DEPARTMENT OF ECONOMICS
OxCarre
Oxford Centre for the Analysis of Resource Rich Economies

Manor Road Building, Manor Road, Oxford OX1 3UQ
Tel: +44(0)1865 281281 Fax: +44(0)1865 271094
oxcarre@economics.ox.ac.uk www.oxcarre.ox.ac.uk

UNIVERSITY OF OXFORD

OxCarre Research Paper 127

_

Institutions and the Location of Oil Exploration

James Cust*

World Bank

&

Torfinn Harding*

NHH Norwegian School of Economics

***OxCarre External Research Associate**

Direct tel: +44(0) 1865 281281 E-mail: oxcarre@economics.ox.ac.uk

Institutions and the Location of Oil Exploration

James Cust and Torfinn Harding*

February 6, 2017

Abstract

We provide evidence that institutions strongly influence where oil and gas exploration takes place. To identify the effect of institutions, we utilise a global dataset on the location of exploration wells and national borders. This allows for a regression discontinuity design, with the key assumption that the position of borders was determined independently of geology. To break potential simultaneity between borders, institutions and activities in the oil sector, we exploit the historical sequence of drilling occurring after the formation of borders and institutions. At borders, exploration companies choose to drill on the side with better institutional quality 58% of the time. The results are consistent with the view that institutions shape exploration companies' incentives to invest in drilling as well as host countries' supply of drilling opportunities. It follows that the observed distribution of natural capital across countries is endogenous with respect to institutions.

Keywords: institutions, investment, oil and gas exploration, regression discontinuity design

JEL-codes: F21, O13, O43, Q32

*We would like to thank four anonymous referees, Jørgen Juel Andersen, Roland Hodler, Guy Michaels, Magne Mogstad, Klaus Mohn, Linda Nøstbakken, Rick van der Ploeg, Simon Quinn, Michael Ross, Shqiponja Telhaj, Ragnar Torvik, Tony Venables and seminar participants at NHH Bergen, University of Bergen, EUDN Berlin, RES Manchester, EGIT Munich, Norwegian Economist Meeting Oslo, ESOP-OxCarre workshop Oxford, University of Oxford, NRG Siegen, University of Stavanger, University of Stirling, NOITS Stockholm, and University of Sussex for useful comments. The authors would also like to thank Wood Mackenzie for academic access to their PathFinder database. Support from the BP funded Centre for the Analysis of Resource Rich Economies (OxCarre) at the University of Oxford and the Statoil chair in economics at the Norwegian School of Economics is also appreciated. Authors: James Cust, World Bank, jimcust@gmail; Torfinn Harding, NHH Norwegian School of Economics, torfinnh@gmail.com.

1 Introduction

Institutions, generally defined as 'the rules of the game in a society' (North, 1990), are widely considered to be a fundamental cause of economic growth. Aspects such as constraints on the executive or the rule of law shape the incentives to invest and therefore growth trajectories (Acemoglu et al., 2005). However, identifying the causal effect of institutions is challenging because of correlated factors and because institutional characteristics are themselves endogenous equilibrium outcomes (Acemoglu et al. 2005, Besley and Persson 2010).

This paper uses the setting of oil and gas exploration to estimate the effect of institutional quality on investment.[1] Exploration and subsequent extraction requires large up-front capital outlays. A potential investor will take into account the probability of discovery and the net present value of a discovered barrel of oil. The former depends upon geology, the latter upon operational costs and risks, which are again partly determined by institutions. The 'supply side', i.e. how potential host countries facilitate drilling, for example through the licensing process and tax deductions for exploration costs, is also plausibly affected by institutional quality. For these reasons, oil exploration is expected to vary with institutional quality for a given set of geological conditions. Aggregate figures show that almost 90% of oil investments have historically taken place in upper-middle and high income countries. OECD-countries have discovered about five times more subsoil natural resources per square km than those in Africa.[2] This paper argues that variation in institutional quality between countries can contribute to understanding the uneven distribution of investments and hence the distribution of known sub-soil wealth between countries.

The ideal experiment in our setting would be to treat a given geology with different institutional quality and observe the response in drilling. We mimic such an experiment by implementing a Regression Discontinuity (RD) design. A new global data set including the geo-coded location of oil exploration wells and national borders allows us to examine investments

[1]We focus in this paper on oil and gas exploration, measured as the number of oil and gas wells drilled. We use 'oil' as a short hand expression for 'oil and gas'.

[2]Collier (2010) estimates that OECD countries have discovered 130 thousand USD worth of subsoil natural resources per square km compared to 25 thousand for those in Africa. See McKinsey Global Institute (2013) for the investment figure.

in areas proximate to national borders. We consider borders which have not been changed since 1965 and only drilling that took place from 1966. If these borders were determined independently of subsoil oil, a discontinuity in the number of wells at national borders can be interpreted as the causal effect of the border on oil exploration.

We follow the literature on the effects of institutions on economic growth and apply broad measures of institutions: political rights, democracy, autocracy and constraints on the executive branch of government. Previous research has also suggested that these dimensions are important for the likelihood of expropriation of foreign assets (Li, 2009).

Like Michalopoulos and Papaioannou (2013), we place the country with the better score for institutional quality on the 'right hand side' of the border. A border-dummy, taking one on the right hand side and zero on the left hand side, is the explanatory variable of interest in the regressions. The institutional variation is bilateral and strictly ordinal, as we merely use the institutional measure to rank the two countries. Simultaneity between activities in the oil sector and institutional quality would induce a bias in our estimates if it shifted the ranking of the two countries. Institutional quality is therefore measured in 1965 or before, prior to the exploration drilling we consider.[3] To estimate the effect of a given change in institutional quality, we use measures of institutional quality as the treatment variable and the border-dummy as the instrument in a two-stage-least-squares estimation.

Our baseline estimates suggest a jump at the border of 36%, from from 4.1 to 5.6 wells on average. This means that, ceteris paribus, exploration companies choose the better side of the border 58% of the time. Scaled with measures of institutional quality, the estimates imply that a one standard deviation increase in the augmented Freedom House Political Rights Index increases the number of wells by 39%. These are close to long-run estimates, since they are based on the accumulated number of wells drilled in the period 1966-2010.

A rough back-of-the-envelope calculation suggests that the average exploration well in our sample results in production of about 2,500 barrels per day. Based on the 2,100 wells we

[3]The issue that previous or current oil activities could affect the institutional quality is emphasised in the literature on the natural resource curse (van der Ploeg, 2011). We note that most of the literature has found a negative correlation between natural resource extraction and institutional quality and such simultaneity-bias would therefore induce a bias towards zero in our estimates.

observe within 10 km of the left hand side of the border, our estimates imply that transplanting the institutions from the right hand side to the left hand side would increase the daily production by an amount that compares to 2% of the global oil production in 2013. The corresponding figure for a distance of 250 km from the borders is 30% of the global oil production in 2013.

Allowing for heterogeneity across companies, drilling by the so called six supermajors of the oil industry (Chevron, Shell, BP, ExxonMobil, ConocoPhillips and Total) is found to be more sensitive to institutional quality than nationally-owned oil companies (such as CNPC, PDVSA, Petrobras or Petronas), as well as the rest of the oil exploration industry. Splitting the sample in terms of income per capita, we find that the effect in general holds, except for democracy and autocracy measures close to borders in the high income countries.

Given that the fundamental causes of economic growth may be classified as 'institutions', 'geography' and 'culture' (see Acemoglu et al. 2005), our contribution is to show that institutions have a positive effect on oil exploration, keeping geography and culture fixed. The effect of institutions itself may be sliced up in three different ways. First, inspired by Acemoglu and Johnson (2005) who unbundle institutions into contracting institutions and property rights institutions, we show that our effect at least partly works via property rights institutions, although it does not seem to be the only relevant aspect. Second, as a fundamental cause of economic growth, institutions operate through a variety of proximate causes such as investments in physical and human capital. We show that controlling for the initial states of such proximate causes reduces but does not take away the estimated effect of institutions. Third, institutions may affect both demand for drilling by the exploration companies and supply of drilling opportunities by the host countries. Disinvestments from Iran by international investors during the 1990s and early 2000s, due to operating terms and escalating political tension, illustrates the demand side. Mexico banned oil drilling by foreign companies between 1938 and 2013 and serves as an illustration of restricted supply. This may have contributed to the sharp asymmetry in drilling across the US-Mexico border in the Gulf of Mexico.[4] In our setting, we observe only the equi-

[4]The Iran example is from the Financial Times 2 November 2016: https://www.ft.com/content/06acb822-95fe-11e6-a80e-bcd69f323a8b. For news coverage of the investment law in Mexico, see the article from BBC 21 December 2013: http://www.bbc.com/news/world-latin-america-25471212. Figure B.1 shows the exploration wells on each side of the US-Mexico border in the Gulf of Mexico.

librium outcome and lack of excludable instruments for demand and supply prevents us from identifying their effects separately.

Our estimates are based on oil exploration close to national borders. There may be a race to drill first at the border, as oil deposits may straddle the border. Although consistent with the hypothesis of this paper, and not posing a threat to the internal validity of our estimates, this would generate an estimate for the effect of institutions that is higher than for oil exploration in general. Empirically we cannot rule out such a competition effect, but show in section 3.6 that this is unlikely to be driving our qualitative conclusion.

To scrutinize the key identifying assumption that national borders are randomly assigned with respect to the underlying geology, we perform a set of tests. We show that the presence of geological basins, where oil can be found, is the same on each side of our borders. Conditional on drilling, there is also no discontinuity at borders in discovery rates. Furthermore, our results are robust to limiting the sample to include "older" borders only (1925, 1938 or 1945), to African countries only, and to so-called 'artificial' states only. The same is true when we exclude offshore borders or borders with interstate conflicts. Finally, the results are robust to the inclusion of previous drilling, previous discoveries or country fixed effects as controls.

The starting point of this paper is the literature on institutions and economic growth (Acemoglu et al. 2005, Acemoglu et al. 2001, Acemoglu and Johnson 2005, Acemoglu et al. 2014, Dell 2010, Hall and Jones 1999, Michalopoulos and Papaioannou 2013, Papaioannou and Siourounis 2008 and Rodrik and Wacziarg 2005). The paper's contribution lies in the identification of investments as a channel through which institutions such as democracy affect growth. We measure investments as the number of exploration wells drilled. The setting of oil exploration lends itself particulary well to a spatial regression discontinuity design, such as the one used by Michalopoulos and Papaioannou (2013). First, we observe the exact geographic location and timing of each exploration well providing us with rich variation in our outcome of interest. Second, the subsoil nature of oil reservoirs together with the timing of borders make the location of borders plausibly exogenous with respect to oil. The borders we study were typically established before oil had high economic value and before its likely locations were known. In contrast, application of other measures of economic activity used in the literature, such as nighttime lights, is more vulnerable to the possibility that the location of people, firms and political borders are simultaneously determined.[5] Our finding suggests that better scores on measures of democracy, autocracy and constraints on the executive branch of government do increase investments in oil exploration. This is in agreement with Acemoglu et al. (2014), who find a positive effect of democracy on GDP per capita with investments as a channel. They are also in line with Acemoglu and Johnson (2005), who find positive effects of property rights institutions, such as constraints on the executive, on GDP per capita and investments. In addition, we present evidence indicating that our findings are not only about foreign direct investments. Both domestic and international companies are responsive to the quality of institutions, although the large international oil companies are more sensitive than the other companies.

In the literature on the use of natural resources, Bohn and Deacon (2000) study the effect of expropriation risk on investments in oil exploration and production. Theoretically, they find a negative effect under plausible assumptions, and confirm this empirically in a cross-country analysis.[6] Our paper improves on their empirical analysis by using micro data on oil exploration, an identification strategy utilising plausibly exogenous variation in institutional quality and more general measures of institutions.[7] Our findings are in line with their findings.

Our finding means that some types of natural capital are economic outcomes. Like human and physical capital, their accumulation (discovery) depends on institutions. This has immediate consequences for the understanding of the determinants of 'Wealth of Nations', as

[5]Related issues are discussed by Alesina et al. (2011), Michalopoulos and Papaioannou (2013), Michalopoulos and Papaioannou (2014) and Turner et al. (2014).

[6]This finding stands in contrast to the earlier conventional view that weak property rights increase current extraction in expectation of higher expropriation risk and hence lower marginal returns in the future (Long, 1975).

[7]The natural resource literature on how resource use varies across levels of development has been concerned with the common pool aspect of natural resources, see for example Bohn and Deacon (2000), Jacoby et al. (2002), Long (1975) and Laurent-Lucchetti and Santugini (2012). A related literature has focused on foreign direct investment under the risk of expropriation, see for example Thomas and Worrall (1994) and Janeba (2002). The literature on oil exploration more generally has either been theoretical or focusing on mature economies like US, UK, Canada and Norway, and has in both strands been less concerned with the institutional setting. See for example Hendricks and Porter (1996), Hurn and Wright (1994), Livernois and Ryan (1989), Mohn and Osmundsen (2008), Pesaran (1990), Quyen (1991) and Venables (2011).

measured by the World Bank (2011). It also has implications for the literature on the natural resource curse (see summary by van der Ploeg 2011). Oil reserves and production must be treated as endogenous in empirical models of the effects of oil. This questions identification strategies that assume oil discoveries and production to be driven by geology, without taking into account that they are themselves outcomes of institutions and associated economic developments.[8]

The idea that oil exploration, and hence the discovery of sub-soil wealth of countries, depends on the institutional environment may not be accepted by policy makers everywhere. Countries like Iraq, Nigeria, Venezuela and Angola, who rank among the top twenty largest oil producers in the world in 2012 (EIA, 2013), score relatively low on measures of institutional quality. Promising geology may therefore be thought of as a sufficient condition for oil exploration. In contrast, our results suggest that countries with identical geology but a different institutional setting will fare very differently. To the extent countries can act to improve their institutional environment, they may be in a position to affect the pattern of oil exploration and discovery. We take this to be an important policy message of this paper.

The paper proceeds as follows. The next section describes the data and outlines the identification strategy. Section 3 presents the main results together with sub-sections dealing with the roles of ethnicity and natural geography, the role of property rights institutions versus other aspects of the institutional setting, and the role of a potential common pool effect at the border. Section 4 investigates in detail potential threats to identification. Section 5 examines heterogeneity across different company types as well as between developing countries and high income countries. Section 6 offers concluding remarks.

[8]We discuss the implications of endogenous institutions with respect to oil in section 4. See Andersen and Aslaksen (2008), Andersen and Ross (2014), Brunnschweiler and Bulte (2008), Haber and Menaldo (2011), Jensen and Wantchekon (2004), Mehlum et al. (2006), van der Ploeg and Poelhekke (2010), Ross (2001) and Tsui (2011) for papers dealing with the resource curse and democracy/institutions. Tsui (2011) and Cotet and Tsui (2013) aim for tackling the endogeneity of oil wealth by using initial resource wealth or discovery rates as instruments and Cassidy (2016) by using information about geological basins.

2 Data and empirical strategy

2.1 Data

Oil exploration data are provided by the PathFinder database owned by Wood Mackenzie (2011) which covers more than 100,000 individual wells in over 120 countries. This proprietary database approaches comprehensive global coverage and is, to our knowledge, the largest collection of the world's exploration wells. It includes a wide range of country and operator reported information assembled historically and updated on a quarterly basis. We utilise information on the exact location of each well, the year when drilling started, and whether the well turned out to be 'dry' or resulted in a discovery.[9] Figure 1 shows the global distribution of wells available.[10]

Data for onshore national borders are from the GADM database of Global Administrative Areas version 2.0 (Hijmans et al., 2010) and for offshore maritime borders the EEZ Maritime Boundaries Geodatabase version 6.1 (Claus et al., 2013).[11] We only include borders which have not changed since 1965, according to Weidmann et al. (2010).[12]

We define our unit of observation as bins of 0.1 km width, stacked from a given national border. We use the distance from the middle of that bin as the distance to the border. We calculate the distance from each well to the closest national border and sum the number of wells drilled in each bin. This sum of wells is our dependent variable. Bins without any wells

[9] We define dry wells (versus those recorded as a discovery) using the industry standard definition: a dry well or dry hole in our data is denoted as such where evaluated to contain insufficient oil for commercial production. Wood Mackenzie analysis supplements the raw well data to include post-drilling well evaluation, allowing for a more accurate picture of the eventual result and assessment of the well - and not just the recorded result at the time of drilling.

[10] We focus in this paper on oil and gas exploration. For solid minerals and metals such as coal, iron ore, gold and copper, we do not possess exploration information equivalent to exploration wells. To run a similar study for such commodities, we would have to rely on data on actual mines, i.e. projects that have moved into the production phase. Furthermore, many of these commodities have held value far back into human history and geologically they are often located closer to the surface than oil and gas. The presence of such commodities may therefore contaminate the location of borders.

[11] All geographic data, including distance and area calculations uses the Cylindrical Equal Area projection. This projection has the advantage of minimal distortion for global projections, while preserving consistent area. We calculate straight-line distances from individual exploration wells to nearest borders and nearest neighbours using the 'near' point-to-line function of ArcGIS.

[12] See the online appendix B.2 for additional discussion of our borders data.

will not be represented in our main specifications.

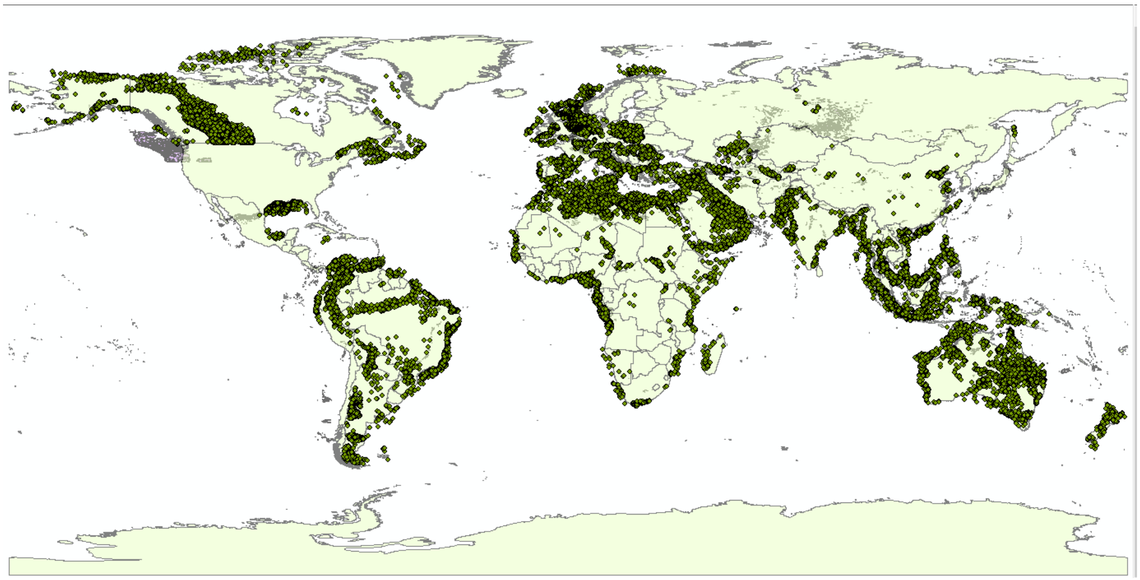

Figure 1: National borders and exploration wells

Our baseline measure of institutional quality, denoted *Freedom House or* FH, is the augmented Freedom House variable from Acemoglu et al. (2008) based on the Freedom House Political Rights Index.[13] As normalised by the authors, the variable is measured on a zero to one scale, with closer to one indicating closer to an ideal set of democratic institutions such as free and fair elections and the presence of electoral competition. For alternative measures of institutional quality we turn to the widely used Polity IV database.[14] *Democ* and *Autoc* are measured on a 0-10 scale, with higher score indicating more democratic or more autocratic institutions, respectively. A composite Polity score is given by *Polity* = *Democ* - *Autoc*, measured on a -10 to 10 scale. *ConEx* captures institutionalized constraints on the Executive and is measured on a 1-7 scale, with higher score corresponding to more constraints.[15] Country-level

[13]The data are available from 1950 and can be found at: https://www.aeaweb.org/articles.php?doi=10.1257/aer.98.3.808

[14]The data start in 1800 and are taken from the 'Polity IV: Regime Authority Characteristics and Transitions Datasets'. They can be downloaded from: http://www.systemicpeace.org/inscr/p4v2012.xls. We follow Acemoglu and Johnson (2005) and treat 'interregnums' as missing values for all the Polity IV variables.

[15]We measure institutions at the country-level and assume that the measures are representative for the institutional quality also close to borders. The issue of dual-institutions pointed to by Michalopoulos and Papaioannou (2013), where the institutions of the 'modern sector' do not penetrate fully to the "countryside' and hence potentially weakens the discontinuity in institutions at the border, is unlikely to be an issue in our setting. Investments in the oil sector are potentially large contributions to the domestic capital formation and likely to attract the attention of the central authorities; laws, policies and procedures relating to the oil sector are likely to be centralised and commonly applied across the country (Daniel et al., 2010).

data on country size and landlocked status are from CEPII.[16]

We use all countries available, except in section 5.2, where we split the sample into developing versus high income countries.[17] We exclude all formally 'disputed areas' in the offshore data, given the unclear and contested ownership of such areas.[18] Table A.1 presents descriptive statistics and table B.7 presents the list of included host and neighbouring countries.

2.2 Empirical strategy

Our conjecture is that institutional quality affects expected profits in the oil sector and an investment location decision will therefore be determined by institutions as well as geology. Three particular sources of endogeneity may induce bias when estimating the effect of institutional quality on oil exploration. The first is that the geology (the actual likelihood of discovering oil) is not observable to us. The second is that institutional quality is itself an outcome that can be affected by the investment activities under study. The third is that the territory of countries may be correlated with both geology and institutional quality. We seek to overcome these three identification challenges by taking advantage of the location of oil exploration relative to national borders and by utilising the particular sequencing of events in our setting.

Data on the precise geographical location of oil wells and national borders allow for a regression discontinuity design, as potential geological deposits of oil often stretch across national borders.[19] By studying oil exploration very close to the border, we effectively control for geology. To deal with the potential simultaneity between activities in the oil sector and the institutional quality, we measure institutional quality before the drilling we consider took place. Institutional quality is then predetermined. Finally, to make sure that the location of the national borders is not affected by the oil exploration we consider, we include only countries we know have not changed in shape during our period of oil exploration.

[16]The data can be found at: http://www.cepii.fr/anglaisgraph/bdd/distances.htm

[17]We use World Bank's definition of developing countries and geographical regions, as of July 2012. For more information, see: http://data.worldbank.org/news/new-country-classifications.

[18]Disputed areas are defined by the maritime border geodatabase Claus et al. (2013)

[19]See Imbens and Lemieux (2008), Imbens and Wooldridge (2009) and Lee and Lemieux (2010) for thorough explanations and discussions of regression discontinuity (RD) designs.

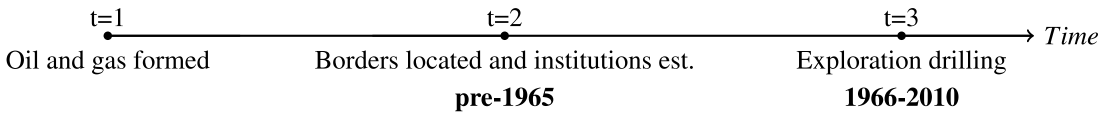

Figure 2: Sequence of events in our sample

Figure 2 details the sequence of events. The true distribution of oil and gas deposits is given from pre-historic geological processes, taking place in period 1. In period 2, we posit that national borders were located and institutions were formed. In period 3, exploration drilling for oil takes place. We choose to define period 2 as 1965 and earlier because Weidmann et al. (2010) provide data on countries that have not changed shapes since 1965.[20] We measure institutions as the country-average across all available observations up to, and included, 1965 (i.e. 1950-1965 for the augmented Freedom House measure and 1800-1965 for the Polity IV measures). Period 3 is defined as 1966-2010 and we include wells drilled in this period only.

We follow Michalopoulos and Papaioannou (2013) and place a given country on the left or right hand side of the border depending on its institutional quality relative to the neighbour. The better country in a country-pair is placed on the right hand side. Pooling observations on the left and right hand side, the effect of national borders on oil exploration is estimated by the following regression:

$$N_{bij} = \alpha_0 + \tau D_{ij} + f_r(X_{bij}) + D_{ij} f_r(X_{bij}) + Z_i' \delta_0 + u_{0,bij} \tag{1}$$

The unit of observation is a distance-bin b in country i close to the border with country j. A distance-bin is defined as a 100 meter wide strip following along the border.[21] N is a variable measuring the number of wells drilled in the bin. α is a constant and D is a dummy variable taking the value of one on the side of the border with the better institutional quality, and zero on the other side. The assignment variable, X, is the distance to the border, taking negative values on the left hand side and positive values on the right hand side. The f-function is used

[20] See section 2 for more information on the data from Weidmann et al. (2010). We provide sensitivity checks on this choice in section 4.1.

[21] The distance from the border is measured to the mid point of the distance-bin. For example, the tenth distance-bin is measured at 950 meters from the border and will include all wells drilled from 900 to 1000 meter from the border.

to pick up the underlying distribution of drilling with respect to the distance to the border and we allow it to differ on both sides of the border by interacting with the border dummy D. Z is a vector of controls for country i and u is the error term. τ picks up a jump in the mean at the discontinuity, i.e. the effect of crossing the national border on the number of exploration wells drilled. Note that institutional quality enters equation 1 only in an ordinal sense, i.e. by determining the ranking of the two countries. Simultaneity between the number of wells and the quality of institutions is therefore only an issue if the ranking is affected.

Control variables should not affect the estimate of τ in a correctly specified RD design, but their inclusion may increase efficiency (Imbens and Lemieux 2008, Lee and Lemieux 2010). For this reason we include the total area of country i, which may be correlated with the length of the border, and the landlocked status of country i, which may affect transport costs asymmetrically across borders.

To estimate how responsive oil exploration is to a given change in institutional quality, we scale the jump at the border with the level of institutional quality by applying standard two-stages least squares estimation, with the border dummy as instrument for the level of institutional quality.

The RD-assumption of equal geology on both side of the border is most likely to be valid close to the border and we focus on 10 km distance from the border. However, we also present estimates for 250 km from the border. We present estimates for different specifications of the f-function and use a second order polynomial as our preferred specification. The density of wells is expected to increase towards the border, as the density of land area increases towards the border (this is confirmed in figure B.2). There are two reasons for this. First, moving away from the border means moving towards the geographic center of a given country. Although the width of each distance-bin is constant, the length of each strip will necessarily decrease as one moves farther into a country (think of a square-shaped country and move towards the centre from all four sides at the same time). Since area size is a quadratic function of the distance to the centre of the country, the first derivative is a linear function of the distance to the centre. This simple geometric fact explains why the area size increases as one approaches the border. Second, all countries in the sample are represented close to the border, while only large countries are represented far away from the border.[22] This imposes curvature, i.e. the area size increases in a convex manner as one approaches the border. A final reason for the shape towards the border may be competition for oil close to the border, as discussed in section 3.6. Gelman and Imbens (2014) warn against high-ordered polynomials and our judgement is that a second order polynomial is appropriate in our setting.[23]

As we measure the distance from the border discretely (for each 100 meters), we follow Lee and Card (2008) and cluster standard errors on distance-bins in all regressions. We show that our results are robust to accounting for spatial correlation using the GMM-method of Conley (1999) as implemented by the program of Hsiang (2010).

3 Results

This section starts with a look at the raw data and a graphical presentation of the jump in exploration wells observed at national borders. We then present econometric estimates of the size of this jump. To get an estimate of the responsiveness of exploration drilling to a given change in institutional quality, we scale our econometric estimates with the level of institutional quality. We do this for several standard measures of broad institutional quality, which also serves to demonstrate that our finding is robust to alternative measures of institutional quality. Then follows a discussion of the role of geography and culture, the two competing fundamental causes of economic growth, as well as a discussion of unbundling the effect of our broad measures of institutions into different sub-aspects of institutions. The section ends with a comparison of the effects close to versus further away from the border.

[22]The distances we are studying in this paper, e.g. up to 250 km, may not seem very large in the context of the size of countries. However, each well is attached to its closest border, and out of a sample of 44,651 wells in developing countries, the median distance to the border is 53 km, while the 95th percentile is 187 km.

[23]In the previous version of this paper (Cust and Harding 2013) we include an extensive discussion of the sensitivity to both the distance from the border and the order of the polynomial for the sample of developing countries. We concluded that 10 km distance to the border in combination with second or third order polynomial was a reasonable compromise between bias and efficiency.

3.1 Graphical evidence of discontinuity at the border

Table 1: Number of wells, left and right

	LHS	RHS	Ratio	p-value
Distance from border →		10 km		
Likelihood of at least one well: D=1 if wells>0	0.08	0.12	1.42	0.00
Mean number of wells	2.55	3.02	1.18	0.00
Total number of wells	2 112	4 724	2.24	
Distance from border →		250 km		
Likelihood of at least one well: D=1 if wells>0	0.023	0.026	1.12	0.00
Mean number of wells	1.64	1.94	1.18	0.00
Total number of wells	16 752	25 435	1.52	

Notes: Based on 50 and 76 home and neighbouring countries for the 10 km sample and 70 and 96 for the 250 km. p-values for simple t-test of differences in means.

Table 1 presents the likelihood of drilling together with the mean and total number of wells on each side of the border. In a given 100 meter-wide distance-bin within 10 km of the border, there is a 8% likelihood that a well is drilled on the left hand side (worse institutional quality) compared to 12% on the right hand side (better institutional quality). The mean number of wells are 2.6 versus 3.0, i.e. 18% more wells on the side with better institutions. In the remainder of the paper, we will focus on the number of wells given that drilling has occurred (the 'intensive margin').

The RD design offers visual evidence of the discontinuity in exploration activity at the national borders. Figure 3 shows the number of exploration wells drilled within bins of distances from the border. We follow the standard in the RD-literature and fit separate polynomials on each side of the border. There is a clear jump at the border, with higher likelihood of a well being drilled on the side of the border with relatively better institutions (i.e., the right hand side of the border, defined as the country with a bilaterally higher score on the augmented Freedom House Political Rights Index in 1965 and earlier).[24] Close to the border, the geological likelihood of discovering oil is the same on both sides and, given the assumption of predetermined

[24]Note that we use as our baseline measure of institutional quality an augmented version of the Freedom House Political Rights Index measure provided by Acemoglu et al. (2008). This is normalised to a zero to one scale, where one means a country comes closest to the political rights ideals suggested by a checklist of questions.

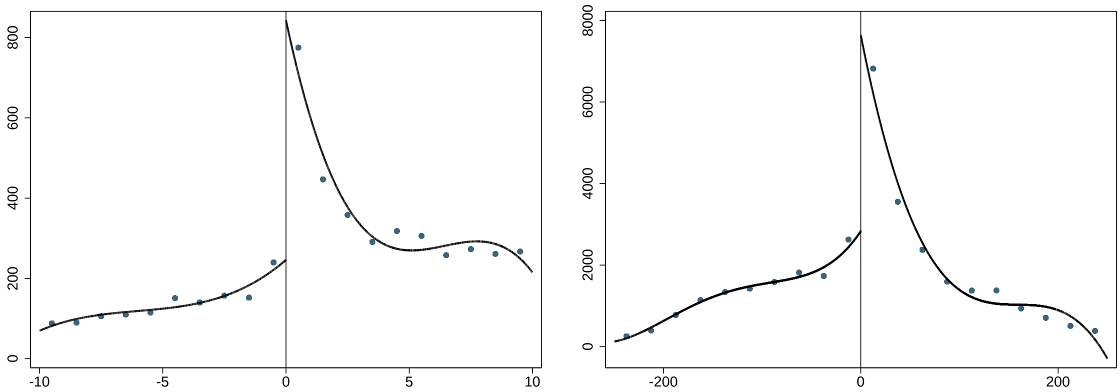

Figure 3: Number of wells around national borders

Notes: Countries placed left or right of zero depending on their bilateral ranking in terms of the Freedom House Political Rights Index. Countries with higher scores are placed on the right hand side of zero. The left hand side panel is based on 1 km bins, 10 km distance from the border and third-order polynomials. The right hand side panel is based on 25 km bins, 250 km distance from the border and fourth-order polynomials. See notes for table 2 regarding the sample and further information about the measure of institutional quality.

borders and institutional quality, the discontinuity indicates the causal effect of differences either side the border.

Figure 4 illustrates the data along a single border. Here we observe recent drilling (over the past decade) on the Ugandan side of the Albert Rift geological basin, but no exploration on the Democratic Republic of the Congo (DRC) side. Given that the oil basin spans both sides of the border, we would not expect, ex-ante, for drilling to be more promising on one side of the border than the other for geological reasons. Our hypothesis is however that institutions influence the drilling decision, and thus observed well distribution would be positively correlated with the quality of institutions locally. For 2013, Uganda's Freedom House Political Rights Index score is 4.5 (*Partly free*), while the score for the DRC is 6 (*not free*). The first well was drilled in pre-independence Uganda in 1938 but was unsuccessful. It was not until the early 2000s that exploration drilling restarted. Since 2002 over 60 wells have been drilled. There is now discussion that exploration drilling may get underway on the DRC-side of the Albert Rift in 2014, twelve years behind Uganda. This lower intensity of drilling and associated delay on the DRC side of the border is consistent with a positive effect of institutional quality on oil exploration.

Several countries have seen evolution in their democratic institutions during our sample pe-

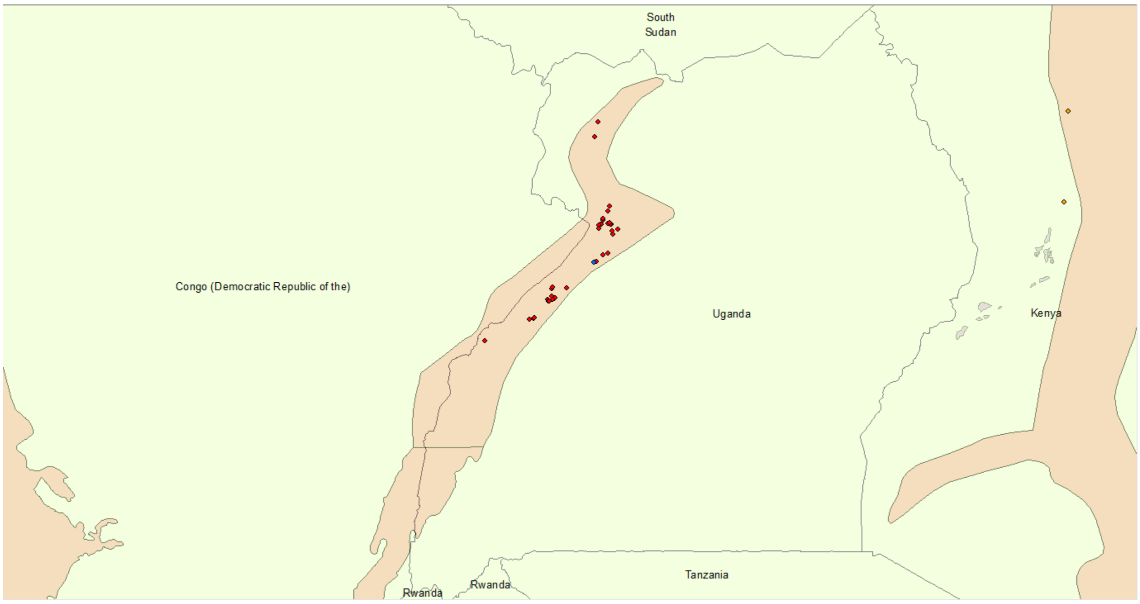

Figure 4: Map of Albert Rift region along Uganda-DRC border.

Notes: Albert Rift basin indicated, along with exploration wells, notably concentrated in the last decade on Ugandan side of the border with DRC.

riod and have recorded associated changes in drilling activity. Ghana, for example, saw strong improvements in their democratic governance scores from 1990s onwards and have seen increased and high sustained drilling activity since the late 1990s, leading to major oil discoveries in 2007 onwards and now a nascent oil sector. Thailand similarly saw progress on democratic institutions in the 1990s and associated increases in drilling activity. Venezuela's shift away from democracy since the 1990s has been accompanied by an associated decline in exploration drilling.

3.2 Econometric estimates of the effect of crossing the border

To formally estimate the effect of crossing the border on oil exploration, we follow the approach outlined in section 2.2. Column 1 is based on 5 km distance from the border, whereas column 2-6 are based on a 10 km distance from the border. Column 1 and 2 are without any controls and polynomials in the distance to the border. Column 3-6 include controls for area size and landlocked status at the country level. Column 4-6 include polynomials of increasing orders, estimated separately on each side of the border.

Table 2: Baseline estimates: Direct effect of crossing the border

Dependent variable →	N: Number of wells					
	(1)	(2)	(3)	(4)	(5)	(6)
Polynomial of order →	Zero	Zero	Zero	First	Second	Third
D = 1 rhs	0.543**	0.467***	0.692***	1.016***	1.490**	1.851**
	(0.256)	(0.159)	(0.168)	(0.386)	(0.617)	(0.821)
ln Area			0.525***	0.559***	0.557***	0.554***
			(0.027)	(0.030)	(0.030)	(0.030)
Landlocked			0.344	0.406	0.533	0.581
			(0.605)	(0.608)	(0.606)	(0.604)
$\bar{N}_{left}$	2.99	2.55	2.55	3.60	4.12	4.63
$\hat{\tau}/\bar{N}_{left}$	0.18	0.18	0.27	0.28	0.36	0.40
Observations	1251	2396	2396	2396	2396	2396
Countries	48	50	50	50	50	50
Neighbours	69	76	76	76	76	76
R-sq	0.00	0.00	0.06	0.11	0.12	0.13

Notes: ***p<0.01, **p<0.05, *p<0.10. Standard errors, shown in parentheses, are clustered at distance-bins (there are 200 of them, 100 for 5 km). The dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. Polynomials in distance from the border included separately for left and right hand side, order as indicated in column headings. Distance from the border: 10 km, except column 1 with 5 km. Only borders that have not changed since 1965 and only wells for which drilling started in the period 1966-2010 are included. Institutional quality is measured as the country average over the years prior to, and included, 1965, based on the augmented Freedom House Political Rights scores used by Acemoglu et al. (2008). Only developing countries included as host country, while neighbours may be either developing or high income countries. Onshore and offshore wells. Country-distance-bins without observed wells are exclude. $\hat{\tau}/\bar{N}_{left}$ is the ratio between the estimated coefficient on the border dummy and the mean on the left hand side of the border. This mean is estimated conditional on a polynomial in distance of the same order as in the regression with the border dummy.

The first row in the lower part of the table, indicated with $\bar{N}_{left}$, provides the mean of the dependent variable on the left hand side of the border. The row below presents the ratio between the estimated coefficient and the left-hand-side mean.[25]

The estimates suggest a jump at the border of 0.5 to 1.9 wells, translating into a relative jump of 18% to 40%. The area and landlocked controls increase the estimated coefficients, as do higher orders of the polynomials.

To investigate further the implications of the specification, we re-estimate the models with OLS and log number of wells as dependent variable, with the Poisson estimator and with the

[25]The mean on the left hand side is based on wells drilled in the period 1966-2010, as the other estimates. It is estimated as the constant in a separate regression for the left hand side including the relevant polynomial as control. The ratio between the estimated coefficient on the border dummy and the estimated left hand side mean therefore gives the relative jump in drilling moving from the left to the right in the relevant period.

negative binomial (NB) estimator. These choices are motivated from the fact that our dependent variable is a count variable. The results are shown in the appendix table B.2 for first and second order polynomials. In terms of scaled estimates, log-linear OLS generate smaller estimates than those in table 2, whereas Poisson and NB create larger estimates.

Based on these empirical results and theoretical reasons for why the distribution of wells around borders may be well described by a second order polynomial (see section 2.2), we choose to go with column 5 in table 2 as our baseline estimate. This estimate suggests that the crossing to the 'better' side of the border increases, on average, the number of wells drilled by about 36%.

Since all our estimates are based on a cross section of wells drilled from 1966 to 2010, i.e. the number of wells accumulated over 45 years, the estimates can be interpreted as close to long-run estimates.

Our results are in line with the literature on institutions and economic growth, which has found that differences in institutional quality helps to explain the differences in economic performance between countries. In particular, a positive effect of democracy on economic growth has been found by Acemoglu et al. (2014), Papaioannou and Siourounis (2008) and Rodrik and Wacziarg (2005). Our results suggest that oil exploration may not be different to other economic activity in its relation to institutional quality measured as democracy.

3.3 Scaling the effect of the border with institutional quality

The estimate of the effect of crossing the border as presented above had the benefit of relying only on ordinal information in the institutional measure. In this section we take one step further and present estimates based on a cardinal interpretation of measures of institutional quality. Although transforming something as complex as institutional quality into a cardinal score may be a challenging task, a cardinal interpretation is not uncommon in the literature (e.g., Acemoglu and Johnson 2005). We scale the border-effect by estimating with 2SLS, as described in section 2.2. The border dummy D is used as an 'instrument' for the institutional quality I. This serves to take into account how much the institutional quality actually jumps at the border.

Table 3: Scaling with measures of institutions

Dependent variable →	N: Number of wells				
	(1)	(2)	(3)	(4)	(5)
I: Inst. measure →	FH	Polity	Democ	Autoc	ConEx
	Direct effect of crossing the border:				
D = 1 rhs	1.490**	2.422***	1.904***	-2.755***	2.443***
	(0.617)	(0.759)	(0.694)	(0.760)	(0.698)
$\bar{N}_{left}$	4.12	3.70	3.92	3.92	3.44
$\hat{\tau}/\bar{N}_{left}$	0.36	0.65	0.49	0.70	0.71
Observations	2396	2348	2347	2351	2352
Countries	50	53	53	53	53
Neighbours	76	66	66	66	66
R-sq	0.12	0.13	0.13	0.14	0.13
Clusters (dbin)	200	200	200	200	200
	Scaled effect (second stage):				
I	5.941**	0.433***	0.735***	-1.132***	0.963***
	(2.575)	(0.136)	(0.273)	(0.325)	(0.273)
F instr	147.59	364.72	168.26	260.15	504.27
	Effect of the border on institutional quality (first stage):				
D = 1 rhs	0.251***	5.601***	2.589***	2.435***	2.536***
	(0.021)	(0.293)	(0.200)	(0.151)	(0.113)
R-sq	0.26	0.36	0.26	0.32	0.41

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. As column 5 in table 2, except for the institutional measures in column 2-5, which are from the Polity IV dataset. For all measures, institutional quality is the country average for all years available up to, and including, 1965. The first stage estimates are based on the bilateral ranking obtained by the measure in question. # See footnote 28.

The 2SLS-estimates are presented in table 3 and we report the unscaled effect of crossing the border in the upper panel ('reduced form' estimates in 2SLS-jargon), the seconds stage estimates in the middle panel and the first stage estimates, i.e. the effect of the border on institutional quality, in the lower panel.

We first focus on column one, which is based on our baseline institutional measure, the augmented Freedom House Political Rights Index (FH). The first stage estimate means that crossing the border gives a jump in FH of about 0.25 on average (the positive coefficient is by construction). For comparison, the score varies on a 0-1 scale and its mean and standard deviation are 0.72 and 0.27 in our sample. The second stage estimate is simply the combination of the first stage and the 'reduced form' estimate, i.e. $1.49/0.25 \approx 6$. A one standard deviation increase in FH, a change of 0.27, corresponds to 39% more wells, moving from the 'bad' side to the 'good' side.[26] For comparison, the difference in FH between DRC and Uganda is of 0.29.

Column 2-5 present estimates based on Polity IV measures of broad institutional quality, a common supplement to The Freedom House Political Rights Index (e.g., in Acemoglu et al. 2008). The polity index (*Polity*) is a composite of indices of democracy (*Democ*) and autocracy (*Autoc*). *ConEx* captures the constraints imposed on the executive branch and is used as a proxy for the quality of 'property rights institutions' by which citizens are protected against expropriation by the elite or the state (Acemoglu and Johnson, 2005).[27]

The upper panel of table 3 demonstrates that the jump in the number of exploration wells at the border holds true also for the Polity IV measures, and varies from 49 to 71%.[28] Our baseline estimate of 36% sits below the estimates based on the Polity IV measures.

The first stage estimates shown in the lower part of the table reveal statistically significant jumps in institutional quality at the borders. The second stage estimates presented in the middle panel suggest that a one standard deviation improvement in the Polity IV measures corresponds to an increase in oil exploration of between 59 and 82%, again is the comparable effect of 39% estimated with the Freedom House measure smaller.

Overall, table 3 confirms our finding across different measures of institutional quality, all broadly defined. In section 3.5 we argue that the effect we estimate works via property rights institutions as well as other aspects. The results presented in this section is an average effect across different types of companies and countries. In section 5.1, we investigate heterogeneity across company types and country groups (developing and high income).

[26] As an example, this is found by $\hat{\gamma} * std.dev./\bar{N}_{left} = 5.941 * 0.27/4.12$

[27] *ConEx* has been used by several other authors, e.g. Acemoglu and Johnson (2005), Acemoglu et al. (2001) and Guriev et al. (2011). Guriev et al. (2011) identify 98 'nationalisations' in 42 countries in the oil industry in the period 1960-2006, of which most took place in the 1970s, some in the 1980s, zero in the period 1990-2005, and again some in 2006 (see also Joff et al. 2009).

[28] For autocracy, a higher score means more autocracy, so we expect a negative sign. Given that we have defined the left hand side of the border as the side with the lower score, the left hand side is in this particular case the side with the 'better' institutional quality. The relative jump from the 'bad' to the 'good' side based on the autocracy measure is therefore 70%: the mean on the 'bad' side is 6.68-2.76=3.92; and the relative jump therefore 2.76/3.92=0.70.

There has recently been a debate in the resource curse literature on whether oil has a detrimental effect on democracy.[29] Institutional quality and oil exploration (and then wealth and subsequent production) are potentially determined simultaneously. We have argued that spatial variation together with the use of predetermined borders and institutions allow us to identify the causality running from institutions to oil exploration (see also section 4.4). For the literature studying the causality running from oil to political (and economic) outcomes, our finding calls for caution. There will be a positive bias in a regression of institutional quality on some oil-measure correlated with exploration if the simultaneity is not tackled properly. Cassidy (2016) finds such a bias in his work on the long-run effects of oil wealth on development.

3.4 Institutions versus geography and culture

The fundamental causes of economic growth are generally classified as either 'institutions', 'geography' or 'culture' (Acemoglu et al. 2005, North and Thomas 1973). In this paper we focus on the effect of institutions and our identification strategy seeks to isolate this from potential effects of geography and culture. We do not attempt to investigate the roles of proximate causes of economic growth, like the accumulation of physical and human capital, the level of technology and the organisation of production. They are considered to be the channels through which the fundamental causes may operate and may be the subject of future work.[30]

Geological prospects, equivalent to a form of geography in the North terminology, is unlikely to induce bias in our setting. The idea behind the RD design is that the geology is effectively identical at the border. We present in section 4.1 evidence in support of no disconti-

[29]For example, Haber and Menaldo (2011), Andersen and Aslaksen (2008), Ross (2013), Cotet and Tsui (2013) and Tsui (2011). Haber and Menaldo (2011) argue that 'increases in resource reliance are not associated with authoritarianism'. In contrast, Andersen and Ross (2014), using the same data, find that oil wealth became a 'hindrance to democratic transitions after the transformative events of the 1970s.' Tsui (2011) finds that 'larger oil discoveries are causally linked to slower transitions to democracy.'

[30]Glaeser et al. (2004) argue that 'human capital is a more basic source of growth than are the institutions.' Note that infrastructure is often purpose-built for the oil sector, and know-how, labour and capital required for oil exploration and extraction are likely to be imported, especially in developing countries. As an example from a developed country, the import-share for equipment to the oil industry in Norway was close to 100% in the first decade of oil production (1970s) and had dropped to just below 40% in 2013, according to the Norwegian Central Bank (governor's annual address, February 14 2013). Rodrik et al. (2004) focus on economic integration alongside geography and institutions. In our analysis, neighbouring country-pairs share bilateral integration. Integration across other borders should to some extent be captured by country fixed effects in column 6 in table 9.

nuity in geology at our borders. We also show in section 4.2 that the likelihood of discovering oil, conditional on drilling, is the same on each side of the border. Our finding is furthermore robust to making the geographic areas around the borders very local and less local (we present estimates for 5 km, 10 km and 250 km from the border). Controls for area size and landlocked status take into account that operational costs may be asymmetric across the border due to non-local geographic features. Finally, the finding is also robust to the inclusion of country fixed effects, which controls for unobserved country-wide geographic features (see section 4.1).

Table 4: Controlling for ethnicity

Dependent variable →	N: Number of wells			
	(1)	(2)	(3)	(4)
Specification →	REF	E FE	REF; g	E FE; g
D = 1 rhs	1.432^{***}	1.071^{***}	1.287^{***}	0.828^{***}
	(0.232)	(0.255)	(0.249)	(0.294)
$\bar{N}_{left}$	1.51	1.51	1.51	1.51
$\hat{\tau}/\bar{N}_{left}$	0.95	0.71	0.86	0.55
Observations	2471	2471	2471	2471
Countries	46	46	46	46
Neighbours	69	69	69	69
R-sq	0.07	0.04	0.09	0.05

Notes: Dependent variable is the number of exploration wells per unit of observation, which is now defined as the bit of a distance-bin in country i at the border with country j, located in the area of ethnic group E. Column 1 presents estimates in this sample with our baseline specification. Column 2 and 4 include Ethnic fixed effects. Column 3 and 4 includes second-order polynomials (indicated g) in the distance to nearest ethnic boundary, estimated separately for left and right hand side of the border. 180 different ethnicities are covered in this table. Controls for country size and landlocked status are included. Based on onshore wells only. Otherwise as described by the note of table 2.

Omitting culture could bias our estimates if: culture changes sharply at national borders; the bilateral ranking of countries in terms of culture is correlated with the ranking based on institutional quality; and culture is correlated with oil exploration.[31] Geo-coded data on ethnicity and the fact that ethnicity often stretches across national borders (Michalopoulos and Papaioannou, 2014), allow us to take into account at least one aspect of culture by including

[31]Michalopoulos and Papaioannou (2014) argue that aspects of culture often stretches across borders instead of varying sharply with them, for a sample of African countries. In robustness checks presented in Table 8 we find that our result is robust to restricting the sample to sub-Saharan African countries and all countries in Africa. Furthermore, we would argue the oil industry is likely to be less driven by local factors such a cultural ties, trust and networks, instead operating in capital-intensive enclaves.

ethnicity fixed effects. In table 4 we split our distance-bins according to ethnicity and the unit of observation becomes then an ethnicity-specific distance-bin located in country *i* close to the border of country *j*. We exclude offshore wells, since ethnicity is only relevant onshore, and all wells in areas without specified ethnicity.[32] As a reference point, the first column presents results with our standard specification. Column two includes ethnicity fixed effects, hence the border effect is estimated with variation within a given ethnicity only. Column three and four modify column one and two by the inclusion of second-order polynomials in distance to the nearest ethnicity boundary. This distance is the mean distance between the wells and their nearest ethnicity boundary for each unit of observation. We allow these second-order polynomials to be different on the left and right hand side of the national border. The estimated coefficient on the border dummy is somewhat smaller when ethnicity is controlled for, but the finding of this paper holds in all specifications. In the strictest specification, crossing the national border towards the side with better institutional quality increases the number of wells by 55%, against 95% without ethnicity controls. To the extent ethnicity captures culture, the results in table 4 provide support that our results are not driven by culture.

A final remark is that our design is not well suited for investigating the effects of geography and culture. Instead of falsifying their role, our finding is that institutional quality makes a substantial difference in attracting investments in oil exploration.

3.5 Unbundling the effect

Acemoglu and Johnson (2005) unbundle the effects of institutions on GDP per capita and investments, as well as other outcomes, into contracting and property rights institutions. As one of their measures on property rights institutions, they use the Polity IV measure constraints on the executive. Column five in table 3 shows that a better score on this measure increase exploration drilling. This is consistent with the theory of Bohn and Deacon (2000), as investors

[32] To measure and locate ethnicity we use the Geo-Referencing of Ethnic Groups (GREG) dataset of global ethno-linguistic regions (Weidmann et al., 2010). This dataset is based upon on maps and data drawn from the Soviet *Atlas Narodov Mira* - a project charting ethnic groups carried out in the 1960s. The Atlas is notable as the basis for the widely utilised Ethno-Linguistic Fractionization index (Taylor and Hudson, 1972).

react negatively to expropriation risk. Measures of property rights and contracting institutions may be correlated, however, and we include in table 5 the border dummy defined on the basis of property rights institutions together with the four other measures of institutions from table 3. The relative jump due to constraints on the executive is reduced from 71% to 12-16%. The relative jump due to polity, democracy and autocracy from the Polity IV database are reduced by 7-10 percentage points, whereas the relative jump due to the Freedom House democracy index increases with six percentage points. Acemoglu and Johnson (2005) use UK legal origin as an instrument for contracting institutions and population density in the 1500s as well as settler mortality as instruments for property rights institutions. They find that the instruments enter in a separable form in the first stages, e.g. the legal origin affects contracting institutions only. We do not find first stages with such separable forms in our setting, inhibiting a precise unbundling into contracting versus property rights institutions. The results in table 5 suggest that there is a separate effect of constraints on the executive. To the extent the measure of constraints on the executive soak up all aspects of property rights institutions, the results in table 5 also suggest sizeable additional effects of other aspects of the institutional setting on oil exploration.

Table 5: Constraints on the executive versus democracy and autocracy

Dependent variable →	N: Number of wells			
	(1)	(2)	(3)	(4)
I: Inst. measure →	FH	Polity	Democ	Autoc
D = 1 rhs I	1.746***	2.151***	1.544**	-2.450***
	(0.645)	(0.745)	(0.668)	(0.745)
D = 1 rhs ConEx	0.533***	0.427***	0.631***	0.478***
	(0.150)	(0.146)	(0.147)	(0.120)
Mean L	4.13	3.70	3.92	3.92
Scaled I	0.42	0.58	0.39	0.63
Scaled ConEx	0.13	0.12	0.16	0.12
Observations	2255	2348	2347	2351
Countries	47	53	53	53
Neighbours	62	66	66	66
R-sq	0.14	0.14	0.13	0.14

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. Identical to the upper panel of table 3, except that the border dummy based on the constraints on the executive is included in all columns.

3.6 Effects away from the border

The effect of institutional quality on oil exploration may be different at national borders compared to elsewhere in the country. Unitization, i.e. 'the joint development of a petroleum resource that straddles territory controlled by different companies', is often regulated by law within countries, e.g. in several states in the U.S. (Boyce and Nostbakken, 2011). This is less common across international borders, although examples exist, for example in the North Sea. It is less common still between developing countries. The consequence may be a race for oil deposits located close to national borders (Libecap and Wiggins 1984, Libecap and Wiggins 1985).

The geology may also be different in border regions than elsewhere. Although the geology may be identical on each side of the border, we cannot exclude the possibility that border regions are associated with geology that differs systematically from the interior of countries. For example, borders could be commonly drawn in rivers or between mountains, which may in turn be associated with relatively better or worse oil prospects than the country on average. We assess that these concerns may affect the external validity of our estimates, but not their internal validity.

Table 6 investigates how the effect changes as we move away from the border. We first expand the distance from the border to 250 km. Column one includes all distance-bins. In column two, we use a 10 km 'thick border', i.e. 5 km on each side of the border are excluded. In column three to eight, we increase the thickness to 20 km and then by 40 km for each column, up to 200 km. The idea is that the resource competition effect should be less relevant the further away from the border we get. The effect in column 1 is 52%. From column two to eight, it increases monotonically from 49% to 177%.[33] Although we cannot rule out the presence of a resource competition effect or special geology in border regions, it does not seem to be the whole story: the effect of institutional quality on oil exploration holds qualitatively also away from the border. The exact size of the estimates in this paper should, however, be interpreted

[33]We compare now with the unconditional mean on the left hand side, as this is most transparent. Use of the conditional mean produces the same pattern of increasing coefficients, varying from 50% to 268%.

Table 6: Thick borders

Dep. var. →	*N*: Number of wells							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Thickness →	0 km	10 km	20 km	40 km	80 km	120 km	160 km	200 km
D = 1 rhs	0.847***	0.775***	0.820***	0.861***	0.884***	1.064***	1.603***	3.027***
	(0.094)	(0.045)	(0.047)	(0.056)	(0.084)	(0.136)	(0.235)	(0.420)
$\bar{N}_{left,uncond.}$	1.64	1.58	1.56	1.55	1.56	1.59	1.64	1.71
$\hat{\tau}/\bar{N}_{left,uncond.}$	0.52	0.49	0.53	0.55	0.57	0.67	0.98	1.77
Observations	23352	21883	20760	18760	15222	12364	9733	7002
Countries	70	69	69	64	54	43	43	36
Neighbours	96	91	91	90	87	75	68	57
R-sq	0.07	0.08	0.07	0.06	0.04	0.03	0.03	0.03

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. Based on 250 km distance from the border to ensure sufficient data. Columns differ by the thickness of the border, e.g. '10 km' indicates that wells within 5 km on each side of the border are excluded. The mean used for scaling is the unconditional mean (polynomial not included). For comparison, the conditional mean is 1.91 and the relative effect 0.44 in the full sample (column 1). The number of distance-bins on which the standard errors are clustered varies from 4681 to 2671. Otherwise as column 5 in table 2.

with the caveat of possible special circumstances in border regions. On the other hand, the bias in the RD-estimate should be the smallest at the border, where it is most plausible that the underlying geology is the same on both sides of the border.

4 Validity of the identification strategy

In this section we examine potential threats to the internal validity of our estimates.

4.1 Exogenous borders

The formal designation of offshore borders has to a large extent been a latter-half of twentieth century phenomenon, in contrast to onshore borders which have existed in various legal forms for centuries. The location of known and expected oil might therefore be thought to have greater influence on the location of offshore borders. Furthermore, offshore borders are arguably more contestable, since no one lives there and legal claim may be less established. In column one in table 7 we exclude offshore wells and the estimate of the effect of the border is, if anything, larger for this sample of onshore wells only. This suggests that our findings are not driven by strategic location of offshore borders by countries with better institutions than their neighbours.

Table 7: Subset of borders: Onshore; historically stable; without military disputes

Dependent variable →	*N*: Number of wells				
	(1)	(2)	(3)	(4)	(5)
	Onshore	1946	1938	1925	No MD
D = 1 rhs	1.833***	0.953***	3.812***	3.856***	2.658***
	(0.635)	(0.339)	(0.720)	(0.727)	(0.674)
Mean L	3.41	2.27	1.65	1.65	3.53
Scaled L	0.54	0.42	2.31	2.34	0.75
Observations	1874	1274	439	427	1724
Countries	46	34	8	7	42
Neighbours	70	39	14	13	59
R-sq	0.09	0.25	0.32	0.31	0.13

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. Column 1 excludes offshore wells. Column 2-4 include only borders not changed since year as indicated (see the description regarding stable borders in the online appendix B.2). Column 5 excludes borders with recorded military interstate dispute after 1945. Column 1 and 5 use borders stable since 1965, as standard elsewhere in the paper. Second-order polynomial and area and landlocked controls are included. There are between 178 and 200 clusters. Otherwise as described by the note of table 2.

Could it be that our findings are driven by borders moving due to interstate territorial conflicts? Nation states with relatively high institutional quality could have been able to redraw borders via invasion and occupation of neighbouring territories with oil or promising geology. To challenge our identification, the border would have to be permanently moved, i.e. a lasting occupation and not merely a threat of or attempted occupation. Caselli et al. (2014) argue that the distance from the border to an oil field is a predictor of bilateral wars, but they do not identify many incidents of borders actually moving, nor that any such effect should systematically work in favour of neighbours with better institutions.[34] However, to deal with the potential issue of moving borders, we consider in this paper only stable borders. This means that we always exclude disputed offshore border areas and include only borders which have not changed since 1965 (see explanation in section 2). To deal with instability before 1965, column 2-4 in table 7 limits the sample to countries with borders that have not changed since 1946, 1938 and 1925. To make comparison with previous results easier, we still use wells drilled after 1965

[34] See Caselli et al. (2014) and Acemoglu et al. (2012) for more on resource wars.

only and the institutional quality measured before 1966. The finding holds and the size of the estimates are larger in all three cases. The results are also robust to controlling for wells drilled before 1966 (available at request from the authors).

Even for our stable borders, inter-state conflict may affect the likelihood of oil exploration. Column five in table 7 excludes borders across which there has been at least one incidence of Militarized Interstate Dispute (MID).[35] The estimated coefficient on the border dummy is positive and significant and somewhat larger than when the MID-borders are included. If anything, this indicates that incidence of conflict at a border reduces the appetite to drill close to that border later on.

Could it be that colonial powers were attracted to a given territory because of oil, creating a positive correlation between certain institutional quality and oil drilling? This is not a great threat to our identification strategy as almost all colonization took place long before the location of oil was plausibly known. Furthermore, the colonial powers would need to know very precisely where the oil was located, and draw the border accordingly, since our analysis relies on comparison of areas very close to national borders. Moreover, instead of claiming that colonial powers carefully assigned borders according to, for example, the presence of natural resources, Alesina et al. (2011) and Michalopoulos and Papaioannou (2016) argues that many states in Africa were imposed "artificial borders" by colonial powers. In table 8 we restrict our sample to Sub-Saharan African countries only (column 1), African countries (column 2), countries belonging to top third partitioning countries (column 3) and countries belonging to the third of countries with the straightest borders (column 4). the two latter criteria is based on data from Alesina et al. (2011). We impose that both countries in a pair satisfy the criteria. We present estimates for onshore wells only, as our judgment is that the artificial borders arguments is mostly valid for onshore borders. Results are the same if we include also offshore borders. As these restrictions reduce our sample considerably, we present estimates for 250 km distance

[35]To define borders with MIDs we employ the Dyadic MID dataset used by and Martin et al. (2008). The dataset codes MIDs with a hostility level ranging from 1 to 5, where 1=No militarized action, 2=Threat to use force, 3=Display of force, 4=Use of force and 5=War. We follow Caselli et al. (2014) in defining conflict as hostility level 4 and 5. The dataset (version 2.0) is presented by Zeev Maoz (2005) at: `http://psfaculty.ucdavis.edu/zmaoz/dyadmid.html`. We consider conflicts 'active' in the post 1945-period, meaning that conflicts that started after 1945, or were not ended in 1946, are included.

from the border in appendix table B.4.

Table 8: Restricting sample to "artificial states"

Dependent variable →	*N*: Number of wells			
	(1)	(2)	(3)	(4)
Sample →	SSA	Africa	PART	FRAC
D = 1 rhs	16.681***	12.545***	23.413***	0.695
	(2.404)	(1.431)	(3.910)	(0.720)
Mean L	1.28	1.17	1.27	4.07
Scaled L	13.04	10.75	18.49	0.17
Observations	183	303	202	322
Countries	8	12	6	14
Neighbours	12	15	7	15
R-sq	0.37	0.32	0.42	0.36

Upper panel 10 km, lower panel 250 km. Columns: 1 SSA, 2 Africa. 3 partdummy=1 Those countries in the top third of partitioning, as measured by Alesina 2011. 4 fractdummy=1 Those countries in the bottom third of fractal measure (i.e. third of countries with straightest borders), as measured by Alesina 2011. 5 artifdummy=1 Those countries Both in the top third of partitioning AND the bottom third of Fractals- as discussed by Alesina et al. (2011). Wells drilled after 1966, FH measured before 1966. Second-order polynomial and area and landlocked controls are included. There are between 117 and 184 clusters. Otherwise as described by the note of table 2.

What about other unobserved characteristics correlated with the location of national borders, the location of subsoil oil deposits, and the quality of institutions? Mountain ranges, rivers or deserts might be candidates. However, these features do not appear to systematically imply the presence of oil on one side of the border versus another. Furthermore, we are not aware of any evidence that topographic features are correlated with the quality of institutions. We have therefore so far been unable to locate a plausible candidate for such co-determination.[36] On the contrary, figure 4 plots the pattern of exploration wells in the Albert Rift, an oil basin on the border between Uganda and the Democratic Republic of Congo. A global map of geological basins and national borders is presented in figure B.3. Both figures reveals to us no obvious relationship between borders and geology. To investigate this assertion more thoroughly, we generate 120 000 random points spread across the globe. We use the map-layer of geological basins used in figure B.3 and count the number of random points landing within geological basins on each side of the border. The result is presented in figure B.2, which is constructed

[36]Alesina et al. (2011), Turner et al. (2014) and Michalopoulos and Papaioannou (2013) also draw inference based on the variation created by political boundaries.

in the exact same way as our border graph of actual exploration wells presented in figure 3. The distribution of random points is symmetric with no discontinuity at the border, equivalent to the surface area of basins not jumping at the border, in stark contrast to the distribution of actual oil wells. This provides strong support for the key identifying assumption of this paper: the location of borders has been determined independently of the location of oil.

4.2 The likelihood of discovering oil and path dependency

Table 9: Discovery rates, path dependency and unobservables

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable →	Disc. rate	Disc. rate	Wells	Wells	Wells (0s)	Country FE
D = 1 rhs	-0.002	-0.004	1.693***	1.284***	0.215***	1.454**
	(0.050)	(0.051)	(0.441)	(0.426)	(0.072)	(0.591)
Wells pre 1966		-0.006	0.890***			
		(0.005)	(0.253)			
Discoveries pre 1966				1.654***		
				(0.510)		
$\bar{dep.var.}_{left}$	0.51	0.51	4.12	4.12	0.28	4.12
$\hat{\tau}/\bar{dep.var.}_{left}$	-0.00	-0.01	0.41	0.31	0.76	0.35
Observations	2396	2396	2396	2396	22811	2396
Countries	50	50	50	50	50	50
Neighbours	76	76	76	76	76	76
R-sq	0.01	0.01	0.26	0.28	0.06	0.08

Notes: Dependent variable is the share of non-dry wells in column 1-2. Tight holes, i.e. wells for which the result is kept secret, are set to missing. In column 3-6, the dependent variable is the number of exploration wells. Both dependent variables are calculated for each 100 meter wide distance-bin in country i at the border with country j. Column 4-5 include zeros for distance-bins without drilling. Column 6 includes country fixed effects. Controls for country size and landlocked status are included in column 1-5. Otherwise as described by the note of table 2.

Consistent with a smooth likelihood of finding a geological basin either side of borders, as presented at the end of the former subsection, column one in table 9 presents an insignificant coefficient on our standard border dummy when the discovery rate is used as the dependent variable.[37] The discovery rate is defined as the share of non-dry wells in a given distance-bin, i.e. the ex-post likelihood of discovering oil. In contrast to the number of wells drilled, the likelihood of finding oil does not appear to differ across the borders. Although the oil

[37] We define dry wells using the industry standard definition: a dry well or dry hole in our data is denoted as such where evaluated to contain insufficient oil for commercial production.

exploration companies have geological information available before they start drilling, there is uncertainty about whether a given well will lead to discovery before it is drilled. It is likely, however, that experience in an area increases the likelihood of discovery. On the other hand, it may be that the most promising wells based on the geological information at hand are drilled first, which would contribute to decreasing the likelihood of discoveries in a given area over time. In column two, we include the number of wells drilled before 1966 as a control variable.[38] The results suggest that previous drilling neither affects the discovery rate nor the effect of the border dummy. In column three we modify our baseline specification by including the number of wells drilled before 1966 as a control variable. In this case, previous drilling serves a similar function as a lagged dependent variable, potentially soaking up unobservable characteristics not otherwise accounted for in our analysis. The effect of previous drilling is positive and significant, which could reflect that additional wells have higher expected value due to, for example, infrastructure already in place. The estimate in column three demonstrates that our results are not driven by some accidental early exploration and path-dependency since then. As common infrastructure could lead to spatial correlation and potentially deflate our standard errors, we present in appendix table B.3 our baseline model with standard errors adjusted for spatial correlation and the result holds. In column four, we include the number of discoveries made before 1966 as a control, resulting in a relative jump of 31%.

As we include only distance-bins with observed wells in our baseline specification, a question may be raised about the continuity of our running variable, the distance to the border. In column five we include all distance-bins, i.e. within 10 km of the border we assign zero wells to distance-bins without observed wells. This increases our sample size tenfold. the relative jump at the border is now estimated to 75%. We therefore conclude that any lack of continuity in the running variable does not seem to bias our estimates away from zero.

Could it be that countries with better institutions are also better at reporting drilled oil wells and our findings are just an artifact of such reporting bias? The PathFinder database from Wood Mackenzie that we use in this paper, is considered to have as comprehensive country

[38]For this control variable, zeros are assigned to distance-bins without drilling before 1966.

coverage as any single dataset on oil exploration. It relies in large part on company reporting, but also desk research, country reporting and cross-checking for verification. If underreporting by companies is an issue, our expectation is that this is least severe at international companies. Table 11 shows that our results are robust across different company types, including the six so-called supermajors. In addition, the results also hold if we run the regressions for the sub-samples of each company type only (results available at request from the authors). Thus, also within each company type, there is more drilling on the better side of the borders. Note also that the underreporting would have to be quite severe to plausibly alter our results. In a sample of 51 876 wells drilled in the period 1966-2010, we observe 16 979 wells (33%) on the left hand side and 34 897 wells (67%) on the right hand side, based on the Freedom House measure.

Reporting bias at the country-level is effectively controlled for by the inclusion of country fixed effects in column 6 of table 9, and the result holds. Further, as reporting bias is presumably smaller in high income countries, the fact that our findings hold also for this sub-sample (except for democracy and autocracy measures very near borders), is reassuring.

Although we cannot rule out bias from mis-reporting or under-reporting, the results above give us confidence that this is not the entire story.

4.3 Controlling for other factors changing at the border

As discussed in section 3.4, we focus in this paper on distinguishing between the effects of institutions on the one hand and geography and culture on the other hand. In this section we seek to illuminate the role of proximate causes of economic growth in our setting. We augment our baseline regression with a second border dummy based on bilateral ranks in terms of *GDP per capita*, *GDP*, employment (*L*), human capital (*H*), physical capital (*K*) and total factor productivity (*A*). Together these factors fully describe GDP, according to a standard production function: $GDP = F(L,H,K,A)$. Table 10 presents the results. Notably, our standard border dummy based on institutions is always significant. As the sample size varies, we re-estimate our standard model on the same samples for comparison (see table B.5). Controlling for the border dummies defined based on *GDP per capita*, *H* or *A* reduces the estimates. This is as expected given, for example, Hall and Jones (1999), who found that variation in institutional quality is an important explanation for varying TFP-levels, and hence income per capita, across countries. However, note that the differences in the estimates are very small and we are cautious to read too much from them. The estimated relative effect is in the range 29%-54%, with the main reason for the variation across columns being the sample size, as seen in table B.5.[39] We conclude that the coefficient on the institutions border dummy is remarkably stable. A note of caution to these estimates is that the two border dummies may be determined simultaneously, imposing a bias in our estimates (they are mutually 'endogenous controls').

Table 10: Controlling for border dummies defined on proximate causes of economic growth

Dependent variable →	N: Number of wells					
	(1)	(2)	(3)	(4)	(5)	(6)
Proximate cause →	GDP/Pop	GDP	L	H	K	A
D = 1 rhs	2.179***	2.348***	1.492**	2.194***	2.335***	1.385*
	(0.695)	(0.709)	(0.740)	(0.691)	(0.707)	(0.769)
D = 1 rhs PWT	0.376**	0.920***	0.214	0.460***	0.939***	1.015***
	(0.150)	(0.141)	(0.133)	(0.140)	(0.140)	(0.126)
$\bar{N}_{left}$	4.31	4.31	4.40	4.31	4.31	4.73
$\hat{\tau}/\bar{N}_{left}$	0.51	0.54	0.34	0.51	0.54	0.29
Observations	2004	2004	1795	2001	2004	1659
Countries	42	42	39	41	42	37
Neighbours	56	56	51	55	56	47
R-sq	0.13	0.14	0.11	0.14	0.14	0.13

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. The second border dummy is defined according the bilateral rank of the variables indicated in the column headings, based on data from the Penn World Tables (PWT). We use the closest variables available, taking the averages of values pre 1966 (maximum 1950-65.), i.e. the same approach as we follow for our measures of institutions. The variables are: *GDP* "Output-side real GDP at chained PPPs (in mil. 2011USD); *Pop*: Population (in millions)"; *L* "Number of persons engaged (in millions)"; *H* "Human capital index"; *K* " Capital stock at current PPPs (in mil. 2011USD)"; *A* "TFP level at current PPPs (USA=1)". Penn World Tables 9.0, downloaded December 2016 from http://www.rug.nl/ggdc/productivity/pwt/. Table otherwise as column 5 of table 2.

[39]Inclusion of several of the proximate causes at the same time, and keeping the sample size fixed across all columns, reduces the coefficient on the institutions border dummy somewhat, but it stays significant in all cases except of one. The estimates suggest a jump at the border of 27%-29% (see table B.6).

4.4 Reverse causality

A separate issue to the position of the border is potential feedback from oil exploration and subsequent production to institutional quality. The institutional quality is an equilibrium outcome and natural resources are likely to affect this outcome. One potential channel of the resource curse is the deterioration of political institutions in the face of natural resource rents (Ross 2001, Collier and Hoeffler 2009).[40] As the correlation between oil extraction and institutional quality is suggested by most papers in the resource curse literature to be negative, this would likely bias our estimates downwards towards zero and hence work against the finding of this paper. However, to accommodate any feedback-bias, we have throughout the paper followed Acemoglu and Johnson (2005) and used prior institutions, i.e. institutions are measured before the oil exploration we consider took place.[41] Furthermore, effects of previous exploration should be accounted for by the previous exploration or previous discoveries included in column three and four in table 9.

5 Heterogeneity: company types and countries

5.1 Heterogeneity across company types

Oil exploration is undertaken by a range of different types of companies, varying in scale, degree of vertical integration and model of ownership. If these companies face different objectives and constraints, a country's institutional setting may affect them differently: large publicly-traded companies may face high potential reputational costs; state owned companies may be insulated from political risk by their state backers or weigh non-commercial factors more heavily than their private counterparts; small specialised exploration companies may be backed by risk-willing owners and capable of handling high risks.

[40] Collier and Hoeffler note that these effects are large: "after 28 years a country with mean income but with resource rents worth 30% of GDP would have a 'checks' score in the 22th percentile instead of in the 34th percentile, and a democracy score in the 25th percentile instead of in the 40th percentile."

[41] Recall also that we to define the border dummy use institutional quality simply to rank the two bordering countries. Any feedback from oil exploration and extraction would therefore need to alter the ranking, and not just the level, to induce a bias.

The so-called six supermajors (Chevron, Shell, BP, ExxonMobil, ConocoPhillips and Total; indicated IOC6) are international oil companies with predominantly non-state ownership. They are vertically integrated by engaging in the whole industry value chain from exploration to production to downstream activities. Companies belonging to this group, including their subsidiaries and predecessors, are listed as operators for 16% of the total 112 thousand wells we observe across developing and developed countries. National oil companies (NOCs) are typically set up to secure oil rents accruing to governments and to carry out national strategic objectives.[42] These companies are listed as operators for about 30% of the full sample of wells, and some of them operate outside of their home countries. A third type is the non-supermajor private companies, which include smaller specialised exploration companies who may seek to win particular licenses, engage in exploration and then offload operations to more integrated companies. Anecdotal evidence suggests that the global oil industry previously moved towards less vertical integration and more out-sourcing of high-risk exploration, although that trend may now be reversing post-financial crisis.

In table 11, we augment with interaction terms our baseline model in column 5 in table 2 to test for heterogeneity across company types. Column one includes an interaction between the border dummy and a dummy, *COMP*, taking one for wells operated by one of the 'IOC6' supermajors.[43] In column two, also the subsidiaries and predecessors of the six supermajors are included in the interaction dummy (indicated IOC6+). Column three does the same for NOCs, while only wells operated by NOCs in the NOC's home country are included in the dummy in column four. Finally, in column five the interaction dummy takes one for wells operated by companies not defined as either IOC6+ or NOC. The coefficient on the border dummy is always positive and significant. The coefficient on the interaction term is relatively large and

[42]Many formerly state-owned oil companies have been privatised, particularly during the 1990s. For the analysis in this section we including in the coding of 'OTH' those no longer under majority state-ownership as NOCs in 2013. In all cases we look at the operating company only. Oil fields may be developed with multiple owners and minority participants, and ownership can change over time, especially moving from the exploration phase to the production phase. We thus limit our analysis to the well operator at the point of exploration only. This excludes any examination of the role of non-operator investors or those who may acquire a stake subsequent to exploration (which can sometimes apply to state-participation, where NOCs may acquire minority equity as part of the production phase).

[43]Note that the dummy indicating company type is also included separately.

Table 11: Heterogeneity across company types

Dependent variable →	N: Number of wells				
	(1)	(2)	(3)	(4)	(5)
Company type →	IOC6	IOC6+	NOC	NOCH	OTH
D(FH)	0.896*	1.046**	1.646***	2.062***	2.008***
	(0.499)	(0.465)	(0.546)	(0.587)	(0.527)
D(FH) x COMP	3.087***	1.822***	-0.491**	-0.535***	-0.635***
	(0.525)	(0.431)	(0.201)	(0.204)	(0.216)
COMP	-1.221***	-0.461***	-0.615***	-0.596***	-0.007
	(0.130)	(0.092)	(0.109)	(0.112)	(0.084)
Observations	2633	2732	2658	2573	2823
Countries	50	50	50	50	50
Neighbours	75	76	75	76	75
R-sq	0.15	0.13	0.13	0.13	0.11

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. As indicated, the COMP-dummy takes the value one for the following company-types: IOC6=major international oil company (one of the six super-majors: Chevron, Shell, BP, ExxonMobil, ConocoPhillips, Total); IOC6+=major international oil company, including their subsidiaries and predecessors; NOC=National oil company; NOCH=National oil company wells at home; OTH=Not IOC6+ or NOC. Controls for country size and landlocked status are included. Otherwise as described by the note of table 2. Note that we to identify heterogenous effects across company types sum wells per company-type per country-distance-bin. The number of observations is therefore different across the columns and compared to the baseline-sample.

positive for the IOC6s and negative for the NOCs and the remaining companies, 'OTH'. These results indicate that all company types are sensitive to institutional quality, but the super-majors more so than the others. Better institutional environments may spur foreign investments, while countries with weaker institutions may rely on home grown NOCs, strategic (foreign) NOC partners or smaller exploration companies.

5.2 Heterogeneity across countries

Throughout the paper, we use the sample of all countries for which we have data. We now investigate whether the effect is different in developing countries than in high income countries. Many of the investments in the oil sector may historically have originated in high income countries. Thus, the information asymmetry issues between source and host countries, known from other work to hamper investment flows, are likely to be more of an issue for developing countries. Developing countries are also known to have less friendly business climates and more cumbersome bureaucracies, also reducing investments. Presumably, the asymmetry and bureaucracy issues are less severe in countries with better institutions.[44]

Table 12: Baseline estimates: Direct effect of crossing the border

Dependent variable →	N: Number of wells			
Distance from border →	10 km		250 km	
	(1) Developing	(2) High income	(3) Developing	(4) High imcome
D = 1 rhs	2.839***	-0.057	0.812***	1.119***
	(0.833)	(0.572)	(0.141)	(0.087)
$\bar{N}_{left}$	4.07	4.21	2.07	1.63
$\hat{\tau}/\bar{N}_{left}$	0.70	-0.01	0.39	0.69
Observations	1370	1026	11216	12136
Countries	33	17	49	21
Neighbours	52	31	71	40
R-sq	0.11	0.25	0.07	0.12

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. Sample split into developing versus high income countries based on World Bank classification. See table B.7 for the list of high income countries (indicated "HI"), developing countries are the remaining countries. Otherwise as column five in table 2.

Table 12 presents our baseline model estimated on the samples of developing countries and high income countries separately. We include wells within 10 km as well as 250 km distances from the border and use our standard Freedom House Political Rights Index as the measure of institutional quality. The results of the paper holds, except for the 10 km high income country sample. Using instead constraints on the executive, we estimate a significant effect also in the 10 km high income country sample.[45] The estimated coefficients in table 12 translate to relative jumps of 70% and 39% in the developing country sample and -1% (insignificant) and 69% in the high income country sample, i.e. a larger jump near borders for the developing country sample and a larger jump away from borders for the high income country sample. We conclude that the qualitative message of more drilling on the better governed side of the border seems to hold in general, with the exception of very close to borders in the high income countries

[44]See Harding and Javorcik (2011) for a discussion of information asymmetry and red tape issue in the context of foreign direct investments across developing vs. high income countries. In a sample of 51 876 wells drilled in the 1966-2010 period, we find that 24% are drilled by national oil companies (NOCs), 21% by one of the six major international oil company, including their subsidiaries and predecessors (IOC+) and 54% by a rest category. (Smaller) international oil companies are likely to make a substantial share of the rest category.

[45]Estimates available at request from the authors.

when we use democracy or autocracy measures of institutions. We leave for future research to investigate why the investment behaviour seems to be different at those borders.

6 Conclusions

A natural experiment of borders, randomly assigned with respect to geology, together with pre-determined institutions allow identification of the responsiveness of oil exploration to differences in institutions. Crossing a national border is found to generate a statistically and economically significant jump in exploration drilling for oil and gas. The number of wells increases by 36% and, all else equal, exploration companies prefer to drill on the side with the better institutions 58% of the time.

This paper contributes to the debate on the drivers of cross-country differences in economic performance. First, it shows that institutions have a strong effect on investments. Second, it reveals that the observed distribution of oil wealth and extraction across countries is endogenous; the natural capital component of the 'wealth of nations' responds to institutional quality. Some regions of the world, such as sub-Saharan Africa, are likely to be 'under-explored' with respect to geology alone, which can explain why they have found relatively little oil per square km. Endogenous oil exploration also poses a challenge to the empirical analyses of the effects of oil wealth and extraction. Third, for governments it is an important message that promising geology may not be sufficient to attract oil exploration. To the extent they can improve the institutional environment, they may accelerate discovery and increase their country's level of natural capital.

References

Acemoglu, D., M. Golosov, A. Tsyvinski, and P. Yared (2012). A dynamic theory of resource wars. *The Quarterly Journal of Economics 127*(1), 283–331.

Acemoglu, D. and S. Johnson (2005). Unbundling institutions. *Journal of Political Economy 113*(5), 949–995.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *The American Economic Review 91*(5), 1369–1401.

Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as a fundamental cause of long-run growth. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*. Elsevier.

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2008). Income and democracy. *American Economic Review 98*(3), 808–42.

Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2014). Democracy does cause growth. Working Paper 20004, National Bureau of Economic Research.

Alesina, A., W. Easterly, and J. Matuszeski (2011). Artificial states. *Journal of the European Economic Association 9*(2), 246–277.

Andersen, J. J. and S. Aslaksen (2008). Constitutions and the resource curse. *Journal of Development Economics 87*(2), 227 – 246.

Andersen, J. r. J. and M. L. Ross (2014). The big oil change: A closer look at the Haber–Menaldo analysis. *Comparative Political Studies 47*(7), 993–1021.

Barbieri, K., O. Keshk, and B. Pollins (2008). Correlates of war project trade data set codebook. *Codebook Version 2*.

Besley, T. and T. Persson (2010). State capacity, conflict, and development. *Econometrica 78*(1), 1–34.

Bohn, H. and R. T. Deacon (2000). Ownership risk, investment, and the use of natural resources. *The American Economic Review 90*(3), 526–549.

Boyce, J. R. and L. Nostbakken (2011). Exploration and development of u.s. oil and gas fields, 19552002. *Journal of Economic Dynamics and Control 35*(6), 891 – 908.

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

Caselli, F., M. Morelli, and D. Rohner (2014). The geography of interstate resource wars. *The Quarterly Journal of Economics*.

Cassidy, T. (2016). The long-run effects of oil wealth on development: Evidence from petroleum geology. Technical report.

Claus, S., N. De Hauwere, B. Vanhoorne, F. Hernandez, and J. Mees (2013). Maritime boundaries geodatabase, version 6.1. *Flanders Marine Institute*. Accessed at http://www.marineregions.org on 2011-05-02.

Collier, P. (2010). *The Plundered Planet*. Oxford University Press.

Collier, P. and A. Hoeffler (2009). Testing the neocon agenda: Democracy in resource-rich societies. *European Economic Review 53*(3), 293 – 308.

Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1 – 45.

Cotet, A. M. and K. K. Tsui (2013). Oil and conflict: What does the cross country evidence really show? *American Economic Journal: Macroeconomics 5*(1), 49–80.

Cust, J. and T. Harding (2013). Institutions and the location of oil drilling. Working paper, University of Oxford.

Daniel, P., M. Keen, and C. McPherson (2010). *The taxation of petroleum and minerals: principles, problems and practice*. Routledge.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica 78*(6), 1863–1903.

EIA (2013, July). U.S. Energy Information Administration. http://www.eia.gov/.

Gelman, A. and G. Imbens (2014). Why high-order polynomials should not be used in regression discontinuity designs. Working Paper 20405, National Bureau of Economic Research.

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

Gleditsch, K. S. and M. D. Ward (1999). Interstate system membership: A revised list of the independent states since 1816. *International Interactions 25*(4), 393–413.

Guriev, S., A. Kolotilin, and K. Sonin (2011). Determinants of nationalization in the oil sector: A theory and evidence from panel data. *Journal of Law, Economics and Organization 27*(2), 301–323.

Haber, S. and V. Menaldo (2011). Do natural resources fuel authoritarianism? a reappraisal of the resource curse. *American Political Science Review 105*, 1–26.

Hall, R. E. and C. I. Jones (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics 114*(1), 83–116.

Harding, T. and B. S. Javorcik (2011). Roll out the red carpet and they will come: Investment promotion and fdi inflows*. *The Economic Journal 121*(557), 1445–1476.

Hendricks, K. and R. H. Porter (1996). The timing and incidence of exploratory drilling on offshore wildcat tracts. *The American Economic Review 86*(3), 388–407.

Hijmans, R., N. Garcia, and J. Wieczorek (2010). Gadm: database of global administrative areas, version 2.0.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of Sciences 107*(35), 15367–15372.

Hurn, A. S. and R. E. Wright (1994). Geology or economics? testing models of irreversible investment using north sea oil data. *The Economic Journal 104*(423), 363–371.

Imbens, G. W. and T. Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*(2), 615–635.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1).

Jacoby, H. G., G. Li, and S. Rozelle (2002). Hazards of expropriation: Tenure insecurity and investment in rural china. *American Economic Review 92*(5), 1420–1447.

Janeba, E. (2002). Attracting fdi in a politically risky world. *International Economic Review 43*(4), 1127–1155.

Jensen, N. and L. Wantchekon (2004). Resource wealth and political regimes in africa. *Comparative Political Studies 37*(7), 816–841.

Joff, G., P. Stevens, T. George, J. Lux, and C. Searle (2009). Expropriation of oil and gas investments: Historical, legal and economic perspectives in a new age of resource nationalism. *The Journal of World Energy Law & Business 2*(1), 3–23.

Laurent-Lucchetti, J. and M. Santugini (2012). Ownership risk and the use of common-pool natural resources. *Journal of Environmental Economics and Management 63*(2), 242 – 259.

Lee, D. S. and D. Card (2008). Regression discontinuity inference with specification error. *Journal of Econometrics 142*(2), 655–674.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*(2).

Li, Q. (2009). Democracy, autocracy, and expropriation of foreign direct investment. *Comparative Political Studies 42*(8), 1098–1127.

Libecap, G. D. and S. N. Wiggins (1984). Contractual responses to the common pool: Prorationing of crude oil production. *The American Economic Review 74*(1), 87–98.

Libecap, G. D. and S. N. Wiggins (1985). The influence of private contractual failure on regulation: The case of oil field unitization. *Journal of Political Economy 93*(4), 690–714.

Livernois, J. R. and D. L. Ryan (1989). Testing for non-jointness in oil and gas exploration: A variable profit function approach. *International Economic Review 30*(2), 479–504.

Long, N. V. (1975). Resource extraction under the uncertainty about possible nationalization. *Journal of Economic Theory 10*(1), 42 – 53.

Martin, P., T. Mayer, and M. Thoenig (2008). Make trade not war? *The Review of Economic Studies 75*(3), 865–900.

McKinsey Global Institute (2013). Reverse the curse: Maximizing the potential of resource-driven economies. Report, McKinsey Global Institute.

Mehlum, H., K. Moene, and R. Torvik (2006). Institutions and the resource curse. *The Economic Journal 116*(508), 1–20.

Michalopoulos, S. and E. Papaioannou (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica 81*(1), 113–152.

Michalopoulos, S. and E. Papaioannou (2014). National institutions and sub-national development in africa. *The Quarterly Journal of Economics 129*(1), 151–213.

Michalopoulos, S. and E. Papaioannou (2016, July). The long-run effects of the scramble for africa. *American Economic Review 106*(7), 1802–48.

Mohn, K. and P. Osmundsen (2008). Exploration economics in a regulated petroleum province: The case of the norwegian continental shelf. *Energy Economics 30*(2), 303 – 320.

North, D. C. (1990). *Institutions, institutional change, and economic performance*. Cambridge: Cambridge University Press.

North, D. C. and R. P. Thomas (1973). *The Rise of the Western World: a New Economic History*. Cambridge: Cambridge University Press.

Papaioannou, E. and G. Siourounis (2008). Democratisation and growth. *The Economic Journal 118*(532), 1520–1551.

Pesaran, M. H. (1990). An econometric analysis of exploration and extraction of oil in the u.k. continental shelf. *The Economic Journal 100*(401), 367–390.

Quyen, N. V. (1991). Exhaustible resources: A theory of exploration. *The Review of Economic Studies 58*(4), 777–789.

Rodrik, D., A. Subramanian, and F. Trebbi (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth 9*(2), 131–165.

Rodrik, D. and R. Wacziarg (2005). Do democratic transitions produce bad economic outcomes? *American Economic Review 95*(2).

Ross, M. L. (2001). Does oil hinder democracy? *World Politics 53*(3), 325–361.

Ross, M. L. (2013). The politics of the resource curse: A review. Volume 1 of *Handbook on the Politics of Development*.

Taylor, C. L. and M. C. Hudson (1972). *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.

Thomas, J. and T. Worrall (1994). Foreign direct investment and the risk of expropriation. *Review of Economic Studies 61*(1), 81–108.

Tir, J., P. Schafer, P. F. Diehl, and G. Goertz (1998). Territorial changes, 18161996: Procedures and data. *Conflict Management and Peace Science 16*(1), 89–97.

Tsui, K. K. (2011). More oil, less democracy: Evidence from worldwide crude oil discoveries. *The Economic Journal 121*(551), 89–115.

Turner, M. A., A. Haughwout, and W. van der Klaauw (2014). Land use regulation and welfare. *Econometrica 82*(4), 1341–1403.

van der Ploeg, F. (2011). Natural resources: Curse or blessing? *Journal of Economic Literature 49*(2), 366–420.

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

Venables, A. J. (2011). Depletion and development: Natural resource supply with endogenous field opening. OxCarre Working Papers 062, Oxford Centre for the Analysis of Resource Rich Economies, University of Oxford.

Weidmann, N., J. K. Rd, and L.-E. Cederman (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research 47*(4), 491–499.

Weidmann, N. B., D. Kuse, and K. S. Gleditsch (2010). The geography of the international system: The cshapes dataset. *International Interactions 36*(1), 86–106.

Wood Mackenzie (2011). Pathfinder database, exploration wells dataset, accessed 10 10 2011. Wood Mackenzie’s PathFinder is a commercially-available database, updated quarterly, that contains worldwide exploration and production data for the petroleum industry.

World Bank (2011). *The Changing Wealth of Nations: Measuring Sustainable Development in the New Millennium*. World Bank.

A Summary statistics

Table A.1: Descriptive statistics, baseline samples

	Distance from the border: 10 km					
	count	p50	mean	sd	min	max
Wells per pair	2396	91	214.5	223.1	1	782
Wells (per 0.1 km dist-bins)	2396	2	2.853	3.629	1	85
Wells pre 1966	2396	0	0.669	1.599	0	35
Wells (0s)	2396	2	2.853	3.629	1	85
Wells (0s) pre 1966	2396	0	0.669	1.599	0	35
Dicoveries pre 1966	2396	0	0.301	0.903	0	24
D = 1 if non-dry	2396	0.667	0.562	0.416	0	1
Area	2396	923768	2921785.8	3545968.0	21621	9976139
Landlocked	2396	0	0.0280	0.165	0	1
Distance	2396	2.300	1.558	5.430	-9.950	9.950
D = 1 rhs FH	2396	1	0.654	0.476	0	1
FH	2396	0.810	0.724	0.269	0.102	1
D = 1 rhs Polity	2348	1	0.546	0.498	0	1
Polity	2348	-0.178	1.345	6.301	-10	10
D = 1 rhs Democ	2347	1	0.589	0.492	0	1
Democ	2347	3.032	4.652	3.612	0	10
D = 1 rhs Autoc	2351	1	0.517	0.500	0	1
Autoc	2351	3.571	3.307	2.823	0	10
D = 1 rhs ConsEx	2352	1	0.616	0.486	0	1
ConsEx	2352	4.250	4.514	2.117	1	7

	Distance from the border: 250 km					
	count	p50	mean	sd	min	max
Wells per pair	23352	745	1374.9	1446.7	1	4736
Wells (per 0.1 km dist-bins)	23352	1	1.807	1.752	1	85
Wells pre 1966	23352	0	0.267	0.923	0	35
Wells (0s)	23352	1	1.807	1.752	1	85
Wells (0s) pre 1966	23352	0	0.267	0.923	0	35
Dicoveries pre 1966	23352	0	.136	0.650	0	32
D = 1 if non-dry	23352	0.778	0.589	0.446	0	1
Area	23352	1098581	3100033.5	3711005.0	431	9976139
Landlocked	23352	0	0.0305	0.172	0	1
Distance	23352	9.250	6.814	100.1	-249.6	249.9
D = 1 rhs FH	23352	1	0.562	0.496	0	1
FH	23352	0.810	0.718	0.296	0.102	1

B Online appendix

B.1 Back-of-the-envelope calculation on production

To get a better sense of what the average jump at the border means in terms of oil production, we use production data reported at the field level and conduct a crude back-of-the-envelope calculation.[46] By associating wells with the producing fields that result from successful discoveries, we are able to calculate a measure of the expected daily production per exploration well.

In our baseline sample of wells within 10 km distance to the border, the calculated production per original exploration well is about 2 500 barrels per day. This is based on 503 wells, after the bottom 5% and top 5% are removed to avoid influence from outliers. The same exercise in the 250 km sample suggest an average production of 2 300 barrels per day (based on 4 336 wells).[47]

We observe in our 10 km sample of 2396 border-distance bins, 2112 wells on the left hand side and 4724 wells on the right hand side. Our estimate of 36% more drilling on the right hand side due to relatively better institutions (column five in table 2), suggests that the number of wells on the left hand side would increase with about 750 wells. Given an average production per exploration well in the 10 km sample of 2500 barrels per day, this would correspond to about 1.9 million extra barrels of oil production per day, about the production of Norway, or about 35 billion USD per year given an oil price of 50 USD per barrel.[48]

In the 250 km sample, we observe 16752 wells on the left hand side and 25435 on the right hand side of the border. We estimate in column one in table 6, moving from the left to the right, a jump of 44%. The corresponding increase in the number of wells on the left hand side is then

[46] Our field level data is drawn from the Wood Mackenzie PathFinder database (Wood Mackenzie, 2011). It contains information on producing fields, which can be associated with our various exploration wells. The dataset contains information on annual production volumes from 1965 to 2011 for both oil and gas production, in barrels of oil equivalent.

[47] Looking at the medians instead, the figures per well are 500 and 800 barrels per day for 10 km and 250 km in the developing country sample, and 800 and 1 250 barrels per day in the maximum sample of all countries. We conclude that 2 300 - 2 500 barrels per day may be reasonable, or perhaps somewhat large.

[48] The calculations are: 2112 wells on the left hand side * 0.36 = 760 extra wells. In USD, 760 wells * 2 500 barrels production per day per well * 365 days per year * 50 USD per barrel = 35 billion USD per year.

about 11 200, or 26 million barrels per day given a daily production of 2 300 barrels per day. This corresponds to 470 billion USD per year.[49]

To put these figures in perspective, the world's total oil production was 87 million barrels per day in 2013, according to BP Statistical Review of World Energy June 2014. The U.S. production was 10 million barrels per day, while the Middle East and Africa produced 28 and 9 million barrels per day, respectively. In other words, our estimates translate into a jump corresponding to about 2 % of the world's oil production in the 10 km sample and about 30% of the world's oil production in the 250 km sample.

Note that the production level per well depends on investments and is therefore endogenous. The above calculations should therefore be interpreted with this important caveat in mind.

Note also that the wells studied in this paper are exploration wells and not production wells. The field data on production necessarily correspond to the production phase rather than the preceding exploration phase. Additional (production) wells are needed to extract oil and gas, and the figures we report are therefore not informative about the total drilling necessary for a given level of production. Furthermore, we divide the production on all exploration wells, not only the non-dry ones. As shown in table A.1, about 56% of the wells drilled in our 10 km baseline sample resulted in discoveries (indicated with the dummy variable 'D = 1 if non-dry'). Therefore, our calculated figures measure the production resulting from total exploration drilling, which also best resembles our interest: the jump in production we can expect when moving from the left to the right hand side of the border.

[49]The calculations are: 25 435 wells on the left hand side * 0.44 $\approx$ 11 200 extra wells. In USD, 11 200 wells * 2 300 barrels production per day per well * 365 days per year * 50 USD per barrel = 35 billion USD per year.

B.2 Data on stable borders and maritime borders

We provide sensitivity checks across a range of different sub-samples characterised by the most recent date of border changes. To define stable borders, we utilise a new dataset with time-varying information on country shapes. The CShapes borders dataset from Weidmann et al. (2010) draws on the Correlates of War project (COW, Barbieri et al. 2008), which captures the territorial changes for the countries of the world. The dataset was refined by Gleditsch and Ward (1999) to include an extended list of independent states. The dataset uses several definitions of territorial change. First, new states can become independent, as in the case of the secession of Eritrea from Ethiopia in 1993. Second, states can merge, as in the case of the dissolution of the (Eastern) German Democratic Republic in 1990 and its ascension to the (Western) German Federal Republic. Third, we may have territorial changes to a state that do not involve the emergence or disappearance of states. CShapes includes all changes of the first two types. For the third type, the CShapes dataset only code changes that (i) affect the core territory of a state, and (ii) affect an area of at least 10,000 square kilometers. For our main specification we use the CShapes dataset taking only countries with stable borders prior to 1965. For additional checks we also use the CShapes dataset for post-1945 borders.

In addition to the CShapes stable borders we extend these cut-off dates backwards using a complementary dataset that extends further back in time. Here we utilise the TerrChange database V4.1 (Tir et al., 1998). The TerrChange database records instances of territory changes, marking them by the country who gained territory or lost territory. This does not contain GIS information on the location of territory loss, but instead records whether the country gained or lost territory, and whether this was to its homeland or dependent territories such as colonial holdings. We construct two (arbitrary) year cut-offs - countries whose border has been unchanged since the years 1925 and 1938; this allows us to generate a reasonably sized sub-sample while going back further in time reduces our sample size considerably.

We construct the sub-samples by calculating the most recent border change per country as either a gain or loss to that countries homeland territory (we exclude changes to its dependent territories). We then include only those countries who have data on them in the database, and had a border change prior to the year stated (this is a conservative approach as we do not include missing countries, even though they might be missing from the dataset due to no border changes since the start of the dataset in 1816). The lists of countries we end up with is presented in table B.1. The estimation results for the two sub-samples limited to these countries are presented in table 7. Note that other data limit the actual samples in our setting to 8 and 7 home countries for 1938 and 1925, respectively.

Table B.1: Countries with stable borders since 1938 and 1925

	1938 (n=20)	1925 (n=18)
1	Afghanistan	Afghanistan
2	Australia	Australia
3	Bolivia	
4	Brazil	Brazil
5	Costa Rica	Costa Rica
6	Cuba	Cuba
7	Denmark	Denmark
8	Dominican Republic	Dominican Republic
9	Ireland	Ireland
10	Liberia	Liberia
11	Monaco	Monaco
12	Mongolia	Mongolia
13	New Zealand	New Zealand
14	Norway	Norway
15	Paraguay	Sweden
16	Sweden	Thailand
17	Thailand	
18	United Kingdom	United Kingdom
19	Uruguay	Uruguay
20	Venezuela	Venezuela

We also use offshore maritime borders in our main specification. For this we utilise the the EEZ Maritime Boundaries Geodatabase version 6.1 (Claus et al., 2013). EEZ Maritime Boundaries are a relatively recent legal phenomena. In the 1970s the UN started systematic efforts to establish offshore national boundaries, resulting in the United Nations Convention on the Law of the Sea United Nations (1982). Legally defined offshore borders for exploitation of sub-sea minerals date back to 1945 when the United States established the principle of sovereignty extending up to the limit of the continental shelf. This approach was subsequently adopted by most countries, and replaced freedom of the seas rights originating in the seventeenth century, such as those extending three nautical miles from a nations coastline - the cannon shot rule. The UN Law of the Sea (United Nations 1958) codified in international law the position of adjacent maritime borders (the median line boundary) and the continental shelf limit (i.e. where international waters begin). A subsequent conference, UNCLOS III 1973-1982, refined the definition of how far sovereign rights extend vis-a-vis international waters to give us present-day Exclusive Economic Zones (EEZ) (United Nations, 1982). This later conference also updated the principle of defining adjacent (or opposite) national boundaries, but in practice this would rarely imply deviation from the 1958 median line approach. Resulting EEZs typically extend 200 nautical miles from a nations coastline. For our identification we do not exploit the precise limit of sovereign waters versus international waters but instead the position of maritime borders vis-a-vis a countrys neighbour. While inherently more contestable than onshore borders, the position of these borders typically pre-dates oil exploration due to median line extrapolation agreed in 1958 and earlier bilateral agreements. Furthermore, the exploration for and exploitation of oil and gas beyond shallow coastal waters only began in 1979 with the first deep-water drilling technology deployed to the Gulf of Mexico.

B.3 Extra graphs and tables

Below we present extra graphs and tables referred to in the text. Figure B.1 presents drilling along the border between US and Mexico in the Gulf of Mexico. Figure B.2 shows, around our borders, the distribution of random points landing inside geological basins, as explained in section 4.1. Table B.2 is discussed in the text in section 3.2 and presents estimates based on log wells as dependent variable as well as the Poisson and NB estimators. In table B.3 we investigate robustness of our inference to spatial correlation across observations, and the results hold. Table B.4 repeats table 8 for 250 km and all the coefficients are now significant. Table B.5-B.6 relate to section 4.3 and are described in more detail below. Table B.7 presents the countries included in our samples. Figure B.3 shows a map of geological basins, where oil may be found, together with onshore and offshore national borders.

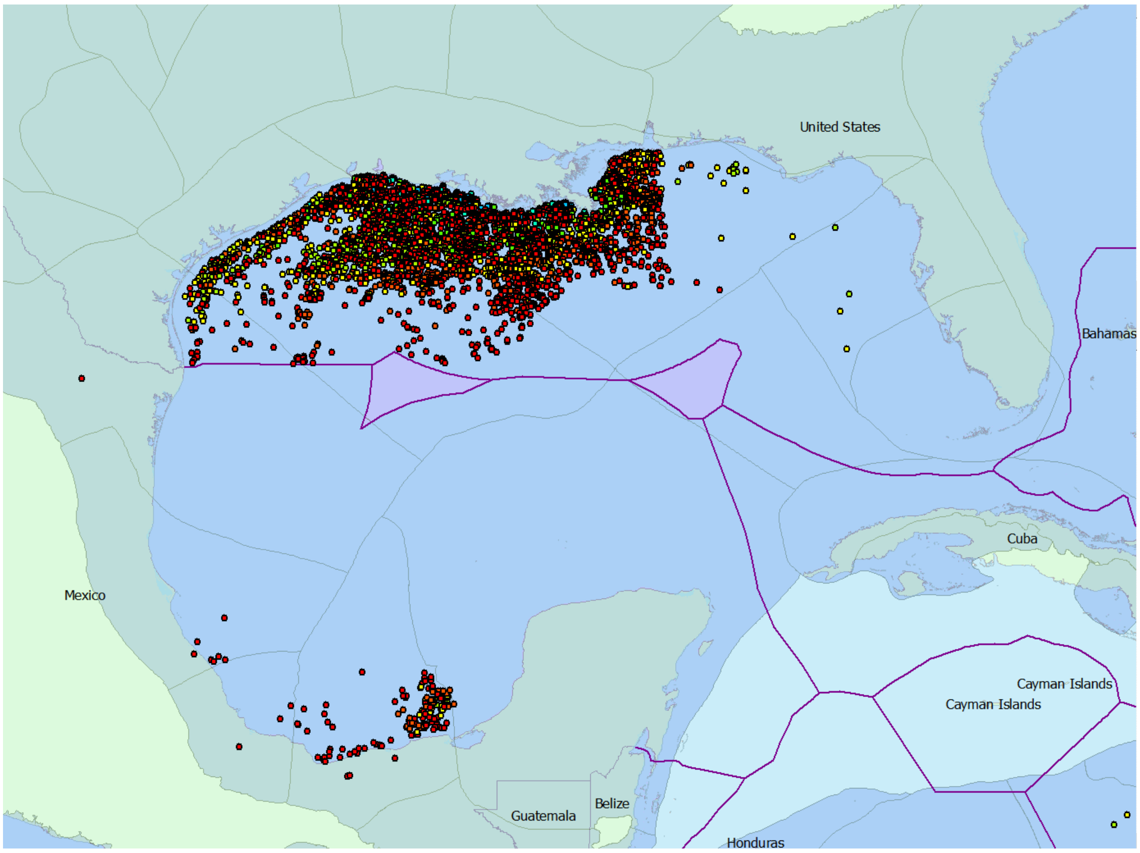

Figure B.1: Drilling at the US-Mexico border in the Gulf of Mexico

Notes: Oil and gas exploration drilling at US-Mexico border in the Gulf of Mexico for all years before 2011. The different colors indicate when the wells were drilled: Blue (pre-1950), Teal (1950s), Green (1960s), Lime green (1970s), Yellow (1980s), Orange (1990s), Red (2000s).

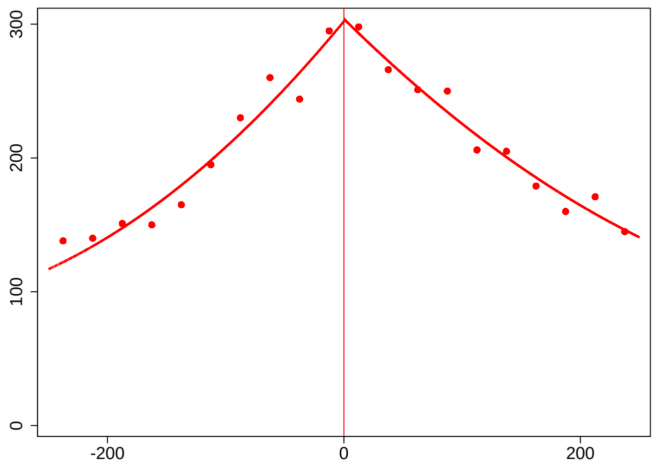

Figure B.2: Random points hitting inside basins

Notes: Random points generated as explained in section 4.1. Sample restricted to our standard 250 km sample.

Table B.2: Comparison OLS, Poisson and NB

	First order polynomial			
Dependent variable	(1)	(2)	(3)	(4)
and estimator →	N OLS	ln N OLS	N Poisson	N NB
D = 1 rhs	1.016***	0.189***	0.257**	0.269***
	(0.386)	(0.053)	(0.105)	(0.096)
Mean L	3.60	2.30	3.76	3.72
Scaled L	0.28			
exp(β)-1		0.21	0.29	0.31
Mean Y	2.85	0.67	2.85	2.85
Var Y	13.17	0.61	13.17	13.17
alpha			0.34	0.34
LR alpha ($\bar{\chi}^2$)			1308.32	
LR alpha (p)				
Observations	2396	2396	2396	2396
R-sq	0.11	0.19		

	Second order polynomial			
Dependent variable	(1)	(2)	(3)	(4)
and estimator →	N OLS	ln N OLS	N Poisson	N NB
D = 1 rhs	1.490**	0.216***	0.361***	0.380***
	(0.617)	(0.083)	(0.138)	(0.125)
Mean L	4.12	2.38	4.26	4.20
Scaled L	0.36			
exp(β)-1		0.24	0.43	0.46
Mean Y	2.85	0.67	2.85	2.85
Var Y	13.17	0.61	13.17	13.17
alpha			0.32	0.32
LR alpha ($\bar{\chi}^2$)			1244.68	
LR alpha (p)			0.00	
Observations	2396	2396	2396	2396
R-sq	0.12	0.19		

Notes: identical to table 2, except dependent variable or estimation method, as indicated. Polynomials in distance to border included as indicated.

Table B.3: Standard errors adjusted for spatial correlation

Dependent variable →	N	
	(1)	(2)
Distance to the border →	10 km	250 km
D = 1 rhs	1.490*	0.847***
	(0.852)	(0.316)
Mean L	4.12	1.91
Scaled L	0.36	0.44
Observations	2396	23352
Countries	50	70
Neighbours	76	96
R-sq	0.46	0.55

Notes: Column one as column five in table 2 and column two as column one in table 6, except we allow for spatial correlation between observations by employing the procedure programmed by Hsiang (2010), which builds on Conley (1999). Standard errors are robust to heteroskedasticity. To measure the distance between observations, we use the average longitude and latitude of the wells drilled in a given country-border specific distance-bin. The distance cutoff is set to 10 km and 250 km, the same as the distance to the border. We use Hsiang's default kernal to weight spatial correlations. For more details about the procedure, see http://www.fight-entropy.com/2010/06/standard-error-adjustment-ols-for.html.

Table B.4: Restricting sample to "artificial states"

Dependent variable →	N: Number of wells			
	(1)	(2)	(3)	(4)
Sample →	SSA	Africa	PART	FRAC
Distance from border →	250 km			
D = 1 rhs	3.980***	3.122***	5.519***	1.077***
	(0.767)	(0.508)	(1.040)	(0.134)
Mean L	1.12	1.07	1.16	1.76
Scaled L	3.55	2.93	4.75	0.61
Observations	803	2237	1036	3479
Countries	12	16	8	19
Neighbours	17	22	11	21
R-sq	0.12	0.11	0.12	0.22
Clusters (dbin)	649	1671	811	2740

Upper panel 10 km, lower panel 250 km. Columns: 1 SSA, 2 Africa. 3 Those countries in the top third of partitioning (i.e., third of countries with most artificial borders based on ethnic groups); 4 Those countries in the bottom third of fractal measure (i.e., third of countries with straightest borders). Measures used in column 3 and 4 from Alesina et al. (2011). Wells drilled after 1966, FH measured before 1966. Second-order polynomial and area and landlocked controls are included. Otherwise as described by the note of table 2.

Table B.5 is described in the text in section 4.3. Table B.6 includes the control-dummies in different combinations. The logic follows the accounting identity in terms of the production function: $Y = F(L,H,K,A)$. Including Y as a control takes into account L, H, K, and A, which are columns 1 and 2 in table 10. We therefore include only combinations of maximum three out of the four variables L, H, K and A in table B.6. The first column includes the three factors L, H and K. Then the next three columns include different combinations of those three factors. In the three next columns we allow A to substitute each of the factors, L, H and K. In other words, in the four first columns, institutions are allowed to work via A. In the three next columns, institutions are allowed to work via L, H and K, respectively. Across the different specifications, the coefficient on the institutions-based border dummy is stable and statistically significant at the ten percent level in all but one of the seven cases. The estimated relative jump is 27%-29%, hardly reduced from the 30% jump estimated in the model without the proximate causes (column 6 in table B.5).

Table B.5: Baseline regression on PWT-samples

	(1) Wells	(2) Wells	(3) Wells	(4) Wells	(5) Wells	(6) Wells
D = 1 rhs	2.314***	2.314***	1.456**	2.313***	2.314***	1.438*
	(0.708)	(0.708)	(0.738)	(0.708)	(0.708)	(0.779)
Mean L	4.31	4.31	4.40	4.31	4.31	4.73
Scaled L	0.54	0.54	0.33	0.54	0.54	0.30
Observations	2004	2004	1795	2001	2004	1659
Countries	42	42	39	41	42	37
Neighbours	56	56	51	55	56	47
R-sq	0.13	0.13	0.11	0.13	0.13	0.11
Clusters (dbin)	200	200	200	200	200	200

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. Table presents estimates of the baseline model based on the samples used in table B.5. Otherwise as described by the note of table 2.

Table B.6: Baseline regression, controlling for border dummies defined on proximate causes of economic growth

	(1) Wells	(2) Wells	(3) Wells	(4) Wells	(5) Wells	(6) Wells	(7) Wells
D = 1 rhs	1.257	1.286*	1.382*	1.337*	1.334*	1.268*	1.288*
	(0.767)	(0.775)	(0.766)	(0.767)	(0.759)	(0.769)	(0.759)
D = 1 rhs L	-1.229***	-1.258***		0.253*		-1.052***	0.109
	(0.288)	(0.286)		(0.148)		(0.293)	(0.143)
D = 1 rhs H	0.088		0.206*	0.379***	0.194*		0.312***
	(0.110)		(0.114)	(0.119)	(0.113)		(0.114)
D = 1 rhs K	1.945***	2.002***	0.774***		0.523***	1.580***	
	(0.284)	(0.284)	(0.140)		(0.131)	(0.285)	
D = 1 rhs A					0.865***	0.786***	0.967***
					(0.113)	(0.116)	(0.117)
Mean L	4.73	4.73	4.73	4.73	4.73	4.73	4.73
Scaled L	0.27	0.27	0.29	0.28	0.28	0.27	0.27
Observations	1659	1659	1659	1659	1659	1659	1659
Countries	37	37	37	37	37	37	37
Neighbours	47	47	47	47	47	47	47
R-sq	0.13	0.13	0.12	0.12	0.13	0.14	0.13
Clusters (dbin)	200	200	200	200	200	200	200

Notes: Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. The additional border dummies are defined according to the bilateral rank of the variables indicated in the column headings, based on averages of the values observed in the period 1950-1965. The data are from the Penn World Tables (PWT) and described in section 4.3. The sample is restricted to be the same in all columns. Otherwise as column 5 in table 2.

Table B.7: Home countries and neighbours included for 10 km and 250 km

			Home country				Neighbouring country			
			10 km		250 km		10 km		250 km	
Region	**Country**	**ISO**	**Obs.**	**Wells**	**Obs.**	**Wells**	**Obs.**	**Wells**	**Obs.**	**Wells**
EAP	Cambodia	KHM					7	8	177	263
EAP	China	CHN			3	3	26	27	21	24
EAP	Indonesia	IDN	119	649	107	546	1267	2537	1085	2038
EAP	Malaysia	MYS	23	26	119	651	506	604	855	1904
EAP	Myanmar	MMR	7	7	48	77	131	156	312	440
EAP	Philippines	PHL	18	21	17	19	56	64	112	135
EAP	Thailand	THA	15	16	4	4	440	614	308	371
ECA	Albania	ALB	3	3	58	73	10	10	171	207
ECA	Armenia	ARM			2	2			14	14
ECA	Belarus	BLR			1	1			87	91
ECA	Bulgaria	BGR			44	58	3	3	260	365
ECA	Georgia	GEO			1	1			8	8
ECA	Kazakhstan	KAZ					13	13	5	24
ECA	Moldova	MDA			4	7			195	289
ECA	Montenegro	MNE			1	1			4	5
ECA	Romania	ROM	32	49			576	879		
ECA	Tajikistan	TJK			2	19			16	66
ECA	Turkey	TUR	93	176	8	9	378	575	17	18
ECA	Turkmenistan	TKM			41	154			219	801
ECA	Ukraine	UKR			63	78			530	711
ECA	Uzbekistan	UZB	45	175			241	892	11	11
HI	Australia	AUS	173	670	26	45	1270	2362	151	210
HI	Austria	AUT	8	8	8	8	58	59	37	39
HI	Barbados	BRB					1	2		
HI	Belgium	BEL			32	41			445	628
HI	Canada	CAN	95	380	92	417	1599	3504	317	746
HI	Cyprus	CYP			2	2			2	2
HI	Czech Republic	CZE			7	7			289	356
HI	Denmark	DNK	4	4	4	5	196	234	167	194
HI	Finland	FIN							54	62
HI	France	FRA	25	37	45	50	478	628	291	353
HI	Germany	DEU	30	38	152	367	242	315	1184	1766
HI	Greece	GRC	65	82	115	228	88	105	291	487
HI	Hungary	HUN			14	26	1	1	76	139
HI	Ireland	IRL	1	1	5	6	146	163	63	73
HI	Israel	ISR	52	80	2	2	65	95	2	2
HI	Italy	ITA	97	146	41	57	423	543	161	201
HI	Japan	JPN							26	27
HI	Luxembourg	LUX							3	3
HI	Netherlands	NLD	117	298	14	20	948	1505	741	1105
HI	New Zealand	NZL	26	45	74	133	150	209	371	543
HI	Norway	NOR					890	1256	1325	2933
HI	Poland	POL	100	155	5	6	1171	1473	11	12
HI	Portugal	PRT	15	24	6	8	43	62	103	116
HI	Saudi Arabia	SAU			4	4	3	3	14	14
HI	Singapore	SGP			14	16			509	745
HI	Spain	ESP	48	53	38	59	360	431	240	305
HI	Sweden	SWE			1	1			210	251
HI	Switzerland	CHE			3	3			185	229
HI	Trinidad and Tobago	TTO							1	2
HI	United Kingdom	GBR	16	17	1	1	1954	3865	978	1379
HI	United States	USA	154	640	95	380	2050	5482	1600	3505
LAC	Argentina	ARG	125	329	58	78	1564	3070	330	426
LAC	Bolivia	BOL	12	14	116	331	326	374	505	1213
LAC	Brazil	BRA	192	781	3	3	1303	3019	193	206
LAC	Chile	CHL	46	64	94	292	79	105	1381	2862
LAC	Colombia	COL	28	32	15	19	391	540	143	162
LAC	Ecuador	ECU	32	38	64	79	303	346	439	593
LAC	Guatemala	GTM			16	18			144	157
LAC	Guyana	GUY	2	2			3	3	4	4
LAC	Jamaica	JAM					2	2		
LAC	Mexico	MEX	16	18	62	223	145	158	1733	4736
LAC	Panama	PAN			9	10			54	59
LAC	Paraguay	PRY			104	484			864	1863
LAC	Peru	PER	46	58	18	20	225	242	249	277
LAC	Uruguay	URY			2	2			34	36
MENA	Algeria	DZA	35	37	66	80	749	832	504	572
MENA	Iraq	IRQ			3	4	5	5	6	9
MENA	Jordan	JOR	8	9	26	32	18	19	35	42
MENA	Lebanon	LBN			26	48			30	53
MENA	Libya	LBY	17	18	38	48	476	520	410	459
MENA	Morocco	MAR	17	20	1	1	207	251	63	63
MENA	Tunisia	TUN	102	137	44	48	399	492	678	766
SA	Afghanistan	AFG			2	2			2	2
SA	India	IND	74	124	8	8	204	285	71	78
SA	Nepal	NPL					1	1	4	4
SA	Sri Lanka	LKA			35	56	2	2	59	88
SSA	Angola	AGO					5	5		
SSA	Benin	BEN	2	2	83	472	11	11	227	665
SSA	Cameroon	CMR	41	51	98	782	127	161	612	2206
SSA	Central African Republic	CAF							14	15
SSA	Chad	TCD			3	3	38	40	31	32
SSA	Ethiopia	ETH			1	1	7	7	7	7
SSA	Ghana	GHA	8	9			11	12	6	6
SSA	Kenya	KEN	2	2	8	8	23	24	17	17
SSA	Liberia	LBR					4	4	3	4
SSA	Malawi	MWI			1	1			7	7
SSA	Mali	MLI							3	3
SSA	Mauritania	MRT					2	2	2	2
SSA	Mozambique	MOZ	6	6	11	19	19	20	18	27
SSA	Niger	NER	2	2			26	27	11	12
SSA	Nigeria	NGA	182	1255	43	53	836	2868	138	172
SSA	Sierra Leone	SLE					3	4	4	4
SSA	Somalia	SOM	2	2	2	2	9	9	24	25
SSA	South Africa	ZAF			3	3			12	13
SSA	Sudan	SDN							2	2
SSA	Tanzania	TZA	18	26	2	2	33	42	4	4
SSA	Togo	TGO			8	9	4	4	12	13
SSA	Uganda	UGA							2	2
SSA	Zambia	ZMB					2	2	7	7
Total			**2396**	**6836**	**2396**	**6836**	**23352**	**42187**	**23352**	**42187**

Notes: Table presents number of observations and number of wells within 10 km or 250 km of the border, summed either per home country (the country with the well) or by the neighbouring country (the country nearest to the well). Based on the Freedom House (*FH*) measure of institutional quality. World Bank regions: East Asia & Pacific (EAP), High Income (HI), Latin America & Caribbean (LAC), Middle East & North Africa (MENA), South Asia (SA), Sub-Saharan Africa (SSA).

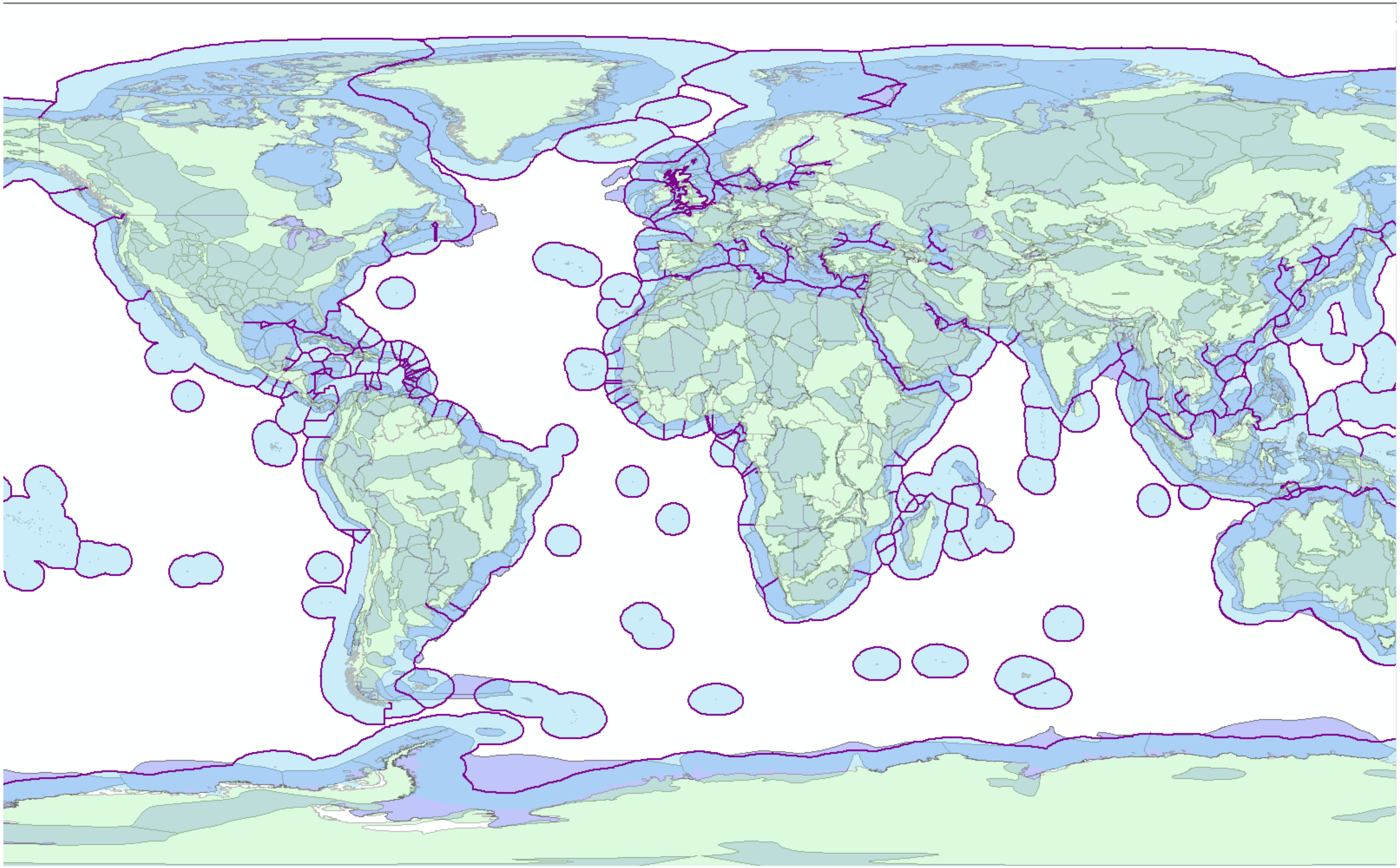

Figure B.3: National borders and geological basins

Notes: Map shows onshore and offshore national borders together with geological basins, structures in which oil can be found, marked as dark areas.