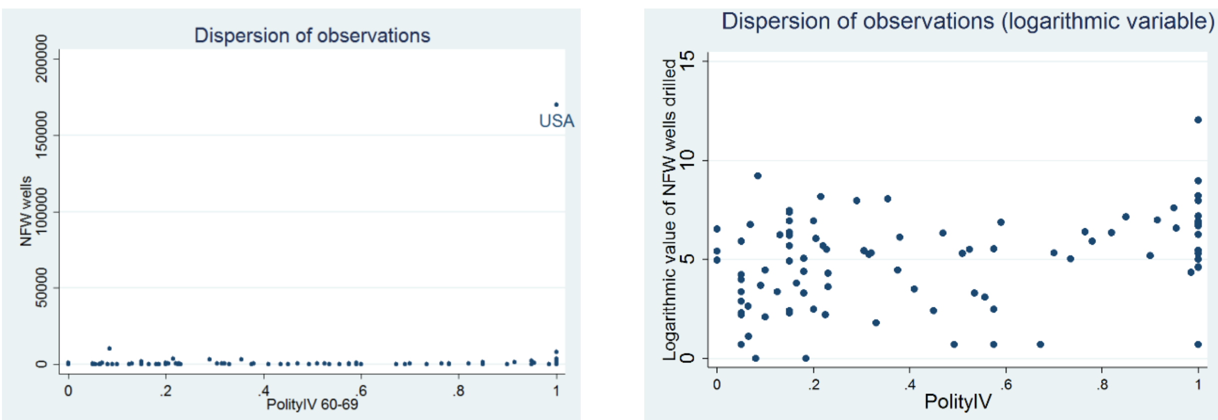
Institutional quality is a variable that has demonstrated to have a large impact on several issues in a country. Sometimes researchers experience problems deciding what is the outcome of what, for instance whether economic wealth is a result of good institutions, or the opposite. There is however a certainty that good institutions and prosperity are closely linked (Acemoglu, et al., 2005). This poses an issue for this model. As mentioned in the theory chapter, institutional quality might just be a proxy variable for wealth when modeling drilling decisions. Yet, it is impossible to add a variable for country wealth to the regressions, since many countries are wealthy because of their natural resources, and this could cause an endogeneity problem. In the robustness chapter, regressions are run on samples of low-income and high-income countries separately, to see if the impact of institutions are strong enough to hold when rich countries are left out of the sample.

Adding irrelevant variable to the regression, will not affect the biasedness of the OLS model, but could lead to multicollinearity. Multicollinearity occurs when there is a high degree of correlation between some of the independent variables, which would lead to a less efficient estimator for the individual parameters, although the dependent variable would still be unbiased (Woolridge, 2009). The variables in the equation do not show high correlation with each other, although the variable for distance to equator has been used as an instrument for institutional quality in some papers (Dollar & Kray, 2003). The correlation coefficient is 0.31, so it seems safe to assume that the model does not experience problems with multicollinearity.

Functional form misspecification

The OLS estimator could be biased if the model is misspecified. This means that the real model has a different functional form than the one presented in the regression. Failing to use the logarithmic version of a variable when this is more appropriate, is an example of a misspecification error (Woolridge, 2009).

Figure 8: Dispersion of observations before and after logarithmic transformation of NFW wells



The figure to the left shows a dispersion of the observations from the dataset of NFW wells drilled and the Polity IV score between 1960 and 69. One outlier stands out, there has been drilled around 150 000 new field wildcat wells in the US, compared to the other countries. This argues for either leaving USA out of the dataset, to treat it with a dummy variable or to apply a logarithmic transformation of the variable. Removing the US would reduce the argument for increased drilling in countries with good institutions. When testing the two non-nested models, level-level and log-level, with a Davidson Mackinnon test, neither of the models were rejected. In that case, the R^2 can be used as an indicator of which model to choose (Balsvik, 2013). The log-level model gave a higher R^2 than the level-level model. Even when leaving the US out, the data observations differ from one to two wells to more than 5000. A level-level model would treat any increase in institutions with the same number of wells increased. This might not be reasonable, as it might be differences between the first 100 wells and the 3000^{th} well. A log-level model on the other hand would rather measure the constant percentage change that comes with institutional improvement (Woolridge, 2009). The variable measuring drilldays has been transformed to a logarithm for the same reason. In the general regressions, USA and Canada are represented in a North America variable who captures much of the variance. A specific US dummy did not give any significant outcomes. The issue with outliers and how it was dealt with is described further in the robustness chapter.

Measurement errors

If there is variation between the reported data used in the regressions and the actual observations, the model will suffer from measurement errors. An example of a measurement error is when some developing countries chose to report a lower GDP per capita in order to receive development assistance (Jerven, 2013). If the measurement error were systematically related to one of the explanatory variables, the OLS estimates would be biased.

I have already commented on the consistency and reliability of the dataset and all data sources seems to be trustworthy. The data coming from oil companies could potentially suffer from some measurement bias, but it is hard to predict in which direction. The assumption for the model is thus that it does not suffer from measurement bias.

Assumption 5

Homoskedasticity

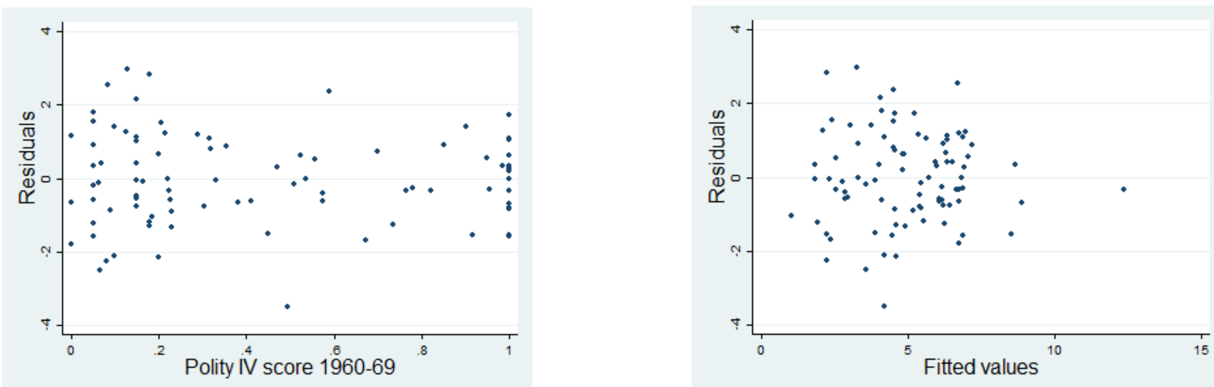
The error term has the same variance given any value of the explanatory variable.

$Var(u|x)=\sigma^2$

If the variance depends on the independent variables, the error term will show sign of heteroskedasticity. Heteroskedasticity does not affect the biasedness of the estimators predicted by the OLS model, but it will affect the variance used to prove statistical inference of the variables. This will break the Gauss-Markov assumptions and the OLS estimators will no longer be the best estimator even though they still are linear and unbiased (Woolridge, 2009).

Residual plots of the OLS regressions can help to investigate whether the assumption of homoskedasticity has been broken.

Figure 9: Scatterplots of residuals and fitted values from the regression with NFW wells



The two plots show the residual from the regression with new field wildcats plotted against the Polity IV variable and the fitted values. The first plot seems not to display any trends in the residuals, although the dispersion of the residuals might be slightly more concentrated on the lower and upper side of the polity score than in the middle. As for the fitted values, the residual plot seems less symmetric which leans towards that assumption 5 might be broken.

A more formal test to detect heteroskedasticity is the Breusch-Pagan test (Balsvik, 2013). When I run the test, the null hypothesis about homoskedasticity is not rejected and none of the variables proves significant in explaining the variance in the residuals. Just to be sure, I use robust standard errors in all the regressions.

4.2 2SLS

If the assumption about unbiasedness of the OLS estimator does not hold and one suspect one of the variables to be endogenous, a possible solution is to replace the endogenous variable with an instrument and use a two stage least squares estimator (2SLS) (Woolridge, 2009). In this paper, I suspect the more current estimate of institutional quality to be potentially endogenous due to its link with the resource curse. This is the reason why an estimate from 1960-69 before the drilling period started has been used as an instrument both in a reduced OLS regression and in a 2SLS regression.

The 2SLS model estimation occurs, as the name indicates, in two stages. In the first stage the assumed endogenous variable, in our case Polity IV 1970-2010, is regressed on the instrument, Polity IV 1960-69.

Estimate of the first-stage 2SLS analysis

$$I_{i(1970\text{-}2010)} = \beta_0 + \Omega_1 I_{i(1960\text{-}1969)} + \beta_2 A_i + \beta_3 D_i + \beta_4 Y_i + \beta_5 E_i + \beta_6 O_i + \beta_7 R_i + \beta_8 L_i + z_i \beta_9 G_i + u_i \quad (5)$$

I use the fitted values provided by equation (4) as an instrument for the institutional quality between 1970 to 2010. The predicted values of $I_{i(1970\text{-}2010)}$ from equation (4) is the part of the variable that can be explained by the instrument; $I_{i(1960\text{-}1969)}$. Since $I_{i(1960\text{-}1969)}$ is uncorrelated with the error term, so is the fitted values of $I_{i(1970\text{-}2010)}$. In the second stage, the predicted part of the variable is used in an ordinary OLS regression (Balsvik, 2013)

The instrument used in a 2SLS regression holds two important characteristics: first, it must be exogenous to the dependent variable, and the equation used in the OLS regression. If the covariance between the instrument and the error term is not zero, the endogeneity problem still holds. In the analysis, I assume that earlier institutional quality is exogenous, which means that it has not been deterred by drilling decisions.

Furthermore, the instrument has to be relevant in explaining the variance in the endogenous variable. In most cases, the institutional quality of a country is likely to be stable over time. This is confirmed when looking at the correlation between the two variables which is at 78.5%. As a rule of thumb, the t-value of the instrument should be above 3.2 (Balsvik, 2013). When Polity IV 1970-2010 is regressed on Polity IV 1960-1969 the t-value is at 14.0. The F-value of the instrument is reported in each table under the 2SLS column.

A valid instrument in a 2SLS regression holds two important characteristics, first it must be exogenous to the equation used in the OLS regression. If the covariance between the instrument and the error term is not zero, the exclusion restriction does not hold (Woolridge, 2009). Since the error term is unobservable, this cannot be tested. It has to be left to the econometrist to evaluate whether the exogeneity restriction holds. In this analysis, that implicates an assumption that earlier institutional quality is exogenous to the error term. Previous institutions only affect oil drilling through its impact on institutional quality later on and there are no other factors that correlates with the error term. It is hard to say whether this assumption holds, but at least the instrument removes the reverse causality associated with the resource curse, which is the main argument for including it.

Furthermore, the instrument has to be relevant in explaining the variance in the endogenous variable. In most cases, the institutional quality of a country is likely to be stable over time. This is confirmed when looking at the correlation between the two variables which is at 78.5%. As a rule of thumb, the t-value of the instrument should be above 3.2. When Polity IV 1970-2010 is regressed on Polity IV 1960-1969 the t-value is at 14.0. The F-value of the instrument is reported in each table under the 2SLS column.

5. Analysis

5.1 NFW wells as the dependent variable

Table number 1 shows the result from five different regressions run with NFW wells as the dependent variable. The three first regressions in column 1, 2 and 3 come from an ordinary OLS regression. The control variables are added in segments to test the consistency of the results. The first two regressions represented in columns 1 and 2 are following equation (6) and (7). While column 3, 4 and 5 follows equation (1) presented in chapter 3.

$$NFW_i = \beta_0 + \beta_1 I_i + \beta_2 A_i + \beta_3 D_i + \beta_4 Y_i + u_i \tag{6}$$

$$NFW_i = \beta_0 + \beta_1 I_i + \beta_2 A_i + \beta_3 D_i + \beta_4 Y_i + \beta_5 E_i + \beta_6 O_i + \beta_7 R_i + \beta_8 L_i + u_i \tag{7}$$

In two out of three of the OLS regressions in columns 1-3, the institutional variable Polity IV is significant at a 5% level. Using the coefficient from column 3 indicates that an increase in the normalized Polity IV variable with 1/100 or 0.01 will increase the drilling activity with 1.41%. An increase in the Polity IV index by one point, for example from 6 to 7, would mean a change in the normalized index by 0.05. If Polity IV increases with 1 unit on average, everything else held constant, then the number of NFW wells drilled will increase with (100*1.41) *0.05=7.05%. If we look at two countries, like Brazil and Norway, there is a difference of 7 points on the Polity IV index between the countries in the period from 1970 to 2010. Norway has a score of 10, while Brazil has an average of 3. Using the estimates from column 3 indicates that the institutional difference would result in 49.4% more wells drilled in Norway compared to Brazil.

As already mentioned, there is a chance that institutional quality deteriorates with oil drilling. Oil companies might choose a country with good institutions, but then as oil dependence increases, the institutional quality decays, which could reduce drilling activity again. The estimated effect of institutional quality when using data from the drilling period could be undermined by the concurrent effect of the resource curse.

To avoid endogeneity problems and the spurious relations coming from omitted variable bias, I use Polity IV scores between 1960 and 1969 as an instrument for Polity IV scores from 1970-2010. Since the estimated drilling period has not yet commenced, the instrument is expected to be exogenous to the resource curse. Institutional quality is a stable variable that does not tend to fluctuate, and is thus a strong instrument in explaining the variation in current estimates.

Columns 4 and 5 show the results where Polity IV from 1960-69 has been used to estimate the effect on NFW wells. Column 4 shows an OLS regression, replacing the endogenous variable with the instrument. The results look similar to the results in column 1 and 3. The variable is significant at a 1% level, and is slightly smaller than the other coefficients.

Table 1: Regression results with NFW wells as the dependent variable

	Dependent variable: log (NFW wells drilled between 1970-2010)				
	(1) OLS simple regression	(2) OLS with political control variables	(3) OLS with geological control variables	(4) OLS reduced form	(5) 2SLS second stage
Polity IV score 1970-2010	1.379*** (0.494)	0.611 (0.771)	1.411** (0.659)		2.613*** (0.820)
Polity IV score 1960-69				1.338*** (0.433)	
Distance from equator	0.031*** (0.009)	-0.029* (0.018)	-0.030** (0.014)	-0.028* (0.014)	-0.038*** (0.013)
Area	0.614*** (0.111)	0.703*** (0.098)	0.728*** (0.081)	0.718*** (0.085)	0.751*** (0.081)
Years with drilling activity	-0.028*** (0.005)	-0.018*** (0.005)	-0.017*** (0.005)	-0.016*** (0.004)	-0.017*** (0.004)
Ethnolinguistic fractionalization		-0.375 (0.710)	-0.818 (0.650)	-0.818 (0.639)	-1.040* (0.601)
Landlocked		-0.847* (0.477)	-0.922** (0.454)	-0.750* (0.432)	-0.836** (0.394)
South Asia		-1.389** (0.612)	-3.096*** (0.821)	-3.139*** (0.806)	-3.466*** (0.862)
East Asia and Pacific		-1.547** (0.662)	-2.551*** (0.699)	-2.633*** (0.714)	-2.849*** (0.713)
Sub-Saharan Africa		-3.274*** (0.899)	-4.029*** (0.884)	-4.184*** (0.872)	-4.133*** (0.850)
Middle East and North Africa		-1.463* (0.803)	-3.106*** (0.944)	-3.049*** (0.838)	-3.209*** (0.942)
Latin America and Caribbean		-2.053** (0.827)	-3.237*** (0.825)	-3.122*** (0.830)	-3.715*** (0.876)
Western Europe		0.446 (0.705)	-0.265 (0.704)	-0.502 (0.697)	-0.753 (0.687)
North America		2.116*** (0.764)	1.233 (0.743)	1.027 (0.797)	0.631 (0.775)
Member of OPEC		0.015 (0.591)	0.166 (0.592)	0.066 (0.563)	0.281 (0.544)

Geological factors (NWF drilled)			1.273***	1.080***	1.505***
			(0.354)	(0.313)	(0.415)
Constant	-1.294	0.328	-0.813	-0.339	-1.506
	(1.712)	(1.800)	(1.620)	(1.574)	(1.539)
Observations	86	86	86	86	86
R^2	0.593	0.717	0.757	0.767	0.746
Adjusted R^2	0.573	0.661	0.704	0.717	0.691
F value (instrument)					44.73

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column 5 represents the second stage from a 2SLS regression, where Polity IV scores from 1970-2010 have been instrumented with the earlier estimates from 1960-69. The results from the first stage, estimated by equation (5) are presented in Table 2, together with the first stage results from the other 2SLS regressions for the other dependent variables.

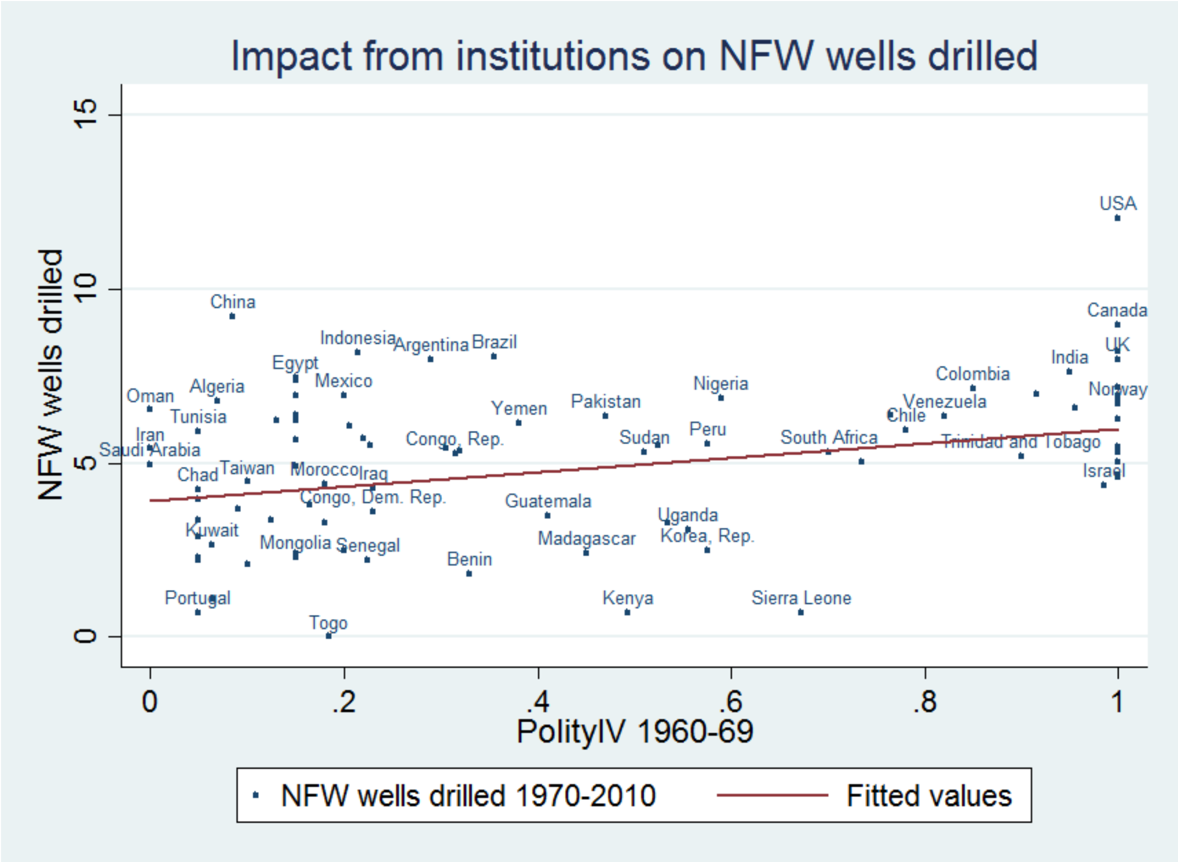
Table 2: Results from the first stage of the 2SLS regression

Dependent variable	Polity IV 1970-2010		
Sample	NFW wells	Drilldays	Successrate
Polity IV score 1960-69	0.512***	0.523***	0.553***
	(0.078)	(0.082)	(0.078)
Distance from equator	0.004	-0.004*	0.004
	(0.002)	(0.002)	(0.002)
Areal	-0.013	-0.011	-0.012
	(0.014)	(0.012)	(0.012)
Years with drilling activity	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Ethnolinguistic fractionalization	0.085	0.039	0.132
	(0.086)	(0.101)	(0.093)
Landlocked	0.033	0.015	0.035
	(0.047)	(0.047)	(0.037)
Member of OPEC	-0.082	-0.072	-0.109*
	(0.057)	(0.066)	(0.063)
South Asia	0.125	-0.085	0.010
	(0.133)	(0.116)	(0.085)
East Asia and Pacific	0.082	-0.035	0.024
	(0.096)	(0.080)	(0.085)
Sub-Saharan Africa	-0.019	-0.119	-0.088
	(0.118)	(0.119)	(0.111)
Middle East and North Africa	0.061	-0.159	0.010
	(0.138)	(0.105)	(0.120)
Latin America and Caribbean	0.227**	0.098	0.204**
	(0.113)	(0.104)	(0.099)
Western Europe	0.096	0.030	0.023
	(0.081)	(0.082)	(0.080)
North America	0.152	0.081	-0.036
	(0.098)	(0.097)	(0.108)
Geological factors (Successrate)	-0.163***	-0.011	1.042**
	(0.062)	(0.033)	(0.375)
Constant	0.447***	0.447	-0.638
	(0.191)	(0.379)	(0.388)
Observations	86	86	86
R2	0.809	0.772	0.805
Adjusted R2	0.768	0.724	0.763
F value (instrument)	42.77	40.79	49.96

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The scatterplot shows the dispersion of NFW wells drilled vs the Polity IV score from 1960-69[1]. The regression line is estimated by using the regression presented in column 4. The line indicates a positive relationship between institutional quality and the number of wells drilled.

Figure 10: Scatterplot of NFW wells drilled vs Polity IV score 1960-69



If the hypothesis about the resource curse is correct, the curse would have a negative impact on the institutional variable in the ordinary OLS regressions in column 1-3. The 2SLS regression uses only the systematic variance from the Polity IV 1970-2010 variable and excludes the

[1] Some names have been excluded from the scatterplot in order to make the graph readable. A full list of the countries in the sample can be found in the appendix.

unsystematic variance where there might be omitted variable bias from the resource curse (Balsvik, 2013). The variable turns out highly significant and the value has increased with 1.2 percentage points from the OLS regressions. This supports the hypothesis that institutional quality has worsened with oil drilling, and that the direct impact from institutions on drilling decisions is in fact stronger than what the OLS estimates indicate. The 2SLS result indicates an even stronger correlation between drilling activity and institutional quality. The omitted variable bias in the ordinary OLS regressions seem to have a negative impact on the drilling activity, which supports the existence of a resource curse.

According to column 5, an improvement in the average Polity score, with 1 point would increase drilling activity, ceteris paribus, with 13%. In some cases, the IV estimation can result in odd results, if the relevance condition of the instrument does not hold (Woolridge, 2009). The significance of the F-value (44.73) presented in column 5, indicates that this is not the case in this 2SLS regressions and that the results can be trusted.

I expected a positive correlation between the number of wells drilled and the number of years with drilling activity, but the correlation turns out negative in all cases. The reason is probably that drilling activity tend to be centralized around certain periods and not spread over several decades. Countries with more years of drilling tend to have a higher number of wells drilled, but the effect is removed when divided by the higher number of drilling years. The coefficients for ethnic fractionalization are negative, but never turns out significant at a 5% level. It does not seem like culture has had a major impact on drilling decisions. Landlockedness has a negative impact as expected, and turns out significant in two out of four regressions.

Most of the regions seem to be associated with a negative impact on wells drilled, except North America, which turns out positive, due to the high number of wells drilled in the US. Sub-Saharan Africa has a strong negative significant impact on drilling, which is also one of the regions associated with high country risk. A more surprising result is that the coefficients for Middle East and North Africa are significant and negative. The OPEC variable, representing many countries in the region, could offset this finding, but does not turn out significant in any of the regressions.

5.2 Drilldays as the dependent variable

The next regressions use the number of drilldays spent in each country between 1970 and 2010 as a dependent variable.

Table 3 shows the results from the regressions. The first three are OLS results with Polity IV 1970-2010 as an explanatory variable, while the last two columns show an OLS regression and the second stage of a 2SLS regression, using Polity IV 1960-69 as an instrument to avoid spurious correlation. The presentation of regressions in columns 1 and 2 follows the same order as in equation (6) and (7), but with drilldays as the dependent variable. Columns 3,4 and 5 follows equation (2) from chapter 3. In the regression using the fewest variables, the Polity IV variable does not show signs of significance. The estimated impact of institutional quality is highly significant in the three last regressions, where all the control variables are included. Column 3 indicates that increasing the average Polity score with 1 point, ceteris paribus, will increase the total number of days spent drilling in a country with 9.5%. If we compare two countries again, for instance the US and Mexico, the difference in institutional quality differs with 9 points, US has a score of 10 points, while Mexico had an average score of 1 point between 1970 and 2010. Although the countries share a border, and thus probably some similar geological conditions, the regression results suggest that the number of drilling days would be 85% higher in the US compared to Mexico.

The results remain robust when using Polity IV 1960-69 as an instrument. The coefficient from the second stage of the 2SLS regression is more than one percentage point larger than the coefficient from regression 3, again supporting the notion of the negative bias from an omitted resource curse variable.

The same clear trend occurs here as in the first regressions with NFW wells. Institutions matter for how many days oil companies decide to drill in a country. It is interesting that the effect of institutional quality stays significant when using a completely different dataset for the dependent variable. It indicates that the impact of institutional quality is robust to big changes.

There are no radical changes in the other control variables compared to the first table. The only difference is that the countries in North America have a significantly lower score. This because the US data point is not as extreme in this dataset as in the NFW case.

Table 3: Regression results with drilldays as the dependent variable

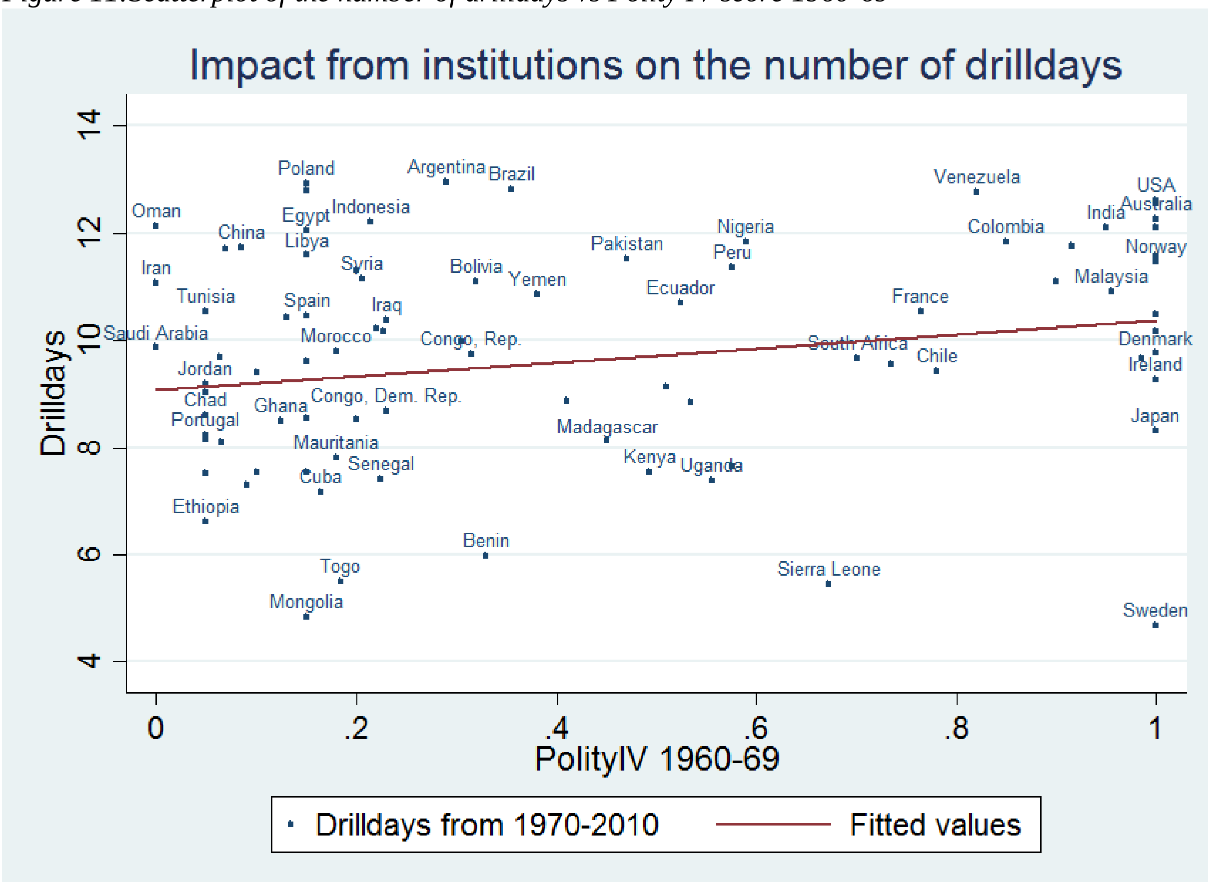
	Dependent variable: Log (Days spent drilling between 1970 and 2010)				
	(1) OLS simple regression	(2) OLS with political control variables	(3) OLS with geological control variables	(4) OLS reduced form	(5) 2SLS second stage
Polity IV score 1970-2010	0.846	1.334*	1.889***		2.959***
	(0.512)	(0.717)	(0.655)		(0.774)
Polity IV score 1960-69				1.556***	
				(0.465)	
Distance from equator	0.013	-0.042**	-0.032*	-0.026	-0.038***
	(0.010)	(0.019)	(0.016)	(0.016)	(0.015)
Area	0.324***	0.456***	0.441***	0.426***	0.458***
	(0.098)	(0.099)	(0.078)	(0.082)	(0.067)
Years with drilling activity	-0.033***	-0.024***	-0.021***	-0.020***	-0.021***
	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)
Ethnolinguistic fractionalization		-0.170	-0.181	-0.204	-0.318
		(0.698)	(0.581)	(0.564)	(0.543)
Landlocked		-0.917**	-0.749*	-0.592	-0.638*
		(0.387)	(0.422)	(0.437)	(0.384)
South Asia		-1.570**	-2.310***	-2.682***	-2.430***
		(0.690)	(0.699)	(0.723)	(0.647)
East Asia and Pacific		-2.188***	-2.620***	-2.872***	-2.768***
		(0.709)	(0.682)	(0.699)	(0.614)
Sub-Saharan Africa		-3.168***	-3.455***	-3.798***	-3.447***
		(0.902)	(0.888)	(0.879)	(0.826)
Middle East and North Africa		-0.757	-1.572**	-1.922***	-1.452**
		(0.716)	(0.726)	(0.677)	(0.700)
Latin America and Caribbean		-1.985**	-2.210***	-2.193**	-2.482***
		(0.867)	(0.829)	(0.874)	(0.743)
Western Europe		-0.462	-0.492	-0.752	-0.841
		(0.812)	(0.764)	(0.742)	(0.707)
North America		-0.636	-1.823**	-2.112***	-2.352***
		(0.772)	(0.740)	(0.737)	(0.670)
Member of OPEC		0.389	0.498	0.379	0.592
		(0.565)	(0.548)	(0.517)	(0.509)
Geological factors (Drilldays)			1.180***	1.250***	1.282***
			(0.330)	(0.330)	(0.291)
Constant	8.543***	8.813***	-2.565	-2.695	-4.017
	(1.880)	(2.018)	(3.775)	(3.651)	(3.382)
Observations	86	86	86	86	86

R^2	0.544	0.676	0.743	0.746	0.732
Adjusted R^2	0.521	0.612	0.688	0.692	0.675
F value (instrument)					40.79

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The scatterplot shows the distribution of observations of the total number of drilldays in a country against the Polity IV score 1960-69, the regression line presented is from column 5 and demonstrates a clear tendency that the number of drilldays increases when the institutional quality rises.

Figure 11:Scatterplot of the number of drilldays vs Polity IV score 1960-69



5.3 Success rate as the dependent variable

The first two dependent variables, NFW wells and drilldays, measured the interest and the willingness to invest in countries with good institutions. The last dependent variable displays something the oil companies cannot control: the likelihood of finding oil. The dependent variable estimates the success rate of NFW wells drilled in the period from 1970 to 2010.

The results presented in table 3 are quite striking. The coefficient for institutional quality is consistently negative, and significant at a 5% level in two OLS regressions and at a 10% level in the second stage of the 2SLS regression. This differs from the results of Cust and Harding (2014), who are not able to prove a significant correlation between the discovery rate and institutional quality. Everything else held constant, an increase in Polity score with 1 point will reduce the probability of finding oil with (-0.2*100) *0.05 = 1 percentage point. If we use the same example with the US and Mexico, the likelihood of finding oil would be 9 percentage points lower in the US compared to Mexico.

There is no difference between the coefficient in the 2SLS regression and the other models where Polity IV is significant. In contrast to NFW wells and drilldays, we can assume that the success of finding oil is quite exogenous and will thus not be affected by institutional quality. Even if the success of finding oil can deteriorate institutions, the coefficients of the institutional variable are negative, which is also the direction of the resource curse bias.

Some of the other control variables differ from the previous regressions. Surface area is a negative explanatory factor to finding oil in the success rate model. Years with drilling activity has turned positive, although the value is low. One possible explanation may be that companies know where to search for oil in countries where a lot of oil has previously been found, although one would expect these countries to have exhausted their reserves already. Most of the regions turn out positive, although rarely significant, while OPEC membership has increased its explanatory value in this case.

The results indicate that although companies have searched more and spent more time drilling in countries with better institutions, there has not been found more oil in these countries. The lack of a positive relationship indicates that geology is not affected by institutional quality. We can assume that countries with good institutions can not affect the natural capital lying in their soil the same way they can affect industrialized growth for instance. What is more interesting is that there is a significant negative correlation between good institutions and oil discovery. The reason may be that countries with good institutions have exploited most of their reserves earlier due to higher drilling rates, while there are more “low-hanging fruits” i.e., oil that is easy to find, in countries with weak institutions. Since exploration is expensive, oil companies choose to drill fewer wells in risky countries, which can lead to very high discovery rates when only wells with very good geological prospects are drilled. It is possible that oil companies have perceived the costs associated with operating in high-risk countries to outweigh the potential of finding more oil.

Table 4: Regression results with success rate as the dependent variable

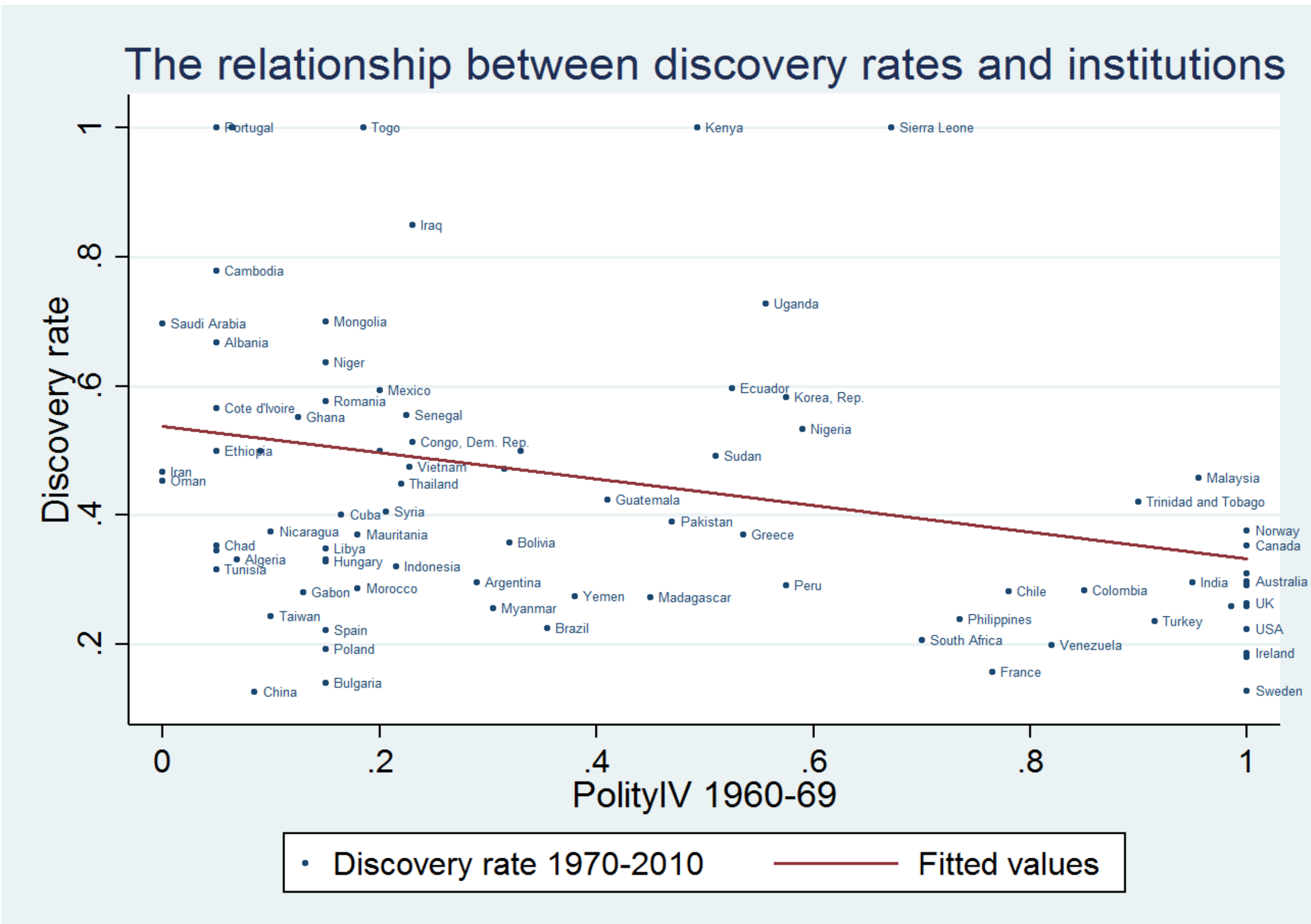
	Dependent variable: Success rate NFW wells drilled 1970-2010				
	(1) OLS simple regression	(2) OLS with political control variables	(3) OLS with geological control variables	(4) OLS reduced form	(5) 2SLS second stage
Polity IV score 1970-2010	-0.206*** (0.070)	-0.147 (0.122)	-0.211** (0.103)		-0.204* (0.114)
Polity IV score 1960-69				-0.113 (0.075)	
Distance from equator	-0.003** (0.001)	-0.001 (0.002)	-0.001 (0.002)	-0.002 (0.002)	-0.001 (0.002)
Area	-0.039*** (0.013)	-0.053*** (0.013)	-0.056*** (0.012)	-0.054*** (0.013)	-0.056*** (0.011)
Years with drilling activity	0.001*** (0.001)	0.001** (0.001)	0.001** (0.000)	0.001** (0.000)	0.001*** (0.000)
Ethnolinguistic fractionalization		-0.054 (0.124)	0.063 (0.107)	0.035 (0.109)	0.062 (0.093)
Landlocked		0.111 (0.071)	0.120 (0.074)	0.113 (0.075)	0.121* (0.066)
South Asia		0.135 (0.120)	0.260** (0.126)	0.257** (0.126)	0.259** (0.117)
East Asia and Pacific		0.057 (0.113)	0.142 (0.115)	0.136 (0.118)	0.141 (0.107)
Sub-Saharan Africa		0.151 (0.140)	0.192 (0.136)	0.209 (0.137)	0.192 (0.123)
Middle East and North Africa		0.050 (0.131)	0.235 (0.150)	0.233* (0.139)	0.235* (0.134)
Latin America and Caribbean		0.088 (0.124)	0.231* (0.136)	0.188 (0.138)	0.229* (0.133)
Western Europe		-0.018 (0.101)	0.012 (0.101)	0.005 (0.104)	0.010 (0.091)
North America		0.199* (0.114)	0.124 (0.130)	0.129 (0.132)	0.122 (0.120)
Member of OPEC		0.184** (0.076)	0.136* (0.077)	0.159* (0.082)	0.137* (0.073)
Geological factors (Successrate)			1.162*** (0.341)	0.944*** (0.327)	1.156*** (0.336)
Constant	0.950***	1.006***	-0.055	0.077	-0.053

	(0.200)	(0.219)	(0.368)	(0.362)	(0.338)
Observations	86	86	86	86	86
R^2	0.339	0.429	0.507	0.489	0.507
Adjusted R^2	0.306	0.316	0.402	0.379	0.402
F value (instrument)					49.96

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The scatterplot shows the distribution between the success rate and instutional quality. The regression line from column 5 indicates a clear negative relationship between the Polity IV score and the success rate.

Figure 12: Scatterplot of the success rate vs Polity IV score 1960-69.



6. Regressions run on sub-samples of the dataset

This chapter presents the results from regressions run on a different sample of countries than the main regressions in chapter 5.

6.1 High-income and low-income countries

To address the potential issues of the omitted variable bias from income, I run a regression using a sample consisting of low income countries only. The categories originate from the dataset made by Davis and Hopkins, which defines different categories for income level. I use the two lowest categories; low income and lower middle income to create a sample consisting of 42 countries. The results from using equation (1) for ordinary OLS and 2SLS are presented in table 5. Since the sample has been reduced to half its size, I chose to leave out the regional variables. The variables for ethnic fractionalization, landlockedness, OPEC membership and the geological vector are included in the regression.

It turns out that the assumption that institutional quality is important for drilling activity does not hold when looking at a sample of low-income countries. The Polity IV variable does not turn significant in any of the regressions. In the 2SLS regression for success rate the coefficient turns positive.

The regression probably suffers from a small sample, which makes it harder to prove significance due to larger standard errors. Another issue is that the Polity IV 60-69 instrument is very weak. While the usual F-value has been at around 40, the first stage regressions return F-values between 4.40 and 7.08. It seems like institutions are less stable in low-income countries and the Polity IV 69-69 loses much of its explanatory power as an instrument for more current Polity ratings.

The results indicate that the findings from the main regressions should be treated with caution. It is not evident that drilling activity will increase in all countries just because institutional quality improves: perhaps it is a more general transition towards a stable, industrialized society that determines drilling decisions. Identifying institutional quality alone as a determinant for FDI does not necessarily capture all the other factors involved in the decision process.

Table 5: Results from regressions run on a sample of low and lower middle income countries

Dependent variable	Log (NFW wells 1970-2010)		Log (Drilldays 1970-2010)		Success rate 1970-2010	
	(1) Ordinary OLS	(2) 2SLS second stage	(3) Ordinary OLS	(4) 2SLS second stage	(5) Ordinary OLS	(6) 2SLS second stage
Polity IV score 1970-2010	0.988	1.335	1.521	1.698	-0.199	0.269
	(1.161)	(2.258)	(0.980)	(2.241)	(0.153)	(0.357)
Distance from equator	-0.029	-0.028	0.002	0.003	-0.002	-0.001
	(0.020)	(0.018)	(0.017)	(0.016)	(0.002)	(0.003)
Area	0.631***	0.643***	0.419***	0.424***	-0.074***	-0.055**
	(0.197)	(0.197)	(0.113)	(0.121)	(0.018)	(0.024)
Years with drilling activity	-0.030***	-0.030***	-0.024***	-0.024***	0.001	0.002*
	(0.007)	(0.007)	(0.007)	(0.008)	(0.001)	(0.001)
Political control variables	Yes	Yes	Yes	Yes	Yes	Yes
Geological control variables	Yes	Yes	Yes	Yes	Yes	Yes
Regional control variables	No	No	No	No	No	No
Constant	-0.466	-0.825	-3.320	-3.643	0.652	0.452
	(2.777)	(3.386)	(3.948)	(5.575)	(0.394)	(0.463)
Observations	42	42	42	42	42	42
R^2	0.706	0.705	0.729	0.729	0.494	0.371
Adjusted R^2	0.635	0.634	0.664	0.663	0.372	0.219
F value (instrument)		5.67		4.40		7.08

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I run the same regressions on the remaining sample of upper middle income and high income countries. The sample holds two more countries than the low-income group, but the results are significant at a much lower level, and with larger coefficients than in the low-income sample. In terms of drilling activity (NFW wells and drilldays) the 2SLS turns out significant at a 5 or 10% level, and the coefficients are almost twice the size, compared to the table above. The most interesting result however, is when we look at column 5 and 6 for success rate. Both the OLS and the 2SLS regressions are significant at a 5% level with coefficients -0.25 and -0.34 respectively. Compared to the results presented in the main findings, significance has increased in the 2SLS regression and the impact measured by the coefficients has changed by -0.15. This indicates that, ceteris paribus, the likelihood of finding oil decreases with 1.25 percentage points when the Polity variable increases with 1 point in average.

Table 6: Results from regressions on a sample of high and higher middle income countries

Dependent variable	Log (NFW wells 1970-2010)		Log (Drilldays 1970-2010)		Success rate 1970-2010	
	(1) Ordinary OLS	(2) 2SLS second stage	(3) Ordinary OLS	(4) 2SLS second stage	(5) Ordinary OLS	(6) 2SLS second stage
Polity IV score 1970-2010	1.288	2.516**	1.754	2.209*	-0.249**	-0.344**
	(0.793)	(1.037)	(1.122)	(1.198)	(0.117)	(0.161)
Distance from equator	0.031*	0.022	0.003	-0.000	-0.003	-0.002
	(0.018)	(0.015)	(0.019)	(0.018)	(0.003)	(0.002)
Area	0.630***	0.635***	0.292**	0.293***	-0.033*	-0.033*
	(0.136)	(0.118)	(0.108)	(0.096)	(0.019)	(0.018)
Years with drilling activity	-0.021***	-0.020***	-0.027***	-0.026***	0.001	0.001*
	(0.008)	(0.007)	(0.009)	(0.008)	(0.001)	(0.001)
Political control variables	Yes	Yes	Yes	Yes	Yes	Yes
Geological control variables	Yes	Yes	Yes	Yes	Yes	Yes
Regional control variables	No	No	No	No	No	No
Constant	-2.602	-3.373	-1.027	-1.696	0.402	0.400
	(2.575)	(2.530)	(5.250)	(4.680)	(0.261)	(0.254)
Observations	44	44	44	44	44	44
R^2	0.558	0.536	0.580	0.577	0.322	0.309
Adjusted R^2	0.457	0.430	0.484	0.480	0.167	0.151
F value (instrument)		40.86		41.06		41.67

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

It seems like there is a stronger tendency using institutional quality as a decision tool for exploration and drilling in high-income countries than in low-income countries. These countries have probably had more drilling activity than those with a lower GDP and some of the countries might owe their income status to its natural resources. The strong negative correlation between success rate and institutional quality could either be explained by the fact that most of the resources have been exhausted earlier, or that rich countries have spent more time and money searching for oil. The results did however, not correspond with the incentives.

6.2 Onshore versus offshore drilling

It is possible that drilling activities offshore will be less affected by incidents, such as civil war or a state coup. For instance, Statoil commenced its offshore operations in Angola in 1991 during the civil war that lasted from 1975 to 2002 (Statoil, 2009) (, 2015). To test the hypothesis that institutional quality is more important for onshore than offshore drilling, I run the regressions using new field wildcats drilled onshore and offshore, respectively.

Again, the regional dummy variables are omitted to reduce the problem with a small sample and high standard errors. In contrast to the hypothesis, the impact of institutional quality on offshore drilling is more significant and with a larger magnitude than in the onshore case. A reason could be that offshore investments require substantially higher capital expenditures, and IOCs will normally be very cautious before making offshore commitments in unstable countries.

Table 7: Results from regressions run on samples of onshore and offshore NFW wells

Dependent variable:	NFW drilled onshore			NFW drilled offshore		
	(1) Ordinary OLS	(2) OLS reduced form	(3) 2SLS second stage	(4) Ordinary OLS	(5) OLS reduced form	(6) 2SLS second stage
Polity IV score 1970-2010	1.181*		2.135**	2.617***		3.431***
	(0.680)		(0.924)	(0.836)		(1.007)
Polity IV score 1960-69		1.281**			2.053***	
		(0.561)			(0.625)	
Distance from equator	0.040**	0.038**	0.033**	0.006	0.008	-0.001
	(0.017)	(0.016)	(0.016)	(0.022)	(0.021)	(0.021)
Area	0.816***	0.808***	0.835***	0.518***	0.474***	0.532***
	(0.129)	(0.126)	(0.126)	(0.134)	(0.147)	(0.123)
Years with drilling activity	-0.027***	-0.026***	-0.026***	-0.009	-0.010	-0.008
	(0.006)	(0.006)	(0.005)	(0.007)	(0.007)	(0.007)
Ethnolinguistic fractionalization	-0.623	-0.881	-0.872	0.305	0.069	0.109
	(1.053)	(0.993)	(0.952)	(0.990)	(0.973)	(0.931)
Landlocked	-0.223	-0.048	-0.140	Omitted	Omitted	Omitted
	(0.598)	(0.566)	(0.570)	(.)	(.)	(.)
Member of OPEC	0.417	0.444	0.578	-0.689	-0.737	-0.580
	(0.501)	(0.483)	(0.461)	(0.790)	(0.709)	(0.782)
Geological factors (NWF drilled)	0.300	0.188	0.531	0.319	0.022	0.470
	(0.353)	(0.334)	(0.410)	(0.447)	(0.461)	(0.444)
Regional variables	No	No	No	No	No	No
Constant	-4.807**	-4.410**	-5.662**	-4.032*	-2.382	-4.661**
	(2.220)	(2.081)	(2.289)	(2.308)	(2.418)	(2.259)
Observations	76	76	76	67	67	67
R^2	0.570	0.585	0.560	0.330	0.333	0.321
Adjusted R^2	0.519	0.535	0.508	0.250	0.254	0.241
F value (instrument)			98.05			87.71

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7. Robustness testing

In this section I test the robustness of the results by changing the dependent variable, the institutional variable and by including and excluding observations.

7.1 Exchanging the dependent variable from IHS with a dependent variable from Wood Mackenzie

I replace the dependent variables coming from the IHS dataset and represented in equation (1) and (3), NFW wells and discovery rate, with the same variables originating from the Wood Mackenzie dataset. The variable used to replace NFW wells is the total number of wells drilled in each country between 1970 and 2010. The number includes both exploratory wells and appraisal wells, which are usually wells drilled after the first NFW wells and used to further investigate whether an area contains oil and gas, and the extent of the reserves. The discovery rate comes from the number of wells that are reported to be dry. The results from the regression are presented in table 8.

Columns 1 and 2 display the results from the ordinary OLS regression, using all variables and the 2SLS regression using the Polity IV score from 1960-1970 as an instrument for institutional quality. The Polity IV variable turns out highly significant in both regressions. The coefficients are larger in magnitude than in the regressions with NFW wells. An increase in the Polity IV index with one point in average would, ceteris paribus, increase drilling activity with 11.3%. The 2SLS coefficients are larger than the ordinary OLS regressions, supporting the presence of a resource curse.

The results in columns 3 and 4 come from the regressions using the discovery rate as a dependent variable. While the IHS data returned a significant and negative correlation between the Polity IV score and the discovery rate, the results are not as robust when using the data from Wood Mackenzie. The ordinary OLS regression returns a significant result at 10% and a negative coefficient, while the 2SLS coefficient is neither significant nor negative. A reason for this could be that the probability of finding oil is less exogenous when drilling appraisal wells compared to new field wildcats.

Table 8: Results from regression replacing the dependent variables measuring number of wells and success rate

	Total number of wells		Success rate	
	(1) Ordinary OLS	(2) 2SLS second stage	(3) Ordinary OLS	(4) 2SLS second stage
Polity IV score 1970-2010	2.252***	3.445***	-0.140*	0.109
	(0.655)	(0.768)	(0.083)	(0.142)
Distance from equator	-0.022	-0.028*	-0.002	-0.003
	(0.018)	(0.016)	(0.002)	(0.002)
Areal	0.522***	0.544***	0.010	0.015
	(0.081)	(0.068)	(0.014)	(0.014)
Years with drilling activity	-0.020***	-0.020***	-0.001*	-0.001**
	(0.005)	(0.004)	(0.001)	(0.001)
Ethnolinguistic fractionalization	-0.572	-0.745	-0.070	-0.107
	(0.618)	(0.583)	(0.144)	(0.133)
Landlocked	-0.592	-0.462	0.142	0.169*
	(0.386)	(0.352)	(0.098)	(0.091)
Member of OPEC	0.266	0.369	0.308***	0.330***
	(0.560)	(0.514)	(0.060)	(0.060)
Regional control variables	Yes	Yes	Yes	Yes
Geological control variables	Yes	Yes	Yes	Yes
Constant	-1.674	-2.561	0.709**	0.524**
	(1.991)	(1.751)	(0.302)	(0.262)
Observations	86	86	86	86
R^2	0.725	0.710	0.353	0.303
Adjusted R^2	0.666	0.648	0.214	0.154
F Value (Instrument)		43.98		43.97

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.2 Replacing Polity IV with Freedom House

To investigate whether the results hold for other types of measures of institutional quality, I replace the Polity IV 1970-2010 variable with an averaged Freedom House (FH) score. The FH score dates back to 1973, so for the 2SLS regression I instrument it with the Polity IV 1960-1969 ratings.

The results are similar to those presented in chapter 6. In the regressions with NFW wells and drilldays the FH variable returns the same significance and in drilldays the magnitude of the variables has increased with a couple of decimal points. The OLS regression for success rate has lost some significance, compared to the results in table 4, but otherwise there are small differences in the outcome.

Table 9: Results from regressions replacing Polity IV with Freedom House

	(1) NFW OLS	(2) NFW 2SLS	(3) Drilldays OLS	(4) Drilldays 2SLS	(5) Successrate OLS	(6) Successrate 2SLS
Freedom House score 1973-2010	1.332**	2.963***	1.924***	3.392***	-0.170*	-0.229*
	(0.638)	(0.981)	(0.660)	(0.885)	(0.100)	(0.132)
Distance from equator	-0.032**	-0.045***	-0.035**	-0.046***	-0.001	-0.001
	(0.015)	(0.015)	(0.017)	(0.015)	(0.002)	(0.002)
Area	0.732***	0.768***	0.448***	0.475***	-0.056***	-0.057***
	(0.082)	(0.087)	(0.079)	(0.069)	(0.013)	(0.012)
Years with drilling activity	-0.016***	-0.016***	-0.020***	-0.020***	0.001**	0.001***
	(0.005)	(0.004)	(0.006)	(0.005)	(0.000)	(0.000)
Regional control variables	Yes	Yes	Yes	Yes	Yes	Yes
Political control variables	Yes	Yes	Yes	Yes	Yes	Yes
Geological control variables	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.709	-1.577	-2.633	-4.643	-0.034	-0.046
	(1.593)	(1.563)	(3.867)	(3.492)	(0.359)	(0.338)
Observations	86	86	86	86	86	86
R^2	0.755	0.735	0.744	0.723	0.495	0.492
Adjusted R^2	0.702	0.678	0.690	0.664	0.387	0.383
F value (instrument)		28.63		25.91		30.99

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The findings indicate that the impact of institutional quality does not rely only on one estimate of democracy and political freedom. This makes sense, as it is not likely that oil companies depended merely on the ratings provided by Polity IV in the decision making. It makes it harder, however, to determine what kind of factors that are the most important for an oil company when making drilling decisions. As mentioned previously, it is possible that institutional quality works as a proxy for both political stability and attractive FDI policies, or for a more general path, where improved institutions and economic development coincides. A potential topic for further investigation would be to separate and identify the impact of these effects on oil investments.

7.3 Excluding outliers

I mention in the results that the value for the United States in the NFW dataset has a much higher value than the rest of the observations. The effect is reduced when applying a logarithmic transformation on the variable. I compute three statistics to help me identify potential outliers in the datasets; Cook's distance, leverage and studentized residuals (Jacoby, 2005). It turns out that none of the observations have a Cook's value close to or higher than one, which indicates that there are no outliers, as a rule of thumb. Four countries return a higher leverage than the estimated limit in all regressions; Afghanistan, Canada, India, and USA, which I exclude from the new sample. After looking at the distribution from added variable plots, I also chose to leave Saudi Arabia out of the new sample. The results from the regressions with the new sample are presented in the appendix, table 11.

Even when reducing the sample to 81 countries, the outcome of the regressions for NFW wells and drilldays are similar to the results presented in the beginning of the analysis. The ordinary OLS for NFW wells is only significant at a 10% level, while the OLS in reduced form and the 2SLS does not change. The same goes for drilldays where neither significance nor the magnitude of the coefficients change that much. The outcome of success rate has changed, the Polity IV variable loses significance in all regressions and is only significant at a 10% level in the ordinary OLS regression. It seems like this variable is more sensitive to excluding outliers than the other two.

7.4 Adding more observations to the regression

Two of the dependent variables used in the analysis have been used in logarithmic form. This means that some observations with the value zero have been excluded, since the logarithm of zero is not defined (Woolridge, 2009). It turns out that ten observations from the NFW dataset have been excluded from the sample because of this. The Wood Mackenzie dataset only enlists wells that have been drilled, countries without any wells are thus not included in the dataset and there are no extra countries with zero wells to be added from the dataset. To include all countries, I add the value of 1 to all the observations of NFW wells and estimate a new logarithm. The new sample includes ten new countries: Costa Rica, Gambia, Guinea, Honduras, Liberia, Mali, Panama, Somalia, Sri Lanka and Uruguay. I run all regressions with the three dependent variables on the new sample to check the robustness of the results when expanding the sample. The results are presented in the appendix in table 12, 13 and 14.

Since all the new variables represent an extreme value (zero) in terms of NFW wells drilled, it is their average Polity score that will determine the outcome of the regression. When looking at the average Polity Score, the 86 countries of the original sample have a normalized score of 0.43 between 1960 and 1969, and an average score of 0.55, in the period from 1970 to 2010. In comparison, the group of ten countries with zero NFW wells drilled has an average Polity score of 0.59 in both the time periods. When using 1960 to 1969 as an instrument, the additional countries lie 0.15 normalized points above the main sample, which on the Polity IV score would mean a difference of 3 points.

The addition of countries with a higher Polity IV rating reduces the significance of the variable when using NFW as a dependent variable. Institutional quality is not significant in any of the regressions, and the coefficients have been reduced to half of their original size.

When looking at drilldays, the original results are more robust. The countries that have been added to the dataset do not hold any zero values in the Wood Mackenzie dataset, and the impact is less extreme. Polity IV turns out significant in 2 out of 5 regressions, and the value of the coefficients are slightly lower than in the main regression.

The new observations of NFW wells are also directly connected to success rate, since it measures the discovery rate of NFW wells drilled. All the observations added have a discovery rate of zero. In contrast to the NFW results, success rate and institutional quality are already negatively connected. Expanding the dataset strengthens the significance of most of the regressions, and the coefficients have doubled in magnitude in most of the cases.

The regressions with the expanded sample indicate that the results are sensitive when adding more variables to the sample. Given the nature of the data, the results depend largely on the Polity IV values to the countries who are added.

8. Conclusion and summary

The analysis in this thesis shows that institutional quality has indeed mattered for drilling activity. The results display that both the number of test wells drilled and the number of days spent drilling in a country are influenced by the institutional quality, as measured by the Polity IV variable. The results are remarkably robust. The institutional variable is significant at a 5% level in 4 out of 5 regressions with NFW wells, and 3 out of 5 regressions when measuring drilldays.

Although companies tend to drill more in countries with good institutional quality, the results show that the success rate in such countries is not higher. Institutional quality has a significant and negative impact on the probability of finding oil. Possible explanations could be that most oil in countries with high institutional quality has already been depleted, and that the remaining reserves are higher in countries with worse institutions. Another reason could be that oil companies would only enter a high-risk country if the likelihood of finding oil is very high.

Other results show that the impact of institutions only hold for rich countries, when dividing the countries into two subsamples, based on income. It turns out that institutional quality matters more for offshore than onshore investments.

When changing the dependent variable for number of wells drilled, the results stay significant and increase in magnitude. The same goes for changing the variable for institutional quality as well to removing outliers. The explanatory value and significance is reduced when countries with zero values are added to the sample and does not hold when the dependent variable for success rates is replaced.

It is important to treat the results with caution, as they might not apply to all countries. There are several factors to consider when making drilling decisions, and an increase in institutional quality alone is not necessarily a guarantee for improved interest from oil companies. Investments in infrastructure, fiscal systems, educated workers and long-term stability might be just as important. Although the analysis show that historically, institutional quality has been an important determinant of oil exploration and activity, there is no guarantee that this will continue.

9. References

Acemoglu, D., Johnson, S., Robinson, J. A. & Yared, P., 2005. *Income and democracy,* National Bureau of Economic Research.

Acemoglu, D. & Robinson, J. A., 2012. *Why nations fail.* 1. ed. New York: Crown Publishers.

Adams, C., 2015. *Financial Times.* [Online]
Available at: http://www.ft.com/cms/s/0/2d35d198-d6b7-11e4-97c3-00144feab7de.html#axzz3fIbUVj7r
[Accessed 07 July 2015].

Alon, I., Gurumoorthy, R., Mitchell, M. C. & Steen, T., 2006. Managing Micropolitical Risk: A Cross-Sector Examination. *Thunderbird International Business Review,* 48(5), pp. 623-642.

Aven, E. & Flørenæs, S., 2004. Country risk: quantitative measurement and analysis. *Global association of risk professionals,* May/June(18), pp. 44-47.

Balsvik, R., 2013. *Heteroscedasticity and Specification problems. Lecture in Econometric techniques at NHH.* Bergen, 10 October 2013.

Balsvik, R., 2013. *Instrumental Variables and 2SLS. Lecture in Econometric techniques at NHH.*Bergen, 30 October 2013.

BBC, 2015. Angola Country Profile- overview. [Online]
Available at: http://www.bbc.com/news/world-africa-13036732
[Accessed 05 August 2015].

BERA, 2005/2006. The oil and gas industry. *Business and Economics Research Advisor*, Issue 5/6.

Bohn, H. & Deacon, R. T., 2000. Ownership Risk, Investment and the Use of Natural Resources. *The American Economic Review,* 90(3), pp. 526-549.

Bret-Rouzaut, N. & Favennec , J.-P., 2007. *Oil and gas exploration and production: reserves, costs, contracts.* 3 ed. Paris: Technip.

Cairn Energy, 2015. *Oil and gas exploration, and production life cycle.* [Online]
Available at: http://www.cairnenergy.com/index.asp?pageid=554
[Accessed 25 September 2015].

Collier, P., 2007. *The bottom billion: why the poorest countries are failing and what can be done about it.* 1st ed. Oxford: Oxford University Press.

Collier, P. & Hoeffler, A., 2009. Testing the Neocon agenda: democracy in resource-rich societies. *European Economic Review (2009),* Issue 53.3, pp. 293-308.

Collier, P. & Venables, A. J., 2011. Key decisions for resource management: principles and practice. In: P. Collier & A. J. Venables, eds. *Plundered Nations? Successes and Failures in Natural Resource Extraction.* London: Palgrave Macmillan.

Crooks, E., 2015. *Discoveries of new oil and gas reserves drop to 20-year low*, s.l.: s.n.

Cust, J. & Harding, T., 2014. Institutionas and the location of oil exploration. *OxCarre Research Paper* , Volume 127.

David, P. A. & Wright, G., 1997. Increasing returns and the genesis of American resource abundance. *Industrial and Corporate Change,* 6(2), pp. 203-245.

De Rosa, D., & Iootty, M., 2012. Are natural resources cursed? an investigation of the dynamic effects of resource dependence on institutional quality. *World Bank Policy Research Working Paper*, (6151).

Deloitte, 2014. *Oil and gas taxation in Norway.* [Online]
Available at: https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Energy-and-Resources/gx-er-oil-and-gas-taxguide-norway.pdf
[Accessed 07 July 2015].

Dollar, D. & Kray, A., 2003. Institutions, trade, and growth. *Journal of Monetary Economics*, pp. 133-162.

Easterly, W. & Levine, R., 1997. Africa's growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics* , pp. 1203-1250.

Faye, M. L. e. a., 2004. The challenges facing landlocked developing countries. *Journal of Human Development* , 5(1), pp. 31-68.

Frynas, J. G., 1998. Political instability and business: focus on Shell in Nigeria. *Third World Quarterly,* 19(3), pp. 457-478.

Gauthier, J. M., 2014. *Geopolitics of natural resources- oil pillars of power. Lecture in Geopolitics and natural resources at HEC Paris.* Paris, 20 October 2014.

Gauthier, J. M., 2015. *The fundamentals of the oil market. Lecture in Certificate for energy and finance at HEC Paris*. Paris, 10 May 2015.

Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F. & Shleifer, A., 2004. Do institutions cause growth?. *Journal of Economic Growth,* 9(3), pp. 271-303.

Haber, S. & Menaldo, V., 2011. Do natural resources fuel authoritarinism? A reappraisal of the resource curse. *The American Political Science Review*, Issue 105, pp. 1-26.

Jacoby, J. G, 2005. *Outliers and Influential data.* Lecture notes from Michigan State University [Online]. http://polisci.msu.edu/jacoby/icpsr/regress3/lectures/week3/11.Outliers.pdf
[Accessed 24 September 2015].

Jerven, M., 2013. *Poor numbers: how we are misled by African development statistics and what to do about it.* s.l.:Cornell University Press.

Johnston, D., 2007. How to evaluate the fiscal terms of oil contracts. In: M. Humphreys, J. D. Sachs & J. E. Stiglitz, eds. *Escaping the Resource Curse.* New York: Columbia University Press, pp. 53-88.

Jojarth, C., 2008. The end of easy oil: estimating average production costs for oil fields around the world. *Center on Democracy, Development, and The Rule of Law-Stanford, published by Program on Energy and Sustainable Development Working Paper*, *72*.

Klausen, L. M., Bøe, M. & Gausel, I. J. H., 2014. *Interview with country risk analysts in Statoil* [Interview] (08 August 2014).

Lasserre, P. (1985). Capacity choice by mines. *Canadian Journal of Economics*, 831-842.

Leite, C. & Weidmann, J., 1999. Does mother nature corrupt? Natural resources, corruption and economic growth. *Working paper of the International Monetary Fund*.

Long, N. V., 1975. Resource extraction under the uncertainty about possible nationalization. *Journal of Economic Theory,* 10(1), pp. 45-53.

Looney, R., 2014. *Can Ghana's democracy save it from the oil curse? Foreign Policy.* [Online] Available at: http://foreignpolicy.com/2014/05/01/can-ghanas-democracy-save-it-from-the-oil-curse/
[Accessed 10 July 2015].

M. Marshall, T. G. K. J., 2014. *Systemicpeace.org.* [Online]
Available at: http://www.systemicpeace.org/inscr/p4manualv2013.pdf
[Accessed 24 June 2015].

Manzano, O. & Monaldi, F., 2008. The political economy of oil production in Latin America. *Economia,* 9(1), pp. 59-103.

Mehlum, H., Moene, K. & Torvik, R., 2006. Institutions and the resource curse. *The Economic Journal,* 116(1), pp. 1-20.

Murray, T. H. J., 1990. North American Drilling Activity in 1989. *World Energy Developments*, October, 74(10B), pp. 7-36.

Norman, C., 2009. Rule of law and the resource curse: abundance versus intensity. *European Association of Environmental and Resource Economists*, pp. 183-207.

Nøstbakken, L., 2013. *Oil and oil markets. Lecture in Petroleum economics at NHH.* Bergen, 26 August 2013.

Radon, J., 2007. How to negotiate an oil agreement. In: M. Humphreys, J. D. Sachs & J. E. Stiglitz, eds. *Escaping the Resource Curse.* New York: Columbia University Press, pp. 89-113.

Reno, W., 2001. External relations of weak states and stateless regions in Africa. In: *African foreign policies: Power and process.* London: Lynne Rienner Publishers, pp. 185-205.

Reuters, 2011. *reuters.com.* [Online]
Available at: http://www.reuters.com/article/venezuela-nationalizations-idUSN1E79I0Z520111201
[Accessed 3 December 2015].

Ross, M. L., 1999. Review: The political economy of the resource curse. *World Politics,* 51(2), pp. 297-322.

Sachs, J. D. & Warner, A. M., 1995. Natural resource abundance and economic growth. *National Bureau of Economic Research*.

Staalesen, A., 2011. *Rosneft picks ExxonMobil in Arctic deal.* [Online]
Available at: http://barentsobserver.com/en/energy/rosneft-picks-exxonmobil-arctic-deal
[Accessed 15 December 2015].

Statoil, 2009. *statoil.com.* [Online]
Available at:
http://www.statoil.com/no/OurOperations/ExplorationProd/internationalFields/Angola/Pages/OurBusiness.aspx
[Accessed 27 August 2015].

Stroebel, J. & van Benthem, A., 2013. Resource extraction contracts under threat of expropriation: Theory and Evidence.. *The Review of Economics and Statistics,* 95(5), pp. 1622-1639.

The Economist, 2013. *Supermajordämmerung.* [Online]
Available at: http://www.economist.com/news/briefing/21582522-day-huge-integrated-international-oil-company-drawing
[Accessed 19 December 2015].

Toft, P. & Duero, A., 2011. Reliable in the long run? Petroleum policy and long-term oil supplier reliability. *Energy Policy,* 39(10), pp. 6583-6594.

Treisman, D., 2007. What have we learned about the causes of corruption from ten years of crossnational empirical research?. *Annual Review of Political Science,* Volume Vol. 10, pp. 211-244.

Tsui, K., 2011. More oil, less democracy: Evidence from worldwide crude oil discoveries.. *The Economic Journal,* 121(March), p. 89–115.

US Energy Information Administration, 2015. *eia.gov*. [Online]
Available at: http://www.eia.gov/tools/glossary/index.cfm?id=P
[Accessed 20 July 2015].

Vaughan, A., 2014. *Ecuador signs permits for oil drilling in Amazon's Yasuni national park.* [Online]
Available at: http://www.theguardian.com/environment/2014/may/23/ecuador-amazon-yasuni-national-park-oil-drill
[Accessed 18 December 2015].

Woolridge, J. M., 2009. *Introductory Econometrics - A Modern Approach.* 4e ed. Mason: South-Western Cengage Learning

9.1 Data sets

NFW wells drilled and country risk:

IHS, Petroleum Economics and Policy Solutions (2014), *Petroleum country risk* [data file]. Retrieved from www.ihs.com through Statoil 28.07.2014

IHS, Petroleum Economics and Policy Solutions (2014), *Exploration activity* [data file]. Retrieved from www.ihs.com through Statoil 25.07.2014

Drilldays and startyear:

Wood Mackenzie (2014), *Well detail* [data file]. Retrieved from www.woodmac.com through Statoil 21.07.2014

Reserves:

US Energy Information Administration (EIA), International energy statistics (2015), *Crude oil proved reserves (billion barrels)* [data file]. Retrieved from http://www.eia.gov/cfapps/ipdbproject/iedindex3.cfm?tid=5&pid=57&aid=6 22.06.2015

Polity IV

Polity IV Project, Political Regime Characteristics and Transitions, 1800-2013 (2014), *Polity IV Annual Time-Series, 1800-2013* [data set]. Retrieved from http://www.systemicpeace.org/inscrdata.html 08.05.2014

Area

World Bank, World Development Indicators (2014), *Land area (sq. km)* [data set]. Retrieved from http://data.worldbank.org/indicator/AG.LND.TOTL.K2 05.05.2014

Distance to equator

Davis, Lewis and Hopkins, Mark "Institutional foundations and economic growth" (2008). *Latitude and languages* [data set]. Retrieved from http://www.macrorules.com/mhopkins/datasets.html 07.05.2014

Ethnolinguistic Fractionalization

Treisman, Daniel "What have we learned about the causes of corruption from ten years of crossnational empirical research" (2006) dataset, originates from Roeder, Philip (2001*)* *Ethnolinguistic fractionalization in 1961 and 1985* [data set]. Retrieved from http://www.sscnet.ucla.edu/polisci/faculty/treisman/Pages/publishedpapers.html 09.05.2014

Landlocked, geological vector, regions and income group

Gallup, John L. and Jeffrey D. Sachs, with Andrew Mellinger, "Geography and Economic Development" (1999), *Geodata* and *Physical Factors* [data set]. Retrieved from http://www.cid.harvard.edu/ciddata/ciddata.html 06.05.2014

OPEC (2014), member countries [data set] retrieved from http://www.opec.org/opec_web/en/about_us/25.htm 12.05.2014

Oil price development 1960-2014 (in chapter 1)

The World Bank (2015) [data set] retrieved from http://www.worldbank.org/en/research/commodity-markets 01.12.2015

and EIA, 2015. *Short-term energy outlook*. [Online]
Available at: https://www.eia.gov/forecasts/steo/report/prices.cfm
[Accessed 12 December 2015].

10. Appendix

10.1 Countries used in the main regressions

Afghanistan
Albania
Algeria
Argentina
Australia
Austria
Benin
Bolivia
Brazil
Bulgaria
Cambodia
Canada
Chile
China
Colombia
Congo, Dem. Rep.
Congo, Rep.
Cote d'Ivoire
Cuba
Denmark
Ecuador
Egypt
Ethiopia
France
Gabon
Germany
Ghana
Greece
Guatemala
India
Indonesia
Iran
Iraq
Ireland
Israel
Italy
Japan
Jordan
Kenya
Korea, Rep.
Kuwait
Libya
Madagascar
Malaysia
Mauritania
Mexico
Mongolia
Morocco
Myanmar
Netherlands
New Zealand
Nicaragua
Niger
Nigeria
Norway
Oman
Pakistan
Peru
Philippines
Poland
Portugal
Romania
Saudi
Arabia
Senegal
Sierra
Leone
South Africa
Spain
Sudan
Sweden
Syria
Taiwan
Tanzania
Thailand
Togo
Trinidad and Tobago
Tunisia
Turkey
UK
USA
Uganda
Venezuela
Vietnam
Yemen

10.2 Summary from the dataset of the dependent variables and Polity IV

Variable	Number of observations	Mean	Standard deviation	Minimum value	Maximum value
NFW wells drilled 1970-2010	86	2 689	18 297	1	169 814
Log (NFW wells drilled 1970-2010)	86	5.08	2.21	0	12.04
Drilldays 1970-2010	86	71 666	104 402	108	417 234
Log (Drilldays 1970-2010)	86	9.83	2.01	4,68	12.94
Success rate 1970-2010	86	0.43	0.22	0.13	1
Polity IV 1960-69	86	0.43	0.36	0	1
Polity IV 1970-2010	86	0.55	0.31	0	1

10.3 Estimation of vectors used as geological variables

The variables were first added separately to the regressions, but none proved to have a significant impact.

Table 10: Regression results from estimating the geological vector

	(1) NFW drilled	(2) Drilldays	(3) NFW Successrate
Soil suitability 1 (% of land area)	0.014	-0.007	0.001
	(0.027)	(0.028)	(0.004)
Soil suitability 2 (% of land area)	-0.032	-0.041*	-0.003
	(0.026)	(0.024)	(0.004)
Boreal regions (% of land area)	-2.457	-3.326	0.912
	(5.698)	(7.276)	(0.782)
Temperate desert (+ polar + boreal)	0.314	3.002	0.454
	(4.456)	(5.761)	(0.554)
Tropical and subtropical desert	-1.606	2.810	0.968
	(4.743)	(5.896)	(0.664)
Dry temperate (% of land area)	-3.164	2.615	1.130
	(5.186)	(6.357)	(0.720)
Wet temperate (% of land area)	-3.355	2.086	1.180
	(5.313)	(6.552)	(0.726)
Subtropics (% of land area)	-1.876	3.119	1.024
	(4.885)	(5.836)	(0.685)
Tropics (% of land area)	-2.108	2.449	1.159*
	(4.729)	(5.743)	(0.649)
Polar non-desert (% of land area)	-5.232	14.873	2.252***
	(6.786)	(10.571)	(0.837)
Other exogenous control variables	Yes	Yes	Yes
Constant	3.818	7.843	-0.126
	(5.058)	(6.037)	(0.705)
Observations	86	86	86
R^2	0.742	0.708	0.471
Adjusted R^2	0.646	0.600	0.275

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Other exogenous control variables refer to distance from equator, area, number of years with drilling activity, ethnic fractionalization, landlocked, OPEC member and regional variables. All polity variables have been excluded from the dataset. The vector has been instrumented with the weights of the coefficients reported in each case. Only the variables who has been listed are included in the vector

10.4 Excluding outliers

Table 11: Regression results when excluding outliers

Dependent variable	NFW wells 1970-2010			Drilldays 1970-2010			Successrate NFW wells 1970-2010		
	(1) OLS ordinary	(2) OLS reduced form	(3) 2SLS	(4) OLS ordinary	(5) OLS reduced form	(6) 2SLS	(7) OLS ordinary	(8) OLS reduced form	(9) 2SLS
Polity IV score 1970-2010	1.222*		2.472***	1.756**		2.930***	-0.200*		-0.180
	(0.687)		(0.855)	(0.676)		(0.822)	(0.111)		(0.128)
Polity IV score 1960-69		1.234***			1.472***			-0.098	
		(0.445)			(0.467)			(0.083)	
Distance from equator	-0.031**	-0.030**	-0.039***	-0.036**	-0.031*	-0.043***	-0.001	-0.002	-0.001
	(0.014)	(0.015)	(0.014)	(0.017)	(0.016)	(0.015)	(0.002)	(0.002)	(0.002)
Area	0.732***	0.724***	0.759***	0.467***	0.451***	0.488***	-0.060***	-0.057***	-0.059***
	(0.084)	(0.086)	(0.084)	(0.083)	(0.086)	(0.070)	(0.013)	(0.013)	(0.012)
Years with drilling activity	-0.017***	-0.016***	-0.017***	-0.021***	-0.019***	-0.021***	0.001**	0.001**	0.001***
	(0.005)	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)	(0.000)	(0.000)	(0.000)
Ethnolinguistic fractionalization	-0.972	-0.970	-1.174*	-0.310	-0.337	-0.427	0.104	0.075	0.101
	(0.677)	(0.668)	(0.625)	(0.585)	(0.563)	(0.545)	(0.114)	(0.119)	(0.101)
Landlocked	-0.870*	-0.708	-0.817*	-0.701	-0.529	-0.628	0.117	0.111	0.118
	(0.483)	(0.460)	(0.419)	(0.435)	(0.448)	(0.392)	(0.081)	(0.081)	(0.073)
Member of OPEC	0.405	0.314	0.494	0.712	0.616	0.779	0.108	0.128	0.110
	(0.560)	(0.532)	(0.517)	(0.507)	(0.469)	(0.474)	(0.079)	(0.084)	(0.074)
Regional control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geological control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.693	-0.309	-1.425	-2.346	-2.539	-3.884	-0.155	-0.039	-0.155
	(1.672)	(1.606)	(1.585)	(3.793)	(3.649)	(3.409)	(0.427)	(0.430)	(0.382)
Observations	81	81	81	81	81	81	81	81	81
R^2	0.720	0.732	0.706	0.737	0.742	0.723	0.505	0.486	0.504
Adjusted R^2	0.661	0.675	0.644	0.681	0.687	0.664	0.400	0.377	0.399
F value (instrument)			39.17			35.77			43.47

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Excluding observations from Afghanistan, Canada, India, USA and Saudi Arabia based on detection of outliers.

10.5 Results from regressions with values of zero included.

All regressions were run with all the control variables as listed in equation (1), (2) and (3).

Table 12: Regressions with NFW wells extended sample

	Dependent variable: NFW wells drilled 1970-2010				
	(1) OLS simple regression	(2) OLS with political control variables	(3) OLS with geological control variables	(4) OLS reduced form	(5) 2SLS second stage
Polity IV score 1970-2010	0.774	0.224	0.736		0.956
	(0.571)	(0.766)	(0.788)		(1.062)
Polity IV score 1960-69				0.468	
				(0.573)	
Observations	96	96	96	96	96
R^2	0.599	0.704	0.719	0.718	0.718
Adjusted R^2	0.581	0.653	0.666	0.665	0.666

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Countries included in new sample: Costa Rica, Gambia, Guinea, Honduras, Liberia, Mali, Panama, Somalia, Sri Lanka and Uruguay.

Table 13: Regressions with drilldays extended sample

	Dependent variable: Drillays 1970-2010				
	(1) OLS simple regression	(2) OLS with political control variables	(3) OLS with geological control variables	(4) OLS reduced form	(5) 2SLS second stage
Polity IV score 1970-2010	0.457	0.979	1.445**		1.872**
	(0.558)	(0.721)	(0.685)		(0.942)
Polity IV score 1960-69				0.923*	
				(0.536)	
Observations	96	96	96	96	96
R^2	0.579	0.681	0.719	0.714	0.717
Adjusted R^2	0.560	0.625	0.666	0.661	0.664
F value (instrument)					44.36

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Countries included in new sample: Costa Rica, Gambia, Guinea, Honduras, Liberia, Mali, Panama, Somalia, Sri Lanka and Uruguay.

Table 14: Regressions with success rate extended sample

	Dependent variable: Success rate				
	(1) OLS simple regression	(2) OLS with political control variables	(3) OLS with geological control variables	(4) OLS reduced form	(5) 2SLS second stage
Polity IV score 1970-2010	-0.281***	-0.209	-0.257**		-0.409***
	(0.076)	(0.129)	(0.115)		(0.153)
Polity IV score 1960-69				-0.208**	
				(0.085)	
Observations	96	96	96	96	96
R^2	0.138	0.219	0.256	0.266	0.241
Adjusted R^2	0.100	0.084	0.116	0.128	0.099
F value (instrument)					54.11

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Countries included in new sample: Costa Rica, Gambia, Guinea, Honduras, Liberia, Mali, Panama, Somalia, Sri Lanka and Uruguay.