The Economic Effects of Resource Extraction in Developing Countries

James Cust

Worcester College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Economics

Michaelmas 2014

Acknowledgements

I would like thank my supervisor, Tony Venables, for all the guidance and support throughout the preparation of this dissertation. I would also like to acknowledge the various teachers and mentors from whom I have learned a great deal, and without I whom could not have completed this work. In particular John Challice, Paul Collier, Guy Durden, Marcel Fafchamps, Rick van der Ploeg and Francis Teal. I must also thank my co-authors, Torfinn Harding, Steven Poelhekke and Ridwan Rusli, who, in addition to collaborating on parts of the work presented herein, have also served as a source of wisdom and inspiration.

Thanks to all the friends and colleagues who have helped me throughout the DPhil in so many ways. And most of all thanks to my family, who's unwavering support cannot be quantified: Mum, Dad, Lucy, Flora and Henry.

Abstract

This thesis presents three core chapters examining different aspects of the relationship between natural resources and economic development. While addressing different questions they share several features in common: a concern with causal inference; overcoming the challenges of endogeneity between resource abundance and other characteristics of developing countries; and the use of new and novel datasets with spatially identified units of analysis. The work contributes to a rich and growing empirical literature seeking to deepen our understanding of the underlying mechanisms affecting the fortunes of resource-abundant countries.

In the introductory chapter I discuss the extensive literature on this topic and in particular focus on the new generation of well-identified within-country studies, seeking to understand the empirical relationship between resources and economic development. Countries typically welcome the news of a resource discovery with joy and indeed, resource discoveries hold great economic potential. But what determines whether a country is resource rich or not? Is it more than just a chance finding, or good geology? In Chapter 2, entitled Institutions and the Location of Oil Exploration I present an investigation into this question. I examine the relationship between governance and choices of where to drill for oil. This work utilises a new dataset on exploration wells and looks at the distribution of drilling close to national borders. This allows me to identify estimates for the effect of differences in governance between neighbours. Two times out of three, investors choose to drill on the side of borders that are better governed, all other things being equal. This suggests that resource-wealth itself may be contingent on factors beyond geology, and indeed may be endogenous to the process of development.

In Chapter 3, entitled The Local Effects of Resource Extraction, I turn my attention to the local economic consequences of industrial mining in Indonesia. I present a simple three-sector general equilibrium model to generate predictions for the local labour market, akin to the Corden-Neary Dutch disease model of the macroeconomy. I test the predicted effects in response to an exogenous resource sector shock by looking at mine opening or mine expansion events across three hundred mines. I test the predictions of the model, first by estimating the economic footprint from

industrial mining; found to be an average of fifteen kilometre radius. I then examine the response of reported labour market activity from households surveyed in nearby communities. Here I find no evidence for a shift of local labour into the mining sector. I do find however a notable movement of labour from the traded sectors (agriculture and manufacturing) to the non-traded service sector, with a strong effect for foreign-owned mines versus domestic ones.

Chapter 4, entitled Disentangling the Effects of Resource Extraction: Local Government and Investment Multipliers, examines the oil and gas boom in Indonesia from 1999-2009. Here I deploy a variety of identification strategies to attempt to disentangle the regional effects of the boom, measured in terms of district GDP. I estimate effects arising from transfers of revenue to local government. Using an instrumental variable approach I isolate the fiscal channel from resource projects. I find a positive and significant effect of increased local government revenues on district GDP over the boom decade. I then examine the spillovers from resource projects, isolating them from fiscal transfers. For districts neighbouring resource rich districts I find evidence for a modest positive effect arising from project investments, rather than fiscal transfers.

In Chapter 5 I present concluding thoughts and discuss a future research agenda. I also summarise the burgeoning landscape of resource data available for within-country and spatially identified studies and offer some thoughts on how this might evolve.

Contents

1 Introduction **1**
1.1 Overview of my research . . . 3
1.2 Literature . . . 9
1.2.1 Resource wealth empirical challenges . . . 10
1.2.2 Evaluating the local effects of resources on the economy . . . 12
1.2.3 Regional spillovers and Fiscal decentralization . . . 20
1.2.4 Fiscal decentralization . . . 24
1.2.5 Reflections . . . 28

2 Institutions and the Location of Oil Exploration **29**
2.1 Introduction . . . 31
2.2 Data and empirical strategy . . . 37
2.2.1 Data . . . 37
2.2.2 Empirical strategy . . . 39
2.3 Baseline results . . . 42
2.3.1 Graphical evidence of discontinuity at the border . . . 43
2.3.2 Econometric estimates of the effect of crossing the border . . . 45
2.3.3 Scaling the effect of the border with institutional quality . . . 46
2.3.4 Institutions versus geography and culture . . . 49
2.4 Validity of the identification strategy . . . 51
2.4.1 The likelihood of discovering oil and path dependency . . . 52
2.4.2 The role of offshore borders, military disputes and unobservables . . . 54
2.4.3 Reverse causality . . . 57
2.4.4 External validity . . . 58
2.5 Heterogeneity across company types . . . 60
2.6 Conclusions . . . 62
2.7 Summary statistics . . . 62
2.8 Appendix . . . 63

2.8.1 Specification analysis 63
2.8.2 Extra tables and descriptive statistics 68

3 The Local Effects of Resource Extraction 72
3.1 Introduction 74
3.1.1 Literature 75
3.2 Analytical Framework 78
3.2.1 Set-up 78
3.2.2 Equilibrium 81
3.2.3 Effects of a productivity shock in resource sector 82
3.2.4 Role of local spending spillovers in local Dutch disease effects 84
3.2.5 Research Hypotheses 85
3.3 Data 87
3.3.1 Data Description 88
3.3.2 Construction of treatment variables 94
3.4 Empirical Methodology 95
3.4.1 Identification Strategy 96
3.4.2 Estimating labour market dynamics 100
3.4.3 Estimating Mine Footprint 102
3.5 Results 102
3.5.1 Estimating size and intensity of the mining footprint 103
3.5.2 Mine closure: estimating the bust versus the boom 108
3.5.3 Estimating the labour market effects 113
3.5.4 Estimating the effect of foreign versus domestic ownership 116
3.6 Robustness tests 118
3.6.1 Alternative specifications 118
3.6.2 Employment and out-migration 123
3.7 Conclusion 124

4 Disentangling the effects of resource extraction: quantifying local government and project multipliers 126
4.1 Introduction 128
4.1.1 Oil and gas in Indonesia 130
4.1.2 Indonesia's decentralization and government revenue sharing scheme . 133
4.1.3 Oil and gas contracts and government revenue sharing 135
4.2 Conceptual framework 138
4.3 Data 141
4.3.1 Gross Regional Domestic Product (GRDP) 143

4.3.2 Government revenues . . . 143
4.4 Empirical strategy . . . 145
4.4.1 Identification of fiscal channel effects . . . 146
4.4.2 Identification of project channel effects via neighbour spillovers . . . 149
4.5 Results . . . 151
4.5.1 Fiscal channel effects . . . 151
4.5.2 Project channel effects via neighbour spillovers . . . 154
4.6 Robustness and Validity of the identification strategy . . . 157
4.6.1 Validity of the instrumental variable . . . 157
4.6.2 Fiscal equalization . . . 159
4.6.3 Robustness checks and channels . . . 160
4.6.4 Difference-in-difference estimation . . . 161
4.6.4.1 Effects on sector-specific GDP . . . 163
4.6.4.2 Estimating the neighbour spillover from gas versus that from oil . . . 164
4.7 Conclusions . . . 165
4.8 Indonesia oil and gas sector background . . . 169

5 Conclusions **171**
5.1 Future research directions and research agenda . . . 171
5.1.1 Oil investment and governance . . . 171
5.1.2 Resource extraction and development . . . 173
5.1.3 Data . . . 174

Bibliography **178**

List of Figures

1.1 Number of wells in developing countries 5

2.1 National borders and exploration wells- excluding data for onshore US, and much of the former USSR. 38
2.2 Sequence of events in our sample 40
2.3 Number of wells in developing countries 43
2.4 Map of the Albert Rift region along the Uganda-DRC border. 44
2.5 Cell data: Area size and number of wells 67
2.6 National borders and geological basins 71

3.1 Illustrative labour market response (resources *LR* and non-traded *LN* sectors). (The traded sector - not shown - contracts proportionately to the expansion of the expanding sectors, since labour supply is fixed at L) 82
3.2 Illustrative effect on the market for non-traded goods and services (*N*) 83
3.3 Map of mines and communities (green/grey=mines, red/dark=survey communities) 92
3.4 Mapping layers of a illustrative coal production area (with distance bands, 5km, 10km and 15km shown in black) 92
3.5 Plot of the distribution of observations in terms of distance (in km) from the nearest open mine 105
3.6 Plot of the proportion of mines open within the 25km distance radius of the community, by year 105
3.7 Batu Hijau Copper and Gold mine. Estimating the economic footprint of mining activity: total change 1995-2010. Distance bands shown in black (in km). The two left-most areas of effect are two ports, while the right-most is the mine itself. The middle concentration of light represents a mining-related settlement. 106
3.8 Plot of density of nighttime lights for open and closed mines. Within 15km band. Y-axis shows the density, X-axis the digital light value within the distance band. 106

3.9 Estimating the economic footprint of mining activity: The closing of Kelian gold mine, pre- versus post- 1995 (date of mine closure) 110
3.10 Plot of density of nighttime lights for mine and non-mining communities, defined by 15km band. Y-axis shows the density, X-axis the digital light value within the distance band. 112

4.1 Production and investment in the oil and gas sector 131
4.2 Petroleum output value and total subnational government revenues, including transfers 132
4.3 Distribution of government petroleum revenue transfers 137
4.4 The location of oil and gas wells in Indonesia 142
4.5 Neighbour identification strategy: Illustration of revenue sharing rules in petroleum producing province 149
4.6 Oil production and government revenue 152
4.7 Oil field investment and production 152
4.8 General allocation fund and oil revenues 160

Chapter 1

Introduction

James Cust[1]

> *Of all those expensive and uncertain projects, however, which bring bankruptcy upon the greater part of the people who engage in in them, there is none perhaps more perfectly ruinous than the search after new silver and gold mines. It is perhaps the most disadvantageous lottery in the world, or the one in which the gain of those who draw the prizes bears the least proportion to the loss of those who draw the blanks: for though the prizes are few and the blanks are many, the common price of a ticket is the whole fortune of a very rich man.*

Adam Smith, The Wealth of Nations, 1776 Book IV, Chapter VII, Part I: Of the Motives for Establishing New Colonies

When a country discovers natural resources, the reaction is often one of exuberance. Resource wealth can herald increased government revenues, inward investments, and the opportunity for accelerated development, especially in low income countries. However, there exist many pitfalls separating a newly resource-rich country from the transformative economic development promised by resource wealth.

The relationship between natural resource abundance and economic performance has been the subject of a long and rich tradition in the literature. Some of early pioneers of modern economics such as David Ricardo and Adam Smith were concerned with both the positive and

[1] Part of this chapter presents a review of the literature on local economic impacts of resource extraction. This review has been prepared for inclusion in a longer survey article authored jointly with Steven Poelhekke, VU University Amsterdam.

negative potential of resource wealth. More recent scholars have revisited this topic and sought to understand the ways in which natural resources can both help and hinder national economic performance (Auty 2001, Sachs and Warner 1995).

The resource curse hypothesis that resource abundance is negatively associated with long-term economic growth has been scrutinized in great detail and linked to a wide-range of economic and political phenomena. The literature has examined a range of potential mechanisms including appreciation of the real exchange rate and de-industrialization, volatility, and consequences related to political institutions, such as deterioration in governance, corruption, rent-seeking and conflict. A further argument posits that resource-rich countries are unable to convert their depleting resources into other productive assets for reasons of policy failure, myopia and time-inconsistency (see van der Ploeg 2011 for a recent survey of the literature).

Whether it is fair to characterize natural resource wealth as a curse is still debated. Most of the evidence for the above mentioned channels derives from cross-country and macroeconomic analyses, providing evidence both for and against a potential detrimental effect on growth. Fuelling this debate is not only the mixed evidence, but also concerns regarding endogeneity of the relationship between observed resource wealth and economic or political factors at the cross-country level. Therefore, scholars are increasingly turning to within-country studies for greater disaggregation of economic responses and exogenous identification of impacts.

This thesis presents several contributions to this debate. First I review these debates and describe both the rationale for within-country studies but also the endogeneity challenges associated with cross-country analysis.

In Chapter 2, entitled *Institutions and the Location of Oil Exploration*, I present an investigation into the question of what the causal factors behind resource exploration, and hence discovery, might be. I examine the relationship between governance and choices of where to drill for oil utilising a new dataset on exploration wells around the world.

In Chapter 3, entitled *The Local Effects of Resource Extraction*, I turn my attention to the local economic consequences of industrial mining in Indonesia, to estimate the size of the economic footprint of industrial mines, and examine the labour market effects of new and expanding mines.

Chapter 4, entitled *Disentangling the Effects of Resource Extraction: Local Government and Investment Multipliers*, examines the oil and gas boom in Indonesia from 1999-2009 to look a both the district GDP effects of increased local government revenues from oil and gas extraction, and the project spillover effects on neighbouring districts.

In Chapter 5 I present some concluding thoughts; I reflect on the work presented and discuss a future research agenda. I also summarise the burgeoning landscape of resource data available for within-country and spatially identified studies and offer some thoughts on how this might evolve.

1.1 Overview of my research

Our empirical understanding of the relationship between resource wealth and economic development has advanced considerably in recent years. Much of the attention for the last three decades had been focused at the national aggregate and cross-country level, however attention in the last few years has shifted. In part motivated by methodological and identification concerns, examination of the spatially specific and within-country effects are now becoming widespread. The rigour of these studies has increased, driving a better understanding of the causal relationships, however within-country dynamics are also important to help us understand whether these channels constitute important transmission mechanisms for the macro resource curse symptoms.

Examples include the possibilities for resource wealth in concentrated geographical locations to drive divergent regional economic performance, allocating the costs and benefits of extraction asymmetrically. Furthermore, governments may choose to allocate the costs and benefits differently within countries, such as through revenue sharing rules or targeted expenditure programs. This may further exacerbate regional inequalities. Finally, many of the hypothesized channels of resource benefits and costs are contingent on factors at the subnational level: rent-seeking by local politicians, myopia and time-consistency, regional favoritism, volatility, real exchange rate appreciation and price differences across space, as well as uneven investments in infrastructure, can all add up to a positive or negative national level experience. While some of the causes of the resource curse may be avoidable via actions by the national government, other factors arise locally or as a consequence of local governments actions.

My own work falls within this strand of the literature. Motivated to gain a deeper and more rigorous understanding of the causal factors at work, and armed with new data sources and methodologies, this is a potentially fruitful avenue of investigation. My main contribution to the literature is to use geographically specific data to identify causal relationships and shed light on several important aspects of the potential relationship- the links between institutions and resources, the impact of resources at the local and regional level, and the role of subnational government transfers of resource revenues.

Chapter 2 Institutions and the location of oil exploration: Chapter 2, entitled *Institutions and the location of oil exploration*, is chronologically the most recent of the three main chapters, however exemplifies the nascent potential of the new datasets. Data on the precise location and drill dates of exploration wells can help us shed light on questions specific to resources- such as understanding the incentives and process by which exploration takes place- but can also enlighten us on questions of more general interest. For example, it is widely thought that institutional quality, or the governance environment of countries, is a key factor in driving economic development (Acemoglu et al., 2005). However it has proven fiendishly difficult to identify and estimate given the fact that economic development and institutions are thought to be equilibrium outcomes, and thus endogenously determined.

By linking the location of wells to narrow strips along national borders I test for spatial discontinuity of drilling with respect to borders. This allows me to identify estimates for the effect of differences in governance between neighbours. Two times out of three, investors choose to drill on the side of borders that are better governed, all other things being equal. This suggests that resource-wealth itself may be contingent on factors beyond geology, and indeed may be endogenous to the process of development.

This approach most closely follows the work of Michalopoulos and Papaioannou (2013), who examine the spatial discontinuity in light emissions at national borders. Instead of a spatial discontinuity in light, the spatial discontinuity in the distribution of exploration wells gives us a clue as to the incentives faced by investors. For narrow strips of land close to national borders, where, arguably geology can be treated as identical *ex ante*, investors choose to drill on the better governed side of the border. Figure 1.1 illustrates this relationship with a plot of the spatial discontinuity in drilling at borders (where wells are placed on the right hand side for better governed countries versus their nearest neighbour).

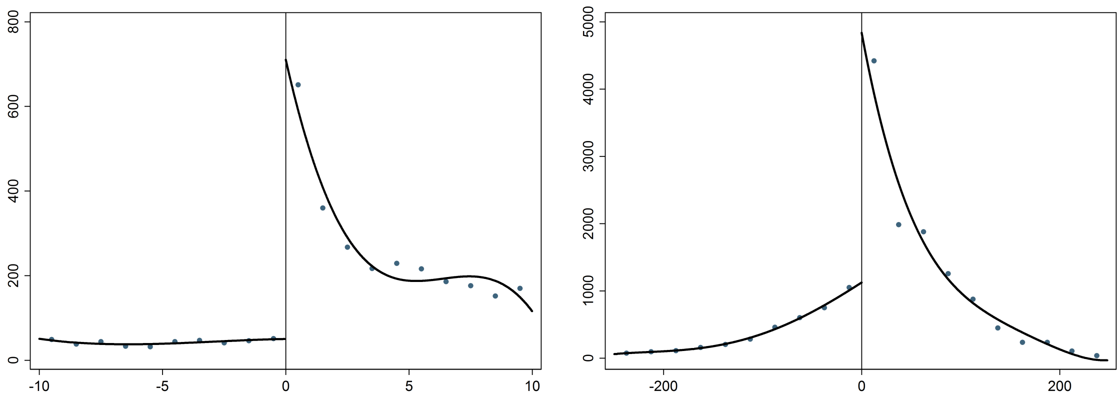

Figure 1.1: Number of wells in developing countries

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. Wells drilled in the better governed side are placed on the right-axis, those drilled on the worse, the left.

Naturally we are concerned about the identification strategy to estimate the causal relationship between institutions and drilling. First, we may worry that border location is endogenous to oil. This may result from them being drawn to secure oil deposits, or wars carried out to adjust borders to capture oil-rich regions - and that doing so works systematically in favour of better governed states. In Chapter 2, I explore a variety of concerns we might have and offer evidence to support the idea that border placement is generally not linked to oil. Second, we may be concerned with our measures of institutions, or indeed what institutional mechanism is driving our effect. This is admittedly a limitation of the paper.

By using broad (and the most widely used) measures of institutions, namely Freedom House and Polity IV, I impose a rather crude approximation of the complexity of institutional environment - but the one favoured in the literature nonetheless. I use measures that reflect democratic rights and constraints on the executive. As noted in Chapter 2, it is not the objective to unbundle the various and complex components of institutions, but rather to present some broad and well-identified measures for an average effect. Indeed it may be somewhat surprising that such a crude measure of institutions generates such an unambiguous effect. Our priors may suggest, for example, that it is not democratic rights that matter, but other mechanisms such as protection of property rights and stability of government, that may, in developing countries, be only weakly correlated with our Freedom House measure.

Nonetheless if exploration is strongly tied to institutional environment, this implies resource abundance measures in turn may be affected. This challenges many of the previous cross-country studies that seek to understand how resource abundance impacts economic growth, treating resources as exogenous.

Within-country studies have a chief advantage in this context- the location and amount of resources within a country is likely to be subject to the same institutional environment, whereas geology varies by region. This is true in contexts where control of the resource sector, such as via licensing or a national oil company is centralised. Indonesia provides one such setting, where for petroleum, and up until 2009 for mining, such administration was controlled by the government in Jakarta.

Chapter 3 The local effects of natural resource extraction: In Chapter 3 I turn my attention to the local economic consequences of industrial mining in Indonesia. I present a simple 3-sector general equilibrium model to generate predictions for the local labour market, akin to the Corden and Neary (1982) Dutch disease model of the macro-economy. I test the predicted effects in response to an exogenous localized resource sector shock by looking at mine opening or mine expansion events across three hundred mines. I test the predictions of the model, first by estimating the typical economic footprint from industrial mining; found to be an average of 15km radius. I then examine the response of reported labour market activity from households surveyed in nearby communities. Here I find no evidence for a shift of local labour into the mining sector. I do find however a notable movement of labour from the traded sectors (agriculture and manufacturing) to the non-traded service sector. The effect appears to be greater in magnitude if mines are foreign-owned. This raises interesting questions about local profits and local procurement, particularly in the context of recent Indonesian legislation seeking to see foreign divestment from mining in the country.

This work builds on the study by Aragon and Rud (2013) who analyse the district-level effects of a single gold mine in Peru, this works seeks to extend this approach in several directions. First, it takes over three hundred industrial mines in Indonesia as the starting point. I combine these with GPS-located household survey data collected for 303 community clusters, and with night-time satellite data for over a decade across Indonesia.

A clear shortcoming of this work is the data limitations. While various household surveys exist for Indonesia, the RAND IFLS study was chosen since it also collected GPS coordinates for survey clusters and is available in panel format (unlike for example, the much larger PODES village census which does not track households and has less detailed labour market information. Or the SUSENAS dataset, which is only identified at the district level). As a consequence the chapter is limited to a study of 303 survey clusters in panel form in a country of 65,000 villages.[2].

As a consequence, measurement error and imprecision of estimates is a major concern. Indeed, the insignificance of the employment effect into the resource sector may be a feature of this shortcoming, or may reflect that mines do not generally source direct employment locally. We seek to mitigate this challenge by supplementing the study with night-time lights data which has the principle advantage of covering the whole of Indonesia, and not subject to the same kinds of measurement error as the RAND IFLS data. Indeed this proves fruitful, both to generate independent estimates of the economic impact, but also to help calibrate the treatment variable, and in particular the appropriate distance cutoff(s).

To understand the local welfare consequences we would ideally supplement any study of footprint and labour market effects with price and wage effects. Furthermore the outcomes of ultimate interest may be local incomes, expenditure and human development measures. Unfortunately the dataset faces severe shortcomings with respect to price and wages with extensive

[2]This number is approximate; the Government's Village Potential Statistics (PODES), a regular village census, covered 68,816 village units in total in 2011

missing data and subject to severe measurement error, rendering it impractical to extend the analysis in this direction.

The plausibility of the identification strategy is a further challenge to any study of this nature. Indeed, as Chapter 2 suggests, resource extraction may be more subject to human factors, versus pure geological factors, than previously thought. One advantage my study has over cross-country regressions is the common institutional setting, whereby mining licensing is centrally administered from Jakarta (up until 2009). We may still be concerned however that the timing and scale of mining may be subject to unobserved local, and even community-level determinants. I can mitigate this concern somewhat via the difference-in-difference estimation strategy, however the relatively small number of community level observations limit the insights we can draw from our panel. The heterogeneous effects from differences in mine ownership has proven fruitful and raise intriguing questions of policy relevant that may warrant further investigation.

Chapter 4 Disentangling the effects of resource extraction: local government and investment multipliers: Chapter 4 examines the oil and gas boom in Indonesia from 1999-2009. Here I deploy a variety of identification strategies to attempt to disentangle the regional effects of the boom, measured in terms of district GDP. I estimate effects arising from transfers of revenue to local government. Using an instrumental variable approach I isolate the fiscal channel from resource projects. I find a positive and significant effect of increased local government revenues on district GDP over the boom decade. I then examine the spillovers from resource projects, isolating them from fiscal transfers. For districts neighbouring resource rich districts I find evidence for a modest positive effect arising from project investments, rather than fiscal transfers.

As discussed in the context of Chapter 3, we again must be concerned with the exogeneity assumptions at the heart of the identification strategy. Is the location, timing and magnitude of the oil and gas boom, for example, tied to district and provincial level characteristics? I deploy various approaches to mitigate these concerns, however these come at a price. The instrumental variable estimates are strictly local to the revenues arising from offshore extraction. This limits the insights generated in the context of most oil and gas production taking place onshore, within the districts also receiving windfalls. Furthermore the neighbour-spillover estimations are constrained by looking further afield than the producing districts. Ideally we would be able to measure both the local spillover in the producing district in addition to neighbourhood spillovers. Here I must rely on diff-in-diff and the shortcomings associated with that approach.

A major potential shortcoming of this chapter, like Chapter 3, is our chosen measure of outcome. In addition to district GDP we may be concerned about the wider socio-economic consequences of the oil and gas boom, and therefore would like to examine education and health indicators as well as other measures including inequality and poverty rates in resource rich districts. District GDP remains a useful, though flawed, measure of economic impact. It is subject to concerns around accuracy given the challenging data collection environment in Indonesia. Furthermore with only limited information on the sectoral breakdown of district GDP, I cannot present the preferred specification which would measure the effects on non-oil GDP, rather than total GDP. Nonetheless it can be informative about competing effects and directions of effect, even if magnitudes should be interpreted cautiously.

1.2 Literature

To place the work into a broader context, the remainder of this chapter surveys the literature on the determinants of resource wealth in addition to studies looking at the local and regional effects of natural resource production. The rest of this chapter is organized as follows. Section 1.2.1 discusses the challenge of measuring resource abundance and its determinants, and how the work presented in this thesis may contribute to this area. Section 1.2.2 examines the literature that seeks to understand the local and direct effects of resource extraction, including how Chapter 3 fits into this picture. Finally section 1.2.5 examines the literature on regional spillovers and the resource-related effects from fiscal decentralization, the subject of Chapter 4 which investigates these in the context of Indonesia.

1.2.1 Resource wealth empirical challenges

The starting point for any study of the effects of resource abundance rest on how we measure a country, or region's resource wealth. A variety of approaches have been adopted by the literature, resting on different arguments of the exogeneity of these measures to the outcome of interest.

Empirical evidence for the resource curse beginning with Gelb (1988) and complemented by Auty (2001) and Sachs and Warner (1995) has spurred a huge literature on the effects of natural resources on economic performance. Much of this work has drawn on cross-country analysis to examine whether they exists a positive or negative relationship; the jury is still out (see a recent survey of this by van der Ploeg 2011).

While this approach has been fruitful in yielding insights into, for example, the effects of volatility (van der Ploeg and Poelhekke, 2009) and the relationship to political institutions (Ross 2001, Ross 2013 and Collier and Goderis 2009) increasingly scholars are turning to within-country and spatially identified data to review and deepen our understanding. In particular, recent years has seen a significant growth in single country studies, typically at the state, province or district level, seeking to compare regional performance in response to resource wealth.

A primary reason for this shift of focus is to better understand the micro-economic foundations of potential resource curse effects at the national level. Do resources drive structural imbalance at the regional level or promote regional inequality? Do the same political challenges of rent-seeking and elite capture manifest with subnational governments as at the national level? And are there welfare gains at the local level that might otherwise offset negative effects on the macro-economy?

Beyond asking important questions, this strand of literature has been facilitated by better within-country data which allows for identification and causal inference. Improved availability of district, country and state level statistics, as well as spatially identified household surveys mean this work is now possible using modern econometric and geographic information systems software. Furthermore, given the potentially subtle and context specific form resource curse mechanisms may take, careful single-county studies are being increasingly sought to understand how these dynamics play out.

Finally, there remains strong concerns that cross-country comparisons are subject to severe methodological limitations, namely arising from the potential endogeneity of measures of resource abundance. In the case of the seminal works, such as Sachs and Warner (1995), they rely on measures of resource dependency from resource exports. More recent work by Brunnschweiler and Bulte (2008) has utilised alternative measures such as the measured value of subsoil reserves (World Bank 2011), others have utilised the total value of resource production (Ross, 2012).

An important requirement for each of these studies is that variation in resource abundance, however measured, must be plausibly exogenous from other drivers of economic performance. As noted by Michaels (2011) resource extraction activity may be determined by factors such as strategic pricing considerations (Libecap, 1989) and the level of technology (David and Wright, 1997). These in turn, may affect the outcomes of interest other than via resource extraction and make it difficult to draw causal inferences. van der Ploeg and Poelhekke (2010) point out in their critique of Brunnschweiler and Bulte (2008), it is reasonable to conjecture that resource exports themselves may be affected by the growth rate of the economy. Furthermore their preferred measure, the World Bank natural capital and rent measures are subject to serious methodological doubts and heroic extrapolations that may exacerbate any bias induced by the inherent endogeneity.

More recent cross-country studies have taken advantage of data on giant oilfields Horn (2004) to study, for example, the link from resource wealth to intra-state conflicts Lei and Michaels (2011). Such studies may be less subject to the concerns noted above given the rare and significant nature of giant oilfield discoveries. Other emerging approaches in the literature to tackle endogeneity concerns include those of Tsui (2011) and Cotet and Tsui (2013) who deploy initial resource wealth or discovery rates as instruments for resource abundance.

The work presented in Chapter 2 offers evidence on the role that institutions play in determining the exploration, and hence discovery. This in turn is initial evidence to demonstrate the endogeneity we have long suspected: many measures of resource wealth that are implicitly conditional on having discovered resource wealth, are outcomes of a process subject to institutions that may simultaneously determine outcomes such as economic growth or conflict.

The chapter also seeks to contribute to the literature on institutions and economic growth, most importantly via the identification strategy. Oil's high economic value and knowledge of its likely location are relatively recent phenomena, which provide a setting with plausibly exogenous variation in institutions. The issue of endogenous location of treated units relative to the discontinuity cutoff, often a challenge in spatial RD designs, is in my setting mitigated by the sub-soil character of the oil reservoirs. In contrast, application of nighttime lights as a proxy for economic activity may be more vulnerable to the possibility that the location of people, firms and political borders are simultaneously determined. For example, a country with better institutions could over time position its borders to better capture places with higher growth potential.[3] My findings are consistent with the emerging consensus that institutions are a fundamental determinant of economic growth.[4]

1.2.2 Evaluating the local effects of resources on the economy

Until recently the study of the effects of natural resource extraction within countries and at the local level has been focused predominantly on case study evidence (ICMM, 2007). This approach has included detailed examination of individual projects, a cluster of sites, or comparison of sites. Methodologically this approach has deployed a variety of tools, from descriptive statistics and correlations, to the analysis of qualitative evidence and stakeholder interviews.

In the last decade the increased availability of large-scale socio-economic survey data, geographic identifiers (such as GPS coordinates of projects and survey respondents) and increasing availability and quality of government data at the subnational have all transformed the possibil-

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

[4]See Acemoglu et al. 2005, Acemoglu et al. 2001, Acemoglu and Johnson (2005), Dell 2010, Hall and Jones 1999, Michalopoulos and Papaioannou 2013, Papaioannou and Siourounis 2008 and Rodrik and Wacziarg 2005. In my setting, institutions are considered to affect investment decisions in particular through political risk, e.g. expropriation risk.

ities for researchers. While many of the new applications using spatially identified information have been focused on natural resources, important contributions have applied this to other questions, including examination of long-run effects of forced labour (Dell, 2010), the impacts of natural disasters (Kirchberger, 2010) and the effects of institutional variation on economic activity in developing countries (Michalopoulos and Papaioannou, 2013), to name but a few.

The particular trend in the local impacts literature has been towards econometric techniques deploying identification strategies that exploit spatial or temporal variation within countries and therefore facilitate causal inference. Researchers have increasingly sought to move beyond estimating structural models or correlations in within-country data. Identification strategies now seek to isolate exogenous variation in policy and governance, and in the discovered resource wealth, production activity or revenues. These treatment effects facilitate causal inference around total effects as well as potential transmission channels.

Notable examples include a study of the Yanacocha goldmine in Peru (Aragon and Rud, 2013) which uses mine output and company procurement practices over time to estimate the welfare effects of the mine on its hinterland. A study of offshore oil production and government revenue windfalls in Brazil (Caselli and Michaels, 2013) allows the authors to isolate the oil windfall effect on municipalities benefiting from fiscal sharing rules applied to oil revenues.

More recently work has expanded the scope to examine the impacts on women (Kotsadam and Tolonen, 2013) or the implications for conflict (Berman et al., 2014). With rich census and household survey data, a variety of new research questions can be answered, including the implications for the environment (Aragon and Rud, 2012), and the role of oil in agglomeration economies and long-run structural change (Michaels, 2011).

Reflecting the new approaches to this empirical work, a variety of units of analysis are now employed across the range of studies. At the more granular spatial dimension, spatial grid cells (Berman et al., 2014) as well as individual households or villages surveyed in panel data (Kotsadam and Tolonen, 2013) can be used to spatially link to resource extraction sites. Firm-level surveys offer similar levels of spatial disaggregation, though are currently less-widely used (exceptions are Allcott and Keniston 2013, and de Haas and Poelhekke 2014).

At greater levels of aggregation, a large number of studies use information collected at the district, county or municipality level, typically in panel form (Aragon and Rud 2013 and Caselli and Michaels 2013). The final strand of within-country analysis uses state or province level information to build a panel dataset to investigate variation in regional performance and economic responses (Papyrakis and Raveh 2014 and Beine et al. 2012).

Local impacts of extraction: At the local level, studies have begun to look at the consequences of resource extraction on a variety of outcomes of interest. In Chapter 3, entitled *The Local effects of resource extraction*, I examine the consequences for the local economy, and in particular the local labour market.

The local impacts of resource extraction has been subject to extensive study via case study analysis, often by companies or governments seeking to understand and manage the spillover consequences are mining or drilling. Such case studies are now ubiquitous, having become a required part of doing business in the form of environmental, social and economic impact assessments.

While case studies have provided a rich source of descriptive statistics and qualitative evidence, they are ill-suited to causal inference and external validity. To achieve a deeper understanding of the consequences of resource extraction, economists have deployed new data and new econometric methods to identify and disentangle the various the effects on the local economy.

As a branch of the resource literature, these studies typically examine the proximate effects from mining or drilling on the surrounding region - this may take the form of households, firms, districts, municipalities or states. Much of this work has helped document the economic consequences of extraction, such as job creation and welfare effects. In contrast to subsequent sections of this chapter, here we focus on those studies that examine the direct channels of effect; namely the spending effects and demand for inputs from mines or oil fields, such as arising from their capital and operating expenditures. In subsequent sections we examine the role of government windfalls arising from resource revenues, and from infrastructure spillovers arising from resource-related infrastructure investments.

The direct channels by which extraction can affect the economy are of particular interest. The development of mines or oil and gas fields can constitute significant shocks to a regional economy, generating jobs and drawing in capital from other regions and countries. Both the development of a project and the subsequent operations can generate large local spending effects, both through local employment and through sourcing non-labour inputs. These backward linkages, to use the language of Hirschman (1958), relate to how these projects connect to the local markets. For people living proximate to resource extraction these backward linkage channels may be the main, and sometimes the only channel by which resource wealth benefits livelihoods.

Studies of the local direct effects of natural resource extraction come from an established strand of the economic literature seeking to understand the labor market effects of demand shocks, often studied in the form of large scale investments. The recent trend in this literature has been to identify specific shocks and hence identify spillover effects using causal inference. Notable contributions include the examination of large-scale construction projects (e.g. the US interstate highway, Michaels 2008, the Alaskan oil pipeline, Carrington (1996)). These studies have documented the observed positive effects on employment, wages and house prices as would be expected from a major localized increase in investment and demand for inputs. More recent developments have provided theoretical foundations for this work (such as work on local multiplier effects in Moretti 2011) and a broader empirical literature has pioneered innovative identification strategies and use of spatial data in other settings (for the effect of dams see Duflo and Pande (2007); for rural electrification: Dinkelman (2011); and for rail privatization: Lowe (2014)).

In contrast, there is some evidence that local projects can have harmful effects too; the study of enterprise zones is one example where a variety of authors have documented neutral and even negative effects on local firms (Bartik, 2003). Other studies of local economic damage include a study of the effects of military base closures in the US Hooker and Knetter (2001).

The welfare implications of resource extraction at the within-country level are not clear a priori since extraction is generally considered an enclave activity (Hirschman, 1958) with limited linkages to the regional economy. This poses an important question for the literature, of whether there is evidence for the resource curse at the subnational level, as opposed to more intuitive welfare and labour market spillovers generated by a local demand shock.

The important study by Aragon and Rud (2013), focusing on a single mine, exploits an exogenous expansion of the mine and change in procurement policies by the mine owner Newmont Mining Corporation to examine the local effects of a change in spending patterns. They find a strong positive relationship between the change in policy and real income levels of the local population that lives within 100km of the mine; they identify the mine‘s demand for local inputs as the main transmission mechanism for these welfare effects. They argue that for their short-run analysis, the backward linkages channel, such as the demand for local inputs, can be a more effective transmission mechanism for local benefits than government spending of revenues. However this is predicated on a mining region already having developed markets for goods and services, a prerequisite for the hinterland being able to respond to increased demand for inputs.

The Yanacocha gold mine is special, in that it represents the second largest goldmine in the world and in the fact that its local procurement policy was instated by one of its shareholders, the International Finance Corporation of the World Bank Group. An important question, therefore, is how externally valid these findings may prove to be. Other studies have followed, extending the Aragon and Rud analysis to different countries and multiple mining projects of differing sizes. Chapter 3, for example, examines the labour market effects of various industrial mines in Indonesia. I find a small average footprint, of around 15km, within which most labour market effects are felt. Inside the typical mine footprint, I document a significant shift of employment from the traded manufacturing sector to the non-traded services sector, with no significant shift of local labour into mining. Mine ownership appears to matter, increasing point estimates for the labour market responses. This may suggest greater local sourcing by foreign than domestic firms. In analogy to this, de Haas and Poelhekke (2014) find in a broad sample of emerging economies that traded sector firms that have no direct link to the extraction sector report more constraints to doing business if they are located within 20km of active mines. These constraints, measured by the Business Environment and Enterprise Performance Survey, include transportation bottlenecks, electricity provision, access to educated workers, crime, and financial constraints. In turn, these constraints tend to stunt firm growth in terms of employment and sales. Firms outside of mining areas report fewer constraints. This constitutes a local variant of the Corden and Neary (1982) resource movement and spending effects.

Backward linkages may constitute the main, and in some cases only means by which the economic benefits of resource extraction are felt by those living proximate to resources. Lippert (2014) studies the extent to which Zambians have benefitted in the copper belt region. The author looks at the effect of the Zambia copper boom using exogenous variation in copper production at the mine level between 1996 and 2010. His study draws upon repeated cross-sectional data drawn from the Living Standard Measurement Survey (LSMS) to create a constituency-level panel. The central result of this paper suggests that an increase in local copper output improves measures of living standards in the respective constituencies via the mines' backward linkages. In particular Lippert estimated an increase in real household expenditure of 2 percent associated with 10 percent increase in copper output at the constituency level. Lippert's study makes the claim of estimating the pure project effects i.e. the backward linkages isolate from any government spending mechanism by reference to the very small revenues accruing to the government during the copper boom.

A similar study by Loayza et al. (2013) uses variation in mining across Peruvian districts to investigate the impact of mining activity and government transfers on local socioeconomic outcomes. The analysis uses a district level dataset that merges administrative data on local mining production, transfers from central to local government, and census and survey-based measures of households average consumption, poverty, and inequality. The authors are able to trace the positive impact of mining on local communities with the strongest effects in producing districts- who benefits from both the mine and government transfers. However, their findings do suggest that mining is exacerbating inequality across districts in Peru- even beyond producing regions. This may be driven by a policy of returning a disproportionate share of resource revenues to the local district, at the expense on non-mining districts.

The documented positive, but modest gains from resource extraction projects challenge some narratives of enclave development. However, there remain concerns that the gains may be distributed unevenly or asymmetrically. These heterogeneous effects are also being examined in recent studies. Kotsadam and Tolonen (2013), for example, use data from the Demographic Health Survey for various countries in Africa to examine the effect of mine opening. They use the RMG Raw Materials Database for information on mine location and opening year, and combine it with DHS clusters using GPS coordinates and straight-line distances. Their study is concerned with the gender impacts of mining, finding evidence for increase female employment from mine opening. Furthermore they find evidence for a shift of women into the service sector, while the effect decays over distance. They find an asymmetric effect for mine closure and suspension with women not fully returning to the agricultural sector, whereas overall employment levels remain low.

The finite nature of resource extraction means that all projects will eventually end and closure prompts concerns of job losses and reversal of economic gains that might have accrued during operations. A study of the coal boom and bust in north-eastern United States (Black et al., 2005) measures the employment spillovers of both mine opening and closure. The authors find a modest number of non-coal jobs created by the boom. During the bust, the spillovers are bigger and more persistent than the boom. This asymmetry of effect is being taken up by other more recent papers (for example, by Toews et al. 2014 for the coal mines closures in the United Kingdom, for mines in Indonesia, and by Kotsadam and Tolonen 2013 in Africa for women and mining, and in Chapter 2 of this work). In contrast, for the Southern US, Michaels (2011) finds evidence that previously oil-producing regions can be better off in the long run, even after production has ceased.

Studies of the relationship between resources and economic development at the macroeconomic and cross-country level have highlighted the important mediating effect of political institutions. This importance seems to be replicated at the subnational level. Libman (2013), for example, uses variation in institutional quality across Russian regions to examine the relationship between resources and governance at the within-country level. His empirical work supports the cross-country hypothesis, that the economic effects of the resource curse are conditioned of the level of institutional quality (see for example Collier and Goderis 2009). Libman finds that resources have a negative association with growth if the quality of regional governance is low. On the other hand, increasing levels of democracy also have negative consequences for regions with substantial resources. However, as the author notes, these estimates cannot be interpreted causally and constitute a very short (seven year) timespan.

Finally, local effects studies are also concerned with the negative potential of resource abundance. The clearest evidence for a resource curse factor being observed at both the cross-country and the local level is that of conflict. Various authors, including Dal Bó and Dal Bó (2011) show theoretically that positive shocks to capital intensive industries like mining may increase the risk of conflict - a prediction empirically confirmed by Berman et al. (2014) for mining in Africa and Dube and Vargas (2006) in their case study of oil windfalls in Columbia. Berman et al. (2014) utilized mine location data for Africa combined with the ACLED conflict dataset (Raleigh et al., 2010). This allows them to examine the relationship between world price shocks to minerals and subsequent conflict incidents. They find evidence to suggest increases in mineral prices trigger increased violence close to mine sites- measured both in terms of riots and unrest, and in terms of battles. Furthermore their evidence suggests that access to mines supports the ability of armed groups to perpetuate violence and this can drive secessionist violence. Oil can be a driver of internal conflict too. Using the Horn (2004) dataset on giant oilfields, Lei and Michaels (2011) find that on average discoveries increase the incidence of internal armed conflicts by about 5-8 percentage points within 4-8 years of discovery, compared to a baseline probability of about 10 percentage points.

Related studies use within-country variation to document the adverse effect of oil windfalls on corruption in Brazil (Caselli and Michaels, 2013), on crime (James and Smith, 2014) or the adverse environmental impact of Gold mining in Ghana (Aragon and Rud, 2012).

The various studies examining the direct effects of resource extraction share many findings in common. Through backward linkages to the local economy, projects can have significant positive welfare effects. Indeed there is evidence to suggest this may be the most effective channel, versus say local government spending. This stands in contradiction to the claims that such projects act as enclaves with very limited ties to the local economy. The magnitude however can be small, and is conditioned on the policies of resource projects as well as the availability of local markets for goods and services. On the other hand, there is evidence that resource projects can fuel internal conflict- both from petroleum and solid mineral production. Extraction can increase and exacerbate inequality, with asymmetric effects on men and women, and there is some initial evidence that now suggests that mine closure or a resource bust can be more painful than the welfare gains during mine opening or the boom.

1.2.3 Regional spillovers and Fiscal decentralization

In Chapter 4 I present a study of the effects of the oil and gas boom in Indonesia on the regional economy, measured at the district level. In particular in my chapter I seek to disentangle the fiscal channel effect, from the impact of project level activity. I present estimates for both channels deploying a variety of identification strategies to do so. This work follows a growing trend in the literature to deploy identification strategies that pursue causal inference on the various channels of effect arising from a resource boom.

The recent shift from cross-country data to subnational data sets has also led to improved identification of Dutch disease, one of the main macroeconomic mechanisms behind a resource curse. Moreover, subnational and micro units of observation allow identification of spillover effects across (local) industries and also across regions. For example, the recent shale revolution has increased interest from policy makers (who decide on drilling licenses) on the costs and benefits of local resource booms on the environment, on employment and on growth. Such local booms may result in migration which mitigates Dutch disease effect. Furthermore, extraction typically requires extensive investment in infrastructure with potentially important side effects on trade and investment.

Allcott and Keniston (2013) examine county-level data for the US to investigate the local and spillover effects of boom-bust cycles in natural resource production.

Several US counties experienced two oil and gas booms after the oil crisis in the 1970s and recently during the shale revolution and one bust during the 1980s and were treated with sizeable changes in natural resource employment. Other counties with less or no initial resource endowments did not experience these cycles directly. However, the high degree of disaggregation allows identification of spillover effects across industries and counties. Their findings suggests pro-cyclical performance of the aggregate county economy: a doubling of resource sector employment predicts a 2.9 percent rise in total employment. In addition, manufacturing employment increases, even though manufacturing wages also rise. The latter reflects imperfect elasticity of labor supply despite significant migration to booming counties which mitigates the rise in wages. Since many manufactures link to the resource sector as up or downstream industries, or produce locally traded goods, these findings are in line with the classic Dutch disease models (i.e. Corden and Neary 1982). However, they find only weak evidence that non-linked traded manufacturing firms suffer, which goes against the predictions of Dutch disease. A related paper on the local effect of the coal boom in the 1970s and bust in the 1980s in coal producing counties of four US states found similar albeit small employment and wage effects across sectors (Black et al., 2005). In addition, and consistent with the relatively low-skill intensity of coal mining at the time, they documented positive effects on poverty during the boom period, suggesting that existing residents benefited from the boom.

Interestingly, one of the main ideas of how Dutch disease may reduce aggregate growth is also not supported by the data in Allcott and Keniston (2013). In Van Wijnbergen (1984), learning by doing, which may mainly take place within the traded sector would also suffer from a resource boom, thereby reducing productivity growth. The paper finds neither contemporaneous nor long run evidence that revenue based TFP of non-linked traded firms relatively declines in booming counties. In addition, none of the positive contemporaneous TFP effect on the other industries hold in the long run. This contrasts the idea based on Chinitz (1961) that mining towns may specialize in heavy industries with large scale economies which in turn crowds out entrepreneurship by reducing access to inputs, capital and investment in skills. Glaeser et al. (2012) show that this reduced employment growth in such cities.

These results may depend on how resource, tradable and non-tradable sectors are defined.

As long as a firm sells some of its output locally, it will also benefit somewhat from the local increase in demand. A narrower definition of non-linked sectors may result in stronger negative spillovers from the resource boom: these firms pay higher wages but cannot benefit from local demand. Alternatively, there may be limited substitutability between the traded and resource sectors, at least in the short run. Another potential explanation is that any negative effects are cancelled by other unobserved positive effects, such as from local infrastructure provision or local tax cuts.

In addition to looking across sectors, Allcott and Keniston (2013) also find evidence for considerable regional spillovers: less resource abundant counties experience positive spillovers from resource abundant counties. Although they may lose population to the higher wage offers in resource counties, their industries also supply more goods and services to meet the increase in regional demand. Michaels (2011) examines US counties in the South and compares the long run economic outcome of oil rich versus oil scarce counties. He finds mostly positive effects however: oil rich counties have higher population density, higher per capita income and higher manufacturing employment density despite inflated factor prices. While he cannot rule out that this may be due to linked industries, the paper also finds that investment in infrastructure has been higher in oil rich counties, which may have led to positive agglomeration externalities and can explain the positive spillover effect on the agricultural sector. This suggests that a local boom can be sustained if local revenues also translate in local investment in public goods, providing another channel of positive spillovers. Such a beneficial public policy may be lacking in less developed countries.

Notwithstanding the clean identification of the results in Allcott and Keniston (2013), Michaels (2011) and Black et al. (2005), the question arises how specific these results are for the US case. One can imagine that in other countries, with fewer up or downstream industries (for example because there is no local refining capacity), a larger share of the manufacturing sector may suffer from Dutch disease effects. Furthermore, the high degree of internal migration limits the rise in wages, which, as the paper notes, explains the relatively small magnitudes of the effects. In contrast, Caselli and Michaels (2013)) found no evidence for interregional migration flows as a result of oil booms. Also, the US is a net importer of hydrocarbons, which limits upward pressure on the exchange rate and reduces electricity prices, further limiting the negative spillovers on the traded sector. Another consideration is that in the US case a large share of resource sector revenues may be spent locally in the form of wages from resource sector workers, but also by land owners whom hold ultimate title to the land and subterranean natural resources and thus receive royalty payments. In other countries, such royalties typically accrue to the national government, and wage earners in the resource sector are not necessarily hired locally. Future research at a similarly disaggregated level from less developed countries may shed more light on these considerations.

A related paper uses provincial level data for Canada (Beine et al., 2012) to investigate how large the mitigating effect of labor mobility are on Dutch disease dynamics. Without internal and international migration, competition for workers would lead to a larger relative change in the size of the non-traded sector (as proxied by the services sector) versus the manufacturing sector in booming regions. Identification comes from the boom in Albertas tar sands oil production and the discovery of oil fields off the coast of Newfoundland. They find that a local resource boom increases the share of the non-traded sectors in the provincial economy, while mainly temporary immigration reduces it. On the other hand, the resulting interprovincial out-migration in non-booming regions leads to negative spillovers in the form of an increase in the share of the non-traded sector.

Further evidence for Canada provided by Papyrakis and Raveh (2014). They find that a provincial boom in natural resources leads to higher local prices and a decline in the share of employment in the tradable sector, although the share of capital in the traded sector increases, thereby also increasing the capital intensity of the traded sector. By estimating a spatial lag model based on distance between the largest cities of each province, they find that a boom in neighboring regions conversely increases the share or labor in the traded sector, while the share of capital in the traded sector decreases, such that capital intensity decreases. Since they make no distinction between linked up and downstream industries and unrelated traded manufacturers, it may be the case that the increase in capital intensity is due to capital intensive linked sectors drawn into the booming region. Alternatively, the positive effect on capital may be due to lower capital taxation which may be financed by resource revenues in the booming province (Raveh, 2013). In addition they find that non-resource exports suffer as a result of a local boom in natural resource production, consistent with Dutch disease. Interestingly, only international exports suffer, while national trade remains unaffected, reflecting the appreciation of the real exchange rate, and the possibility that international trade may have been substituted by national trade to meet increased local demand.

1.2.4 Fiscal decentralization

A relatively recent development is that natural resource revenues have directly benefited local government budgets and the producing administrative units. Prompted by the wave of fiscal decentralization in the 1990s and early 2000s countries have increasingly adopted fiscal sharing rules and budgetary methods for allocating resource revenues around the country. These typically result in disproportionate shares of revenues being returned to local governments proximate to resource extraction activity.

Resource extracting firms are typically taxed by the national government or they are state owned, resulting in resource rents accruing purely to the national government. However, several countries have seen a move to the decentralization of executive and fiscal powers to local administrative units to foster regional autonomy, improve service delivery, and improve accountability of government. To some extent, this was driven by a process of democratization and the wish of (ethnic) minorities to have more say in political decisions. In addition, resource extraction itself created demand for decentralization. Observing steeply rising natural resource revenues during the 2000s, local governments demanded a larger say in the way in which revenues were spent and a larger stake in the revenues themselves. The combination of these trends resulted in a higher impact of resource extraction on their region, through the revenue channel. However, the size of this impact depends on the sharing rules between regions and the way in which local governments spend the revenue. On the one hand, there is relatively little research that compares sharing rules and fiscal arrangements making it unclear what the optimal policy should be. On the other hand, the available evidence suggests that, for a given degree of revenue sharing, the impact of resource rent transfers to local governments has resulted in mixed blessings.

In theory, rent transfers to local governments have the potential to generate more public support for mining by compensating environmental costs. Moreover, the classic case for fiscal federalism (Stigler 1957; Tiebout 1956; Oates 1972) is that local (elected) governments are better informed of local needs and ways to spend the rents effectively, although public goods provision such as infrastructure that transcends local boundaries should be provided by higher levels of government. Countering this argument is that local governments may not have the capacity to administer the rents, and fragmentation of rents across many executive powers may result in ineffectively small amounts being made available for investment. At the same time, public investment may not be effective at a local level due to absorptive capacity constraints. Finally, the different levels of governments may have different priorities: local development needs (such as water provision) are not necessarily in line with national development goals (such as higher education). These differences may also result in political conflict over what is considered to be the more equitable division of resource wealth.

Recently, a small literature has emerged on the effectiveness of rent sharing with local governments. For Brazil, Caselli and Michaels (2013) show that, unfortunately, corruption and embezzlement drive a wedge between the amount of fiscal transfers or royalty payments received and local outcomes. They find that municipal revenues increase with royalties from oil production, which amount to roughly 3% of gross oil output. To identify the effect of this on socio-economic outcomes, they compare coastal municipalities with no oil to those with only offshore oil. The reason is that offshore oil, as opposed to onshore oil, is shown to have no systematic effect on non-oil GDP, suggesting that the effect of oil works through revenues only and not through spillovers from extraction activity. The main result is that municipal spending increases much less than what can be expected from the increase in royalty payments. Spending does increase, but, on average, survey reported measures of real outcomes do not improve, nor do oil municipalities attract migrants, suggesting that welfare does not increase. They argue that the lack of a positive effect cannot be accounted for by offsetting reductions in taxation or federal spending in municipalities, neither of which takes place. They suggest that mayors use money to create fictitious public employment for certain groups to improve reelection chances. Also, mayors of municipalities with oil income are found to be more commonly mentioned in the news in relation to corruption. An open question however, is whether in this respect natural resource windfalls are different from other windfalls such as aid or federal transfers. The former may be worse because there tends to be less transparency over the amount of revenue involved.

In addition to corruption, Arellano-Yanguas (2011) show, based on field research in mining regions in Peru, that redistribution of natural resource rents to regional and local levels of government can create new problems. On the one hand, it increases public support for mining because revenues no longer disappear to the national government. However, it may also fuel new forms of conflict. Where previously typical conflicts involved the mining company and the local community over issue such as land ownership, access to water and pollution, new forms of conflict involve access to natural resource revenues and their use. Partly, the issue is that the streams of revenues increased quickly during 2005-08 without increased administrative capacities in government, such that corruption and absorption capacity constraints led to inefficient spending. Disagreement between levels of government on how the revenue should be spent could, for example, be paid off by raising public salaries. Also, the structure of governance is poorly equipped to execute development strategies due to fractured mandates, overlapping jurisdictions, and high staff turnover. Other forms of conflict involve the rules of allocation which result in political conflict over territorial boundaries or fights over the criteria themselves, fueled by increased inequality between mining and non-mining regions. The basic message is that decentralization of revenue spending can only work if government capacity is improved simultaneously.

This latter view seems consistent with the experience in Colombia, which divides oil and coal royalties between the national government, departments, and local municipalities. Unlike federal transfers, royalty payments are not earmarked specifically although education and health has priority over other public investment. Nevertheless, in a panel of Colombian departments, Perry and Olivera (2009) find that high oil production and high royalty payments correlate negatively with GDP per capita growth. Royalties only appear to increase income in departments which already have a strong capacity to generate tax income. Possibly, departments with governments able enough to raise revenue through taxes are also better equipped to spend royalties relatively wisely. In contrast, in Chapter 4, I present evidence for Indonesia that fiscal transfers related to oil production boost local GDP. In particular they appear to have positive and significant effects on district GDP.

Transfers increased markedly in Indonesia after 1999 when the fiscal system was decentralized to give local governments more autonomy over spending, and resulted in larger disparities between producing and non-producing districts. As in Brazil, this has led to the endogenous creation of more local (and even provincial) administrative units to carve out larger local shares of the revenue. Direct local elections foster accountability over spending, but have also led to increased patronage and collusion with businesses. Moreover, smaller administrative units may not have the capacity to spend revenue effectively. In Kazakhstan, Toews (2013) finds that oil windfalls can lead to increased dissatisfaction, however this may be driven by inflated expectations rather than simply government under-performance.

Finally, in Chile rent transfers are aimed at decreasing inequality between provinces. However, Aroca and Atienza (2011) make a case that the mining region of Antofagasta loses more potential mining related wage income through commuting than the region receives in the form of fiscal transfers through the regional development fund. The reason is that reductions in travel costs have allowed city-mine commuting to replace the creation of traditional mining towns close to mines, especially for more remote areas. These allow the wage income earned by labor to spread further from the mine. Naturally, this reduced the scope for local spillovers. For the case of Chile, which has 44 percent of workers concentrated in the capital city, leaving rest of the country fairly empty, he finds that regions that send workers are the main beneficiaries of mining in the region of Antofagasta, where 10 percent of the workforce commutes from out of the region.

The evidence suggests that positive local outcomes as a result of revenue sharing with local government are possible but in no sense are they guaranteed. Further research is warranted on how to distribute rents in the best way and how to equip local governments with the capacity to spend them effectively. For this reason, none of the papers above allow strong welfare statements, because no benchmark exists. Optimal fiscal policy design with resource rents has been analyzed at the national level, but has been under researched at the local level.

1.2.5 Reflections

A growing body of theoretical and empirical literature has shifted-attention from examination of the resource curse hypothesis across countries, to exploring the potential dynamics at the local and regional levels. In particular, new datasets and empirical methodologies have allowed researchers to conduct causal inference of the effect of resource projects, revenue windfalls, and infrastructure associated with resource abundance. These studies find strong evidence for impacts positive and negative extending beyond the project enclave.

Through backward linkages to the local economy, projects can have significant positive welfare effects. Indeed there is evidence to suggest this may be the most effective channel, versus local government spending. The magnitude however can be small, and is conditioned on the policies of resource projects as well as the availability of local markets for goods and services. On the other hand, there is evidence that resource projects can fuel internal conflict - both for petroleum and solid minerals. Extraction can increase and exacerbate inequality, with asymmetric effects on men and women, and evidence suggesting that mine closure or a bust, can be more painful than the welfare gains during mine opening or the boom.

While mines and oil fields appear to generate positive labour market and welfare effects on the surrounding economy in the short run, the magnitude is possibly only large in developed countries such as the US, which have more potential for developing local up- and downstream industries. Also, internal migration can limit Dutch disease effects. There is some evidence that developed countries are better able to prolong boom periods by investing in local infrastructure and other public goods.

Chapter 2

Institutions and the Location of Oil Exploration

James Cust[1]

[1]This chapter constitutes joint work with Torfinn Harding, NHH Bergen.

Abstract

We provide evidence that institutions strongly influence where investors drill for oil and gas. At national borders, investors drill on the side with better institutional quality two times out of three. To identify the effect of institutions, we utilise a global data set 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 utilise the historical sequence of drilling occurring after the formation of borders and institutions. Our 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. They imply that the observed distribution of natural capital across countries is endogenous with respect to institutions.

> The good Lord didn't see fit to put oil and gas only where there are democratically elected regimes friendly to the United States. Occasionally we have to operate in places where, all things considered, one would not normally choose to go. But, we go where the business is.
>
> Dick Cheney, 1998

2.1 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 chapter uses the setting of oil and gas exploration to estimate the effect of institutional quality on investment.[2] 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.[3] 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.[4]

[2]We focus in this chapter 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'.

[3]Investment in exploration drilling is an equilibrium outcome, determined by 'demand' by exploration companies and 'supply' by potential host countries. We study in this chapter how these equilibrium outcomes vary depending on institutional quality.

[4] 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

This chapter argues that variation in institutional quality between countries can contribute to understanding the uneven distribution of investments and hence the distribution of known subsoil 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 wells and national borders allows us to examine investments in areas proximate to national borders. We posit that national borders were determined independently of subsoil oil. A discontinuity in the number of oil exploration wells at national borders can then be interpreted as the causal effect of the border on oil exploration.

The location of recent borders may have been affected by knowledge about the likelihood of discovering oil, posing a threat to the assumption that borders are random with respect to the location of oil. We therefore only consider borders which have not been changed since 1965 and only drilling that took place from 1966 and onwards.

To be able to use randomly located borders to estimate the effect of institutions, we need to introduce measures of institutional quality. We follow Michalopoulos and Papaioannou (2014) and place the country with the better score for institutional quality on the 'right hand side' of the border, and the other country on the 'left hand side' of the border. The institutional variation is therefore bilateral and strictly ordinal; we merely use the institutional measure to rank the two countries. We define a border-dummy, taking one on the right hand side and zero on the left hand side of each border and we use this as the explanatory variable of interest in the regressions. 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. To get around this, institutional quality is measured in 1965 or before, prior to the exploration drilling we consider.[5]

---

the investment figure.

[5]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.

In addition to using the institutional quality measures to simply bilaterally rank countries, we also estimate a model where we take into account the actual jump in the measured institutional quality at the border. We do this by estimating with two-stage-least-squares, using measures of institutional quality as the treatment variable and the border-dummy as the instrument. This approach implicitly assumes that the measures of institutional quality are interpretable in a cardinal way, which is a widely made assumption in the literature (e.g., Acemoglu and Johnson 2005).

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). These broad measures are also available with good country and time-coverage, which is essential for our identification strategy. Finally, the approach has the benefit of yielding general and broad findings. Unbundling of the effects of institutions into more granular characteristics is left for future research.

Our baseline estimates suggest 95% more exploration well drilling on the side of the border with relatively better institutions. For every well drilled on the 'bad' side of the border, two are drilled on the 'good' side. Scaled with measures of institutional quality, the estimates imply that a one standard deviation increase in, for example, the augmented Freedom House Political Rights Index score increases the number of wells by around 100%. These are close to long-run estimates since they are based on the accumulated number of wells drilled in the period 1966-2010.

We interpret this as the effect of institutions broadly defined. Held as a fundamental cause of economic growth, institutions operate through a variety of transmission channels, e.g. laws and policies that again affect investments in physical and human capital. We do not attempt to unbundle the effect into these various channels and our results should therefore be interpreted as the total effect of institutional quality. We show that the competing explanations of 'geography' and 'culture', the two other fundamental causes of economic growth proposed by the literature, are unlikely to drive our results. Note that both demand for drilling by the exploration companies and supply of drilling by host countries may be affected by institutional quality. We only observe the equilibrium outcomes on each side of the border and do not attempt to separate the effect into demand and supply effects.

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 covered by our data.

A key question regarding our identification strategy is whether national borders are randomly assigned with respect to the underlying geology. We first demonstrate that the actual likelihood of discovery, conditional on drilling taking place, does not vary at the borders. Second, we show that path dependency, or previous drilling, does not drive our results. Third, we deal with offshore borders. These were formally established only relatively recently, when more information about the underlying geology was available, and may be less constrained by natural structures or historic human activities. Thus they might have been systematically influenced by the quest to secure oil by countries with 'better' institutions. We show, however, that the results are robust to excluding offshore borders and omitting onshore borders that have changed since 1965. Fourth, interstate conflicts may be correlated with oil drilling and institutional quality and thereby impose a bias in our estimates. We show that our results are robust to dropping borders that have experienced severe military interstate disputes. Finally we deal with unobservable country and neighbour characteristics by inclusion of country and neighbour fixed effects.

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

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 chapter.

This chapter contributes to the literature on institutions and economic growth most importantly by the econometric identification strategy. Oil's high economic value and knowledge of its likely location are relatively recent phenomena, which provide a setting with plausibly exogenous variation in institutions. The issue of endogenous location of treated units relative to the discontinuity cutoff, often a challenge in spatial RD designs, is in our setting mitigated by the sub-soil character of the oil reservoirs. In contrast, application of nighttime lights as a proxy for economic activity is more vulnerable to the possibility that the location of people, firms and political borders are simultaneously determined. For example, a country with better institutions could over time position its borders to better capture places with higher growth potential.[6] Our finding is consistent with the emerging consensus that institutions are a fundamental determinant of economic growth.[7]

Regarding the literature on oil exploration, this chapter complements the innovative contribution of Bohn and Deacon (2000), who model theoretically how expropriation risk may affect oil exploration and production decisions. The effect of higher expropriation risk is ambiguous,

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

[7] See Acemoglu et al. 2005, Acemoglu et al. 2001, Acemoglu and Johnson 2005, Dell 2010, Hall and Jones 1999, Michalopoulos and Papaioannou 2013, Papaioannou and Siourounis 2008 and Rodrik and Wacziarg 2005. In our setting, institutions are considered to affect investment decisions in particular through political risk, e.g. expropriation risk.

but is negative under plausible assumptions. Their cross-country empirical analysis confirms a negative relationship. 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). Our chapter improves on the empirical analysis of Bohn and Deacon (2000) by using micro data on oil exploration and plausibly exogenous variation in institutional quality. We also utilize more general measures of institutions, as motivated above.[8] Our finding is in agreement with their results.

The literature on the resource curse has been concerned with the effects of the abundance of natural resources on economic and political performance, as summarized by van der Ploeg (2011). The focus with regards to institutions has been on either their role in mitigating adverse effects of natural resources on economic growth, or on the potential deterioration of political institutions caused by natural resource revenues. Our evidence of an effect running from institutions to oil exploration implies that cross-country variation in oil reserves and production must be treated as endogenous in empirical models of the effects of oil.[9] Our finding also suggests that at least some types of natural capital are similar to human and physical capital in the sense that their accumulation (discovery) depends on institutions. This has immediate consequences for the understanding of the determinants of the 'Wealth of Nations', as measured by the World Bank (2011).

The chapter proceeds as follows. The next section describes the data and outlines the identification strategy. Section 2.3 presents the main results. Section 2.4 investigates in detail potential threats to identification and discusses the issue of external validity. Section 2.5 examines

[8]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).

[9]We discuss the implications of endogenous institutions with respect to oil in section 4.6. 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.

heterogeneity across company types. Section 2.6 offers concluding remarks.

2.2 Data and empirical strategy

2.2.1 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, the depth of each well and whether the well turned out to be 'dry' or result in a discovery.[10] Figure 2.1 shows the global distribution of wells available.

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). 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 drilled will not be represented in our main specifications, although we include these in robustness checks.[11]

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

[10] 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.

[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.

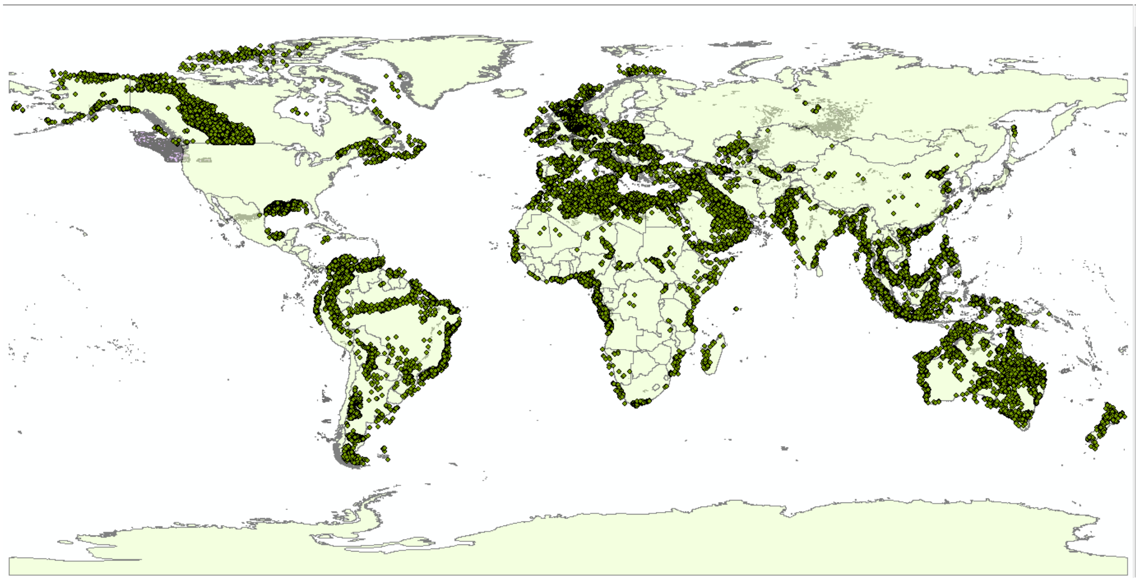

Figure 2.1: National borders and exploration wells- excluding data for onshore US, and much of the former USSR.

Political Rights Index.[12] 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.[13] *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.[14] Country-level data on country size and landlocked status are from CEPII.[15]

We focus on developing countries as host countries (the country in which drilling takes

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

[13]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.

[14]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).

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

place) and include all relevant borders, i.e. also borders with developed countries.[16] We use all data available, with two exceptions. First, all formally 'disputed areas' in the offshore data are excluded given the unclear and contested ownership of such areas.[17] Second, countries in the region of the former Soviet Union are excluded for simplicity, due to potentially complicated variation in institutional quality in the USSR and its satellites.[18] Separate analysis for the sample of developed countries is discussed in section 2.4.4. Table 2.8 presents descriptive statistics and table 2.12 presents the included host and neighbouring countries.

2.2.2 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 or 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 the oil exploration relative to national borders and by utilising the 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

[16]We use in this chapter the 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.

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

[18]This means that we drop all countries belonging to the Europe and Central Asia (ECA) region (8178 wells in total): Albania, Azerbaijan, Belarus, Bulgaria, Georgia, Kazakhstan, Kyrgyz Republic, Romania, Russian Federation, Tajikistan, Turkmenistan and Uzbekistan. Some countries have limited coverage in the PathFinder dataset on wells, notably onshore United States and parts of Eastern Europe, however since we also exclude developed countries this does not present problems for our estimation strategy. Our results hold if we include developed countries and ECA, see footnote 45.

[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.

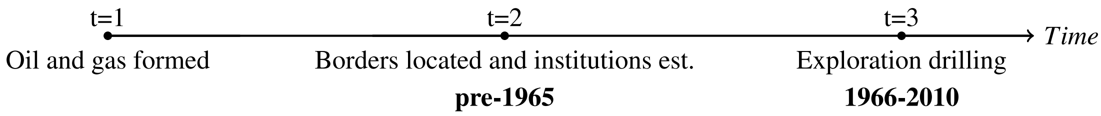

Figure 2.2: Sequence of events in our sample

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.

Figure 2.2 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 use data for only new exploration activity.

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 governed 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:

$$N_{bij} = \alpha_0 + \tau D_{ij} + f_r(X_{bij}) + D_{ij} f_r(X_{bij}) + Z_i' \delta_0 + u_{0,bij} \tag{2.1}$$

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

[20]See section 4.6 for more information on the data from Weidmann et al. (2010). An alternative would be to use the list of countries that have not changed shape since 1945, also provided by Weidmann et al. (2010). This would further reduce the number of countries we could include. Reducing the sample to borders not changed since 1945, expanding the sample by measuring institutions for more recent dates than 1965, or include wells drilled before 1966 do not change the conclusion of this study.

[21]The distance from the border is measured to the mid point of the distance-bin. For example, the tenth distance-

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 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.

The idea is to estimate whether there is a significant jump in the mean at the discontinuity, the national border, captured as a statistically significant τ-coefficient. Its magnitude can be interpreted as the effect of crossing the national border on the number of exploration wells drilled. Note that institutional quality plays a role for the estimate of tau, but only in an ordinal sense by determining the ranking of the two countries. Simultaneity bias is therefore only an issue at this stage if the ranking of the two countries is affected. We discuss threats to identification in section 2.4.

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.[22]

To estimate how responsive oil exploration is to a given change in institutional quality, we apply two-stages least squares estimation. To formalize, the effect of institutions, I, on oil exploration is estimated by:

$$N_{bij} = \alpha_2 + \gamma I_i + f(X_{bij}) + D_{ij} f(X_{bij}) + Z_i' \delta_2 + u_{2,bij}, \tag{2.2}$$

$$I_i = \alpha_1 + \phi D_{ij} + f(X_{bij}) + D_{ij} f(X_{bij}) + Z_i' \delta_1 + u_{1,bij}, \tag{2.3}$$

The estimate of ϕ in equation 2.3 (the 'first stage') indicates whether the border has an

bin is measured at 950 meters from the border and will include all wells drilled from 900 to 1000 meter from the border.

[22]We discuss local area size in appendix 2.8.1

effect on institutions. The dummy D is still defined as one for the side with relatively better institutional quality and ϕ will therefore be positive by construction. γ provides us with an estimate of the jump in the number of wells due to a unit change in institutional quality. By the estimation of equation 2.2 (the 'second stage'), the measures of institutional quality are assumed to have a valid cardinal interpretation.[23]

Based on a specification analysis, we specify the f-function as a third-order polynomial and include wells within 10 km from the border. Appendix 2.8.1 discusses sensitivity to these choices. Although the magnitude of the estimated effect does vary somewhat, the finding of this chapter does not change qualitatively across different specifications.

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.

2.3 Baseline results

This section starts with 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. The section ends with a discussion of the role of geography and culture, the competing fundamental causes of economic growth, in our estimates.

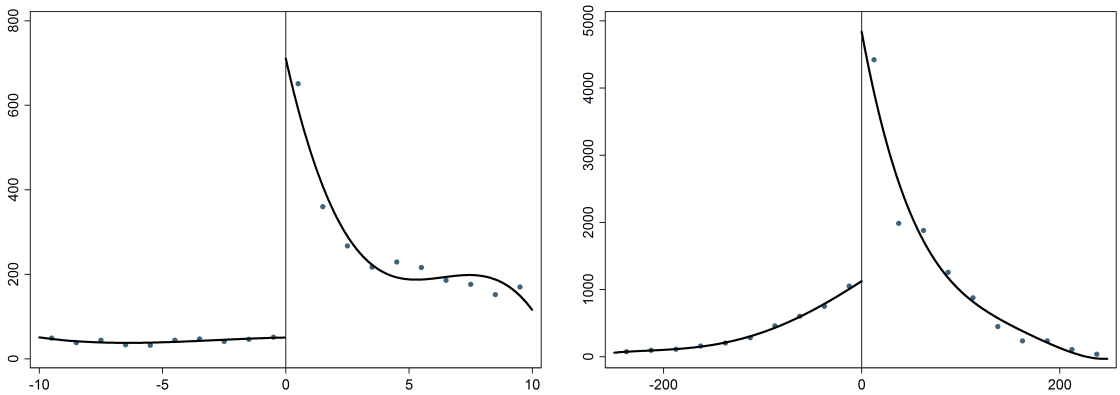

Figure 2.3: Number of wells in developing countries

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 2.1 regarding the sample and the measure of institutional quality used to place countries on either the left or the right hand side of the border.

2.3.1 Graphical evidence of discontinuity at the border

The RD design offers visual evidence of the discontinuity in exploration activity at the national borders. Figure 2.3 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 borders and institutional quality, the discontinuity indicates the causal effect of differences either side the border.

Figure 2.4 illustrates the data along a single border. Here we observe recent drilling (over

[23] D in equation 2.1 and 2.3 is our instrument and is defined based on institutions measured pre 1966. I in equation 2.2 and 2.3, on the other hand, could have been the institutional quality present in the 1966-2010 period. The instrument D deals with endogeneity and measurement error. However, the actual level of institutions factored into decisions vary depending on, at least, the timing of drilling. Measuring I as some average across the years 1966-2010 would therefore also be inapt. In lack of a clearly better alternative, we stick to the transparent alternative of measuring I in the pre-drilling period.

[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.

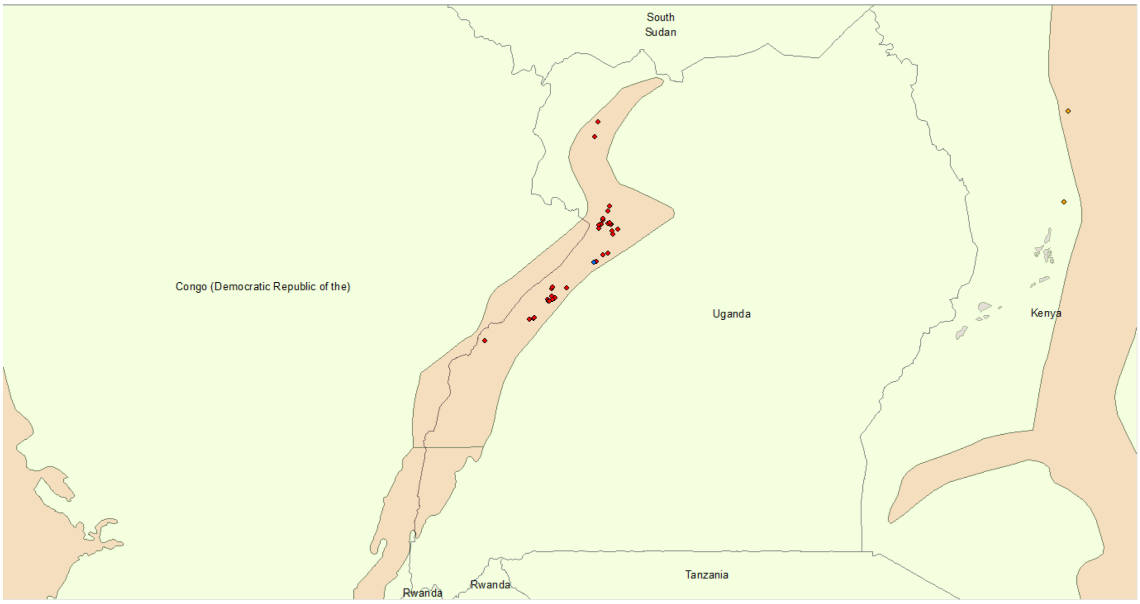

Figure 2.4: Map of the Albert Rift region along the Uganda-DRC border.

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

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.

2.3.2 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 2.2.2. The first and second column of table 2.1 present our baseline estimates of the border on the number of wells (N), with and without controls. The estimate changes very little when we include controls, increasing confidence that our RD design is valid.

Table 2.1: Baseline estimates: Direct effect of crossing the border

Dependent variable →	N: Number of wells	
	(1)	(2)
D = 1 rhs	4.722***	4.572***
	(1.243)	(1.246)
ln Area		0.588***
		(0.063)
Landlocked		-1.992***
		(0.317)
$\bar{N}_{left}$	4.81	4.81
$\hat{\tau}/\bar{N}_{left}$	0.98	0.95
Observations	1197	1197
Countries	29	29
Neighbours	40	40
R-sq	0.13	0.15

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). Dependent variable is the number of exploration wells for each 100 meter wide distance-bin in country i at the border with country j. Third-order polynomials in distance from the border included separately for left and right hand side. Distance from the border: 10 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 developed countries. Onshore and offshore wells. Excludes country-distance-bins without observed wells. $\hat{\tau}/\bar{N}_{left}$ 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 third-order polynomial in distance.

The first row in the lower part of the table, indicated with $\bar{N}_{left}$, provides the mean number of wells drilled on the left hand side of the border. The row below, $\hat{\tau}/\bar{N}_{left}$, presents the ratio between the estimated coefficient and the left-hand-side mean.[25] Crossing the border creates a jump of about 5 wells. This estimate suggests that the crossing to the 'better' side of the border

[25]The mean on the left hand side, $\bar{N}_{left}$, 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.

increases, on average, the number of wells drilled by about 95% or a factor of 2; from 4.81 wells on the 'bad' side of the border, to 9.38 on the 'good' side of the border. These estimates are based on a cross section of wells drilled from 1966 to 2010. The number of wells is in other words an accumulated figure over 45 years and the estimates can therefore be interpreted as close to long-run estimates.

These 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 than other economic activity in its relation to institutional quality measured as democracy.

2.3.3 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 2.2.2. 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.

The 2SLS-estimates are presented in table 2.2 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 0.27 on average. For comparison, the score

Table 2.2: 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	4.572***	6.944***	5.380***	-7.820***	7.214***
	(1.246)	(2.038)	(1.605)	(1.954)	(2.156)
$\bar{N}_{left}$	4.81	4.45	4.83	4.61#	4.45
$\hat{\tau}/\bar{N}_{left}$	0.95	1.56	1.11	1.70	1.62
Observations	1197	1228	1228	1228	1228
Countries	29	30	30	30	30
Neighbours	40	39	39	39	39
R-sq	0.15	0.17	0.15	0.19	0.18
	Scaled effect (second stage):				
Institutional quality	16.957***	1.009***	1.783**	-2.833***	2.930***
	(5.442)	(0.351)	(0.716)	(0.882)	(0.977)
F instr	131.77	112.79	44.30	72.04	89.90
	Effect of the border on institutional quality (first stage):				
D = 1 rhs	0.270***	6.885***	3.017***	2.760***	2.463***
	(0.023)	(0.648)	(0.453)	(0.325)	(0.260)

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 2 in table 2.1, except for the institutional measures in column 2-5, which are all 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.

varies on a 0-1 scale and its mean and standard deviation are 0.64 and 0.20 in our developing country sample. The second stage estimate is simply the combination of the first stage and the 'reduced form' estimate, i.e. $4.57/0.27 \approx 17$. To get a sense of the magnitude, a difference of 0.29 in the Freedom House score between DRC and Uganda translates into a predicted difference of 16 wells over the period 1966-2010. A one standard deviation increase in FH corresponds to an increase of around 72% more wells, moving from the 'bad' side to the 'good' side.[26]

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.

[26]As an example, this is found by $\hat{\gamma} * std.dev./\bar{N}_{left} = 16.957 * 0.204/4.81$

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 2.2 demonstrates that the jump in the number of exploration wells at the border holds true also for the Polity IV measures, and varies from 111% for democracy to 170% for autocracy.[28] Note that the number of observations is the same in all four columns using the different Polity IV measures and the variation in the estimates is entirely driven by the fact that the ranking of certain countries depend on the measure used. Our baseline estimate of 95% 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 positive coefficient is by construction. 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 118% and 173%. The comparable effect for the Freedom House measure is 72%, again smaller than estimated with the polity IV measures.

Overall, table 2.2 confirms our finding across different measures of institutional quality, all broadly defined. Future research could usefully conduct a detailed decomposition of which aspects of the institutional setting that may be most critical for oil exploration decisions.

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

[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 170%: the mean on the 'bad' side is 12.43-7.82=4.61; and the relative jump therefore 7.82/4.61=1.70.

[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

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 2.4.3). For the literature concerned with the causality running in the opposite direction, 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.

The results presented in this section is an average effect across different types of companies. In section 2.5, we investigate heterogeneity across company types.[30]

2.3.4 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 chapter 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.[31]

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 show in section 2.4.1 that the likelihood of discovering

discoveries are causally linked to slower transitions to democracy.'

[30]Moving beyond the extensive margin of exploration drilling, we have also estimated the effect on the number of meters drilled (in total and average per well). We find positive effects for both, although the effect on the average depth per well is smaller in magnitude and statistically less robust. These results are presented in the online appendix 2.8.2. Unfortunately, we do not have data on capital expenditures (CAPEX) and operational expenditures (OPEX) at the well level.

[31]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 be captured by the country and neighbouring fixed effects in column 3-5 in table 2.5.

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 less local: the specification test reported in appendix table 2.9 presents results for distances from the border of up to 250 km.[32] 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 and/or neighbour fixed effects, which controls for unobserved country-wide geographic features (see section 2.4.2).

Table 2.3: Controlling for ethnicity

Dependent variable →	N: Number of wells			
	(1)	(2)	(3)	(4)
Specification →	Ref	E FE	g	E FE; g
D = 1 rhs	3.440***	3.036***	2.999***	2.657***
	(0.473)	(0.526)	(0.444)	(0.572)
$\bar{N}_{left}$	1.54	1.54	1.54	1.54
$\hat{\tau}/\bar{N}_{left}$	2.23	1.97	1.95	1.72
Observations	1357	1357	1357	1357
Countries	27	27	27	27
Neighbours	38	38	38	38
R-sq	0.22	0.15	0.27	0.18

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 includes Ethnic fixed effects. Column 3 and 4 includes third-order polynomials (indicated g) in the distance to nearest ethnic boundary, estimated separately for left and right hand side of the border. 112 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 2.1.

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. In practice, these three premisses seem unlikely to be present simultaneously.[33]

[32] In the sample of developing countries, the mean distance from a well to the nearest border is 53 km and the 95th percentile is 187 km.

[33] 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 not presented to save space, we found that our result is robust to restricting the sample to the Sub-Saharan Africa region and to the Middle East and North Africa region. Furthermore, 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.

Nevertheless, geo-coded data on ethnicity and the fact that ethnicity often stretches across national borders (Michalopoulos and Papaioannou, 2014), allow us to control for at least one key aspect of culture by including ethnicity fixed effects. In table 2.3 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.[34] 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 third-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 third-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 chapter 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 172%. To the extent ethnicity captures culture, the results in table 2.3 provide strong 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 big difference in attracting investments in oil exploration at the margin.

2.4 Validity of the identification strategy

In this section we first examine potential threats to the internal validity of our estimates and finish with a brief discussion of the issue of external validity.

[34]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).

2.4.1 The likelihood of discovering oil and path dependency

Table 2.4: Discovery rates and path dependency

	(1)	(2)	(3)	(4)	(5)
Dependent variable →	Disc. rate	Disc. rate	N	N (0s)	N (0s)
D = 1 rhs	-0.026	-0.026	3.579***	0.298***	0.241***
	(0.100)	(0.099)	(0.666)	(0.087)	(0.046)
Wells pre 1966		-0.000	1.205***		1.692***
		(0.004)	(0.219)		(0.071)
$\bar{N}_{left}$	0.66	0.66	4.81	0.27	0.27
$\hat{\tau}/\bar{N}_{left}$	-0.04	-0.04	0.74	1.11	0.90
Observations	1197	1197	1197	11800	11800
Countries	29	29	29	29	29
Neighbours	40	40	40	40	40
R-sq	0.02	0.02	0.35	0.05	0.39

Notes: Dependent variable is the share of non-dry wells in column 1-2 and the number of exploration wells in column 1-4. Both are calculated for each 100 meter wide distance-bin in country i at the border with country j. Column 5-6 include zeros for distance-bins without drilling. Controls for country size and landlocked status are included. Otherwise as described by the note of table 2.1.

In contrast to the number of wells drilled, whose distribution is shown to be discontinuous across borders, the likelihood of finding oil does not appear to differ across the borders. Column one in table 2.4 presents estimates with the discovery rate as the dependent variable. The discovery rate is defined as the share of non-dry wells in a given distance-bin, which is the ex-post likelihood of discovering oil and the closest we get to observing the true geology.[35] The insignificant border-dummy suggests that the discovery rate is continuous across the border and provides support for our RD-design. In other words, there is no evidence to suggest that countries with better institutions see more drilling due to better geology or better ex-ante knowledge of the geology.

Although the oil 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

[35] 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.

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.[36] 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, this serves a similar function of 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 may indicate that additional wells have higher expected value due to, for example, infrastructure already in place. The coefficient on the border dummy is now somewhat reduced, which could indicate that issues related to weak institutions can be partly overcome, for example through concerted efforts to improve the relationship between the oil industry and the government. However, the message of the chapter is not altered and the estimate in column three demonstrates that our results are not driven by some accidental early discoveries and path-dependency since then. Common infrastructure could also lead to spatial correlation and potentially deflate our standard errors. In appendix table 2.11 we present our baseline model with standard errors adjusted for spatial correlation and the result holds.[37]

As we include only distance-bins with observed wells in our baseline specification, a question can be raised about the continuity of our running variable, the distance to the border. In column four and five we now include all possible distance-bins for the borders included in our baseline sample, i.e. within 10 km of the border we assign zero wells to distance-bins without observed wells. This increases our sample size tenfold. Column five differs from column four by the inclusion of the number of wells drilled before 1966. The coefficient on the border dummy now takes a lower value, but so does the mean number of wells. The effect of interest is the relative jump at the border presented in the row $\hat{\tau}/\bar{N}_{left}$, as usual. Our baseline estimate is a jump of 95% at the border (column two, table 2.1), while the estimate when zeros are included is 111%. When wells drilled before 1966 are included, the estimate is 74% in the baseline sample (column three, table 2.4) and 90% when the zeros are included. We therefore conclude

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

[37]In appendix table 2.11, we also include a column including zeros, as in column four of table 2.4.

that any lack of continuity in the running variable caused by excluding distance-bins without drilling does not seem to bias our estimates away from zero.

2.4.2 The role of offshore borders, military disputes and unobservables

Table 2.5: Offshore borders, military disputes and unobservables

Dependent variable →	N: Number of wells				
	(1)	(2)	(3)	(4)	(5)
Sample/Specification →	Onshore	Non-MID	C FE	N FE	C, N FE
D = 1 rhs	5.176***	5.719***	4.080***	3.364***	4.340***
	(1.167)	(1.247)	(1.204)	(1.143)	(1.360)
$\bar{N}_{left}$	4.07	4.93	4.81	4.81	4.81
$\hat{\tau}/\bar{N}_{left}$	1.27	1.16	0.85	0.70	0.90
Observations	939	983	1197	1197	1197
Countries	27	25	29	29	29
Neighbours	38	34	40	40	40
R-sq	0.16	0.18	0.16	0.17	0.18

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 excludes borders with recorded military interstate dispute after 1945. Column 3-5 includes country fixed effects or/and neighbour fixed effects as indicated. Controls for country size and landlocked status are included, except for in column 3 and 5. Otherwise as described by the note of table 2.1.

The formal designation of offshore borders has been a largely latter-half of twentieth century phenomenon, in contrast to onshore borders which have existed in various legal forms for centuries. 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).[38] By this time, the location of oil, both known and expected, might be thought to have

[38] 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 nation's 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 EEZ's typically extend 200 nautical miles from a nation's 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 country's neighbour. While inherently more contestable than onshore borders, the position of these borders typically pre-dates oil exploration

greater influence on the location of offshore borders, thus undermining a clean exogeneity argument for all offshore borders. Furthermore, offshore borders are arguably more contestable, since no one lives there and there is a less long established legal claim. In column one in table 2.5 we exclude offshore wells. The estimate of the effect of the border is, if anything, larger for this sample of onshore wells only, suggesting that our findings are not driven by the strategic location choice of offshore borders by countries with better institutions than their neighbours.

A related argument is that colonial powers may have been 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.

Could it be that our findings are driven by borders moving due to interstate territorial conflicts? The argument could be that nation states with relatively high institutional quality compared to their neighbours have been able to redraw borders via invasion and occupation of neighbouring territories who possess oil or promising geology. The border would have to be permanently moved due to the presence of oil, i.e. a lasting occupation (not merely threat of, or attempted occupation). Caselli et al. (2012) 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.[39] To deal with the potential issue of moving borders we consider throughout this chapter only stable borders. This means that we always exclude disputed offshore border areas, as coded as such in the EEZ Maritime Boundaries dataset (Claus et al., 2013). In addition, we only include borders which have not changed since 1965, according to Weidmann et al. (2010).[40]

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.

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

[40] 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 (COW, 2008) project which

Even for our stable borders, inter-state conflict may affect the likelihood of oil exploration. Column two in table 2.5 therefore excludes borders across which there has been at least one incidence of Militarized Interstate Dispute (MID).[41] 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 countries with better institutions are also better at reporting drilled oil wells and our findings are just an artifact of such reporting bias? We re-estimate our baseline model with country fixed effects, which effectively control for reporting bias at the country-level and other country-specific unobservable heterogeneity. The resulting coefficient on the border dummy, presented in column three, is very similar to our baseline estimate. The PathFinder database is considered to have as comprehensive country coverage as any single dataset on oil exploration, although we cannot rule out measurement error resulting from mis-reporting or under-reporting of some exploration wells. Similarly, to take into account potential unobservable heterogeneity of the neighbours, e.g., developed country status, we include in column four neighbour fixed effects and in column five country and neighbour fixed effects together. In both these cases, the estimated coefficient on the border dummy is similar to our baseline estimate.

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

captures the territorial changes for the countries of the world. This dataset is extended by Gleditsch and Ward (GW) (1999) to include an extended list of independent states which addresses some limitations of the COW list, and has also gained a wide acceptance in the research community. 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.

[41] 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. (2012) 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.

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.[42] On the contrary, Figure 2.4 plots the pattern of exploration wells in the Albert Rift, an oil basin on the border between Uganda and the Democratic Republic of Congo, and provides a visual illustration of what seems to be a typical pattern. Borders have been historically determined with no obvious regard for the (subsequently discovered) location of oil basins. In this spirit, visual examination of the global map presented in figure 2.6 reveals to us no obvious relationship between geological basins and national borders.

2.4.3 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).[43]

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.[44] Furthermore, effects of previous oil exploration should be accounted for by the lagged dependent variable included in column three in table 2.4.

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

[43] 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.”

[44] 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.

2.4.4 External validity

Is the effect of institutional quality on oil exploration different at national borders to elsewhere in the country? Within countries, unitization, i.e. 'the joint development of a petroleum resource that straddles territory controlled by different companies', is often regulated by law, for example 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).

A second reason why effects may be different in border regions than elsewhere, is geology. 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 the internal validity.

Table 2.6: Thick borders

Dependent variable →	N: Number of wells				
	(1)	(2)	(3)	(4)	(5)
Thickness of border →	0 km	10 km	20 km	40 km	80 km
D = 1 rhs	0.940***	0.356***	0.441***	0.536***	1.227***
	(0.227)	(0.086)	(0.102)	(0.144)	(0.276)
$\bar{N}_{left}$	2.39	1.93	1.75	1.60	1.38
$\hat{\tau}/\bar{N}_{left}$	0.39	0.18	0.25	0.33	0.89
Observations	9995	9390	8798	7759	6138
Countries	43	43	43	38	32
Neighbours	57	56	56	55	53
R-sq	0.10	0.10	0.09	0.07	0.08

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 number of distance-bins on which the standard errors are clustered varies from 3468 to 2670. Controls for area size and landlocked status are included. Otherwise as described by the note of table 2.1. For highest comparability with our other results, we include in this table third order polynomials, although we use a distance of up to 250 km from the border.

Table 2.6 seeks to investigate 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, while we in column two to five successively exclude distance-bins closest to the border. 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, four and five we increase the thickness to 20 km, 40 km and 80 km. The idea is that the resource competition effect should be less relevant the further away from the border we get. The estimated jump in the number of wells is 39% in column one. It falls to 18% for the 10 km thick borders, and increase again to 25%, 33% and 89% for the 20, 40 and 80 km thick borders. 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 chapter should, however, be interpreted 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 identical.

Recall that we have restricted our sample to include developing host countries only, excluding Eastern Europe and Central Asia (ECA).[45] For developed countries, the difference in measured institutional quality across the borders turns out to be very small on average and we do not estimate a robust effect of crossing the border for this sample. This is consistent with our hypothesis. A small gap in institutional quality observed at the border generates a small expected payoff of choosing one side of the border over the other. It may also be that the effect of institutions on oil exploration is non-linear, i.e. improvements beyond a certain level of institutional quality may have little effect.

[45]We have estimated the models presented in the upper panel of table 2.2 for 10 km and 250 km distances from the border on the following pooled samples: i) the sample of developed plus developing countries excluding ECA (up to 64 host countries and 87 neighbouring countries) ii) the sample of developing countries including ECA (up to 49 host countries and 71 neighbouring countries) and iii) the global sample including ECA (up to 70 host countries and 96 neighbouring countries). The coefficient on the border dummy takes the same sign as in table 2.2 and is significant at the 5% or 1% level in 29 out of these 30 regressions. The one exception is *Democ* in the 250 km distance case in the developing countries including ECA sample, in which the border dummy is positive but insignificant. In the interest of saving space, these results are available on request.

2.5 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, like hedge-funds, and therefore capable of handling high risks.

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.[46] 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.

Starting out from our baseline results in table 2.1, we test in table 2.7 for heterogeneity

[46] 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).

Table 2.7: Heterogeneity across company types

Dependent variable →	*N*: Number of wells				
	(1)	(2)	(3)	(4)	(5)
Company type →	IOC6	IOC6+	NOC	NOCH	OTH
D(FH)	6.269***	6.179***	8.213***	10.087***	7.987***
	(1.880)	(1.894)	(1.870)	(2.159)	(1.920)
D(FH) x COMP	4.580***	4.508***	-2.335***	-2.391***	-1.709***
	(0.739)	(0.677)	(0.401)	(0.406)	(0.396)
COMP	-0.852***	-1.066***	1.206***	1.615***	-1.241***
	(0.170)	(0.135)	(0.179)	(0.221)	(0.158)
Observations	1248	1250	1323	1254	1324
Countries	30	30	30	30	30
Neighbours	31	31	31	31	31
R-sq	0.27	0.26	0.16	0.19	0.22

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 2.1. 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.

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. 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.[47] The coefficient on the border dummy is always positive and significant. The coefficient on the interaction term is relatively large and 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.

[47] Note that the dummies indicating company types are also included separately in the regressions in table 2.7.

2.6 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 in developing countries. A one standard deviation increase in the augmented Freedom House score corresponds to around twice as much drilling over the sample period.

This chapter 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. This helps to explain why some regions of the world, such as sub-Saharan Africa, are likely to be 'under-explored' with respect to geology alone. 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. Although countries with poorer institutional quality may eventually catch up and discover oil, imperfect capital markets and current development challenges mean that delayed discovery may carry an economic cost for citizens.

2.7 Summary statistics

Table 2.8: Descriptive statistics, baseline sample

	Observations	Median	Mean	Std. dev.	Min	Max
Wells per pair	1197	68	252.8	260.6	1	782
Wells (per 0.1 km dist-bins)	1197	2	3.137	4.462	1	85
Wells pre 1966	1197	0	0.655	1.777	0	35
Wells (0s)	1197	2	3.137	4.462	1	85
Wells (0s) pre 1966	1197	0	0.655	1.777	0	35
D = 1 if non-dry	1197	0.667	0.608	0.402	0	1
Wells x Depth	1143	4.205	8.169	14.09	0.190	281.0
Depth	1143	2.487	2.521	0.955	0.190	6.151
Area	1197	1285216	2521446.0	2771062.1	91862	8511920
Landlocked	1197	0	0.0117	0.108	0	1
Distance	1197	1.750	1.097	5.653	-9.950	9.950
D = 1 rhs FH	1197	1	0.609	0.488	0	1
FH	1197	0.680	0.644	0.204	0.120	0.943
D = 1 rhs Polity	1228	1	0.507	0.500	0	1
Polity	1228	-1.524	-0.415	5.863	-9	10
D = 1 rhs Democ	1228	1	0.570	0.495	0	1
Democ	1228	2.048	3.374	3.204	0	10
D = 1 rhs Autoc	1228	1	0.555	0.497	0	1
Autoc	1228	3.571	3.788	2.820	0	9.074
D = 1 rhs ConsEx	1228	0	0.480	0.500	0	1
ConsEx	1228	3	3.725	2.130	1	7

2.8 Appendix

2.8.1 Specification analysis

This appendix explains our choice regarding the distance from the border and the order of the polynomial used in our regressions. There is in principle a trade-off between bias and precision regarding the distance limit from the border. Data far away from the border may induce a bias in the estimated jump at the border, while employing more data increases the precision of the estimate (Lee and Lemieux, 2010). The polynomials, the f-functions in equation 2.1-2.3, are included to control for the underlying distribution of the number of wells as a function of the distance to the border. We search for the appropriate specification by varying these two dimensions and compare results. To help in evaluating the different combinations, we follow Lee and Lemieux (2010) in implementing two formal criteria. First, the Akaike Information Criterion (AIC) summarizes the balance between the goodness of fit versus the order of complexity of the model, and we seek a lowest possible AIC.[48] Second, the appropriateness of the polynomials in picking up the relationship between the number of wells and the distance to the border is compared to the non-parametric alternative of bin-dummies; we include a dummy for each 1 km from the border and test their joint significance by an F-test. The idea is that a higher order polynomial makes it more likely to reject the significance of the bin-dummies. If the bin-dummies are insignificant, the polynomial does a good job in controlling for the density of wells across different distances from the border.

Our estimates of the effect of crossing the border across different distances and for different polynomials are presented in table 2.9. Going from the left to the right, we vary the distance from the border from 5 km to 250 km. Going from the top to the bottom, we vary the order of the polynomial from six to zero.[49] Order zero corresponds to a comparison of the means on each side of the border, while order one corresponds to a 'local' linear regression with a rectangular kernel (Lee and Lemieux, 2010).

[48]We calculate: $AIC = Nln(\hat{\sigma}) + 2p$, where N is the number of observations, $\hat{\sigma}$ is the mean squared error of the regression and p is the number of parameters in the regression model (Lee and Lemieux, 2010)[p. 326].

[49]We stopped at a polynomial order of six to keep the specification relatively simple.

Table 2.9: Specification test: Estimated coefficient on the border dummy

Dependent variable →	N: Number of wells					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance from border →	5 km	10 km	25 km	50 km	100 km	250 km
Polynomial order ↓						
Zero	1.305***	0.666**	0.544***	0.427***	0.419***	0.354***
	(0.442)	(0.268)	(0.138)	(0.088)	(0.063)	(0.052)
	0.42	**0.25**	**0.26**	**0.23**	**0.25**	**0.22**
	0.00	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
One	3.353***	2.171***	0.930***	0.798***	0.761***	0.585***
	(0.877)	(0.591)	(0.358)	(0.224)	(0.148)	(0.099)
	0.79	**0.58**	**0.31**	**0.32**	**0.36**	**0.30**
	0.04	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Two	4.709***	3.564***	1.967***	1.145***	0.821***	0.813***
	(1.349)	(0.918)	(0.569)	(0.383)	(0.254)	(0.160)
	0.96	**0.82**	**0.53**	**0.36**	**0.32**	**0.39**
	0.38	*0.01*	*0.00*	*0.00*	*0.00*	*0.00*
Three	5.014***	4.572***	3.265***	1.737***	1.193***	0.940***
	(1.917)	(1.246)	(0.745)	(0.544)	(0.366)	(0.227)
	0.96	**0.95**	**0.79**	**0.48**	**0.40**	**0.39**
	0.32	*0.44*	*0.00*	*0.00*	*0.00*	*0.00*
Four	5.853***	5.108***	3.862***	2.682***	1.691***	0.943***
	(2.106)	(1.603)	(0.973)	(0.675)	(0.474)	(0.296)
	0.94	**1.00**	**0.88**	**0.68**	**0.50**	**0.34**
	0.52	*0.60*	*0.00*	*0.00*	*0.00*	*0.00*
Five	7.198***	5.310***	4.202***	3.516***	1.921***	1.152***
	(2.379)	(1.945)	(1.204)	(0.794)	(0.591)	(0.368)
	1.14	**0.95**	**0.88**	**0.82**	**0.51**	**0.38**
	0.08	*0.59*	*0.00*	*0.00*	*0.00*	*0.00*
Six	8.293***	5.816***	5.022**	4.033***	2.624***	1.506***
	(2.242)	(1.519)	(2.076)	(0.852)	(0.692)	(0.439)
	1.27	**1.02**	**1.01**	**0.92**	**0.67**	**0.46**
	0.06	*0.04*	*0.00*	*0.00*	*0.00*	*0.00*
Observations	605	1197	2684	4583	7595	9995
Suggested poly. order:						
Bin dum. not joint. sign.	2, 3, 4	3, 4, 5				
AIC	2 (4)	4 (3)	6	6	6	6

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 distance from the border as indicated in the top row. The number of distance-bins on which the standard errors are clustered varies from 200 to 3468. Separate polynomials for left and right hand side included as indicated in the first column. Figures in bold is $\hat{\tau}/\bar{N}_{left}$, i.e. the coefficient on the border dummy divided on the mean number of wells on the left hand side. Figures in italics are p-values from a joint significance test of bin-dummies. See the text for further explanation. Otherwise as described by the note of table 2.1.

A first observation in table 2.9 is that all coefficients are positive and significant, providing support that the finding of this paper is qualitatively robust.

A second observation is about the magnitude of the estimated coefficients. Inclusion of distance-bins far away from the border decreases the size of the estimated coefficient. Given figure 2.3, this is no surprise as the density of wells, on both sides, increase towards the border. This is caused by the fact that the density of land area increases towards the border, as shown in the left hand side panel of figure 2.5.[50] There are two reasons for this. First, moving away from the border means moving towards the geographical 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 squared 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 geometrical 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.[51] This imposes curvature, the area size increases in a convex manner as one approaches the border. Decreasing the bandwidth and increasing the order of the polynomial increases the size of the estimated coefficient. A final reason for the shape towards the border may be competition for oil close to the border, as discussed in section 2.4.4.

The optimal order of the polynomial as suggested by the AIC criterion is listed in the bottom

[50] To make figure 2.5 we construct a grid-map of the globe, with cells of 50 km x 50 km as the unit of observation. The distance from the closest border to a well is approximated by the distance from the border to the centroid of the grid cell in which the well is located. Although 50 km x 50 km may seem like a relatively small cell size in the context of the surface of the globe, it is a relatively large size in the context of exploration licenses and oil fields. As an example, average field size in the Norwegian part of the North Sea is about 500 km^2, one fifth of our cell size. However, there seems to be large variation across countries in terms of the area covered by exploration licenses. The fact that cells are smaller close to borders, as the standard grid map oriented orthogonally in the north-south direction is separated into different countries by borders not oriented orthogonally in the north-south direction, makes in practice the measure of distance between the centroid and the closest border continuous. As measuring distance to the centroid rather than to a well introduces considerable noise into our measurement of distance, we prefer the well-based approach used in the main specifications. For completeness, the right hand side panel of figure 2.5 presents the distribution of the number of wells in the cell-level data.

[51] 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.

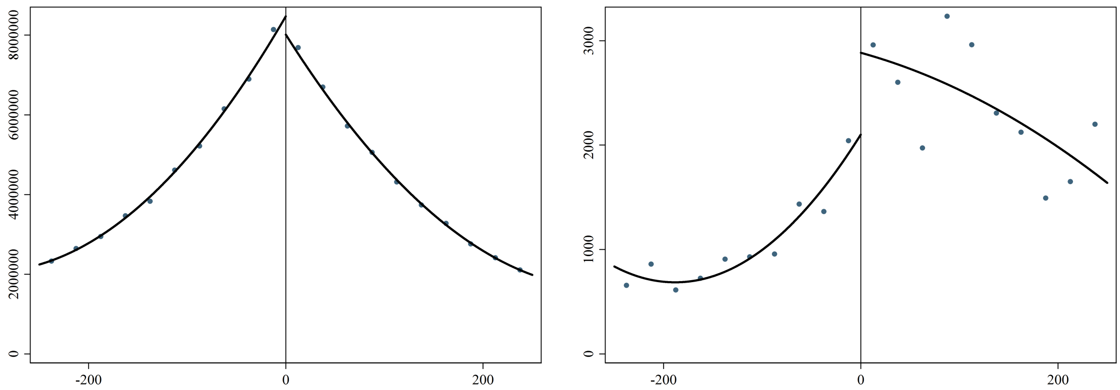

Figure 2.5: Cell data: Area size and number of wells

Notes: left panel shows the area covered per 25 km bins, while the right hand panel shows sum of wells per 25 km bins. Based on 250 km distance from the border. Lines show 2nd order polynomials. Based on cell level data and augmented Freedom House Political Rights score.

row of the table. Order six is suggested to be optimal for distances higher than 10 km. For 5 and 10 km, order 2 and 4 are suggested, respectively, although the AIC is almost identical for order 4 for 5 km and for order 3 for 10 km. Ceteris paribus we have a preference for lower ordered polynomials to keep the model as simple as possible.

The p-values from the joint significance test of the bin-dummies are reported in the fourth column for each polynomial order. The bin-dummies are often found to be jointly significant, indicating that even the high-powered polynomials are not able to fully account for the curvature in the underlying distribution of wells. However, for the distances 10 and 5 km, the bin-dummies are found to have less explanatory power.

Although table 2.9 reveals that the exact size of the jump at the border is quite sensitive to the choice made regarding the distance from the border and the order of the polynomial, the economically interesting estimate is the relative increase in drilling as the border is crossed from the left to the right. We arrive at a measure of the relative jump by judging the coefficient in terms of the mean on the left hand side. This mean varies across different bandwidths and polynomials and we report the estimated coefficient divided on this mean, $\hat{\tau}/\bar{N}_{left}$, in bold types (third row).[52] Across all configurations, the estimated increase in drilling varies from 22% (250

[52]As in previous tables in this paper, the mean on the left hand side, $\bar{N}_{left}$, is estimated as the constant in a separate regression for the left hand side including the relevant polynomial as controls.

km, zero order) to 127% (5 km, sixth order).

We suggest that a distance of 10 km from the border in combination with a third order polynomial represent a reasonable choice for our baseline results. This is however a judgement call and table 2.9 together with figure 2.3 should allow the reader to make an independent judgement.

2.8.2 Extra tables and descriptive statistics

Above we estimated the effect on the number of wells drilled, i.e. the responsiveness along the extensive margin of oil exploration. We have also estimated the effect on the depth of drilling, which is a proxy for costs; the deeper or longer the well, the higher the investment sunk in the well (Boyce and Nostbakken, 2011). The results, presented in table 2.10, show that crossing the border increases the sum of meters drilled (number of wells * depth) by 172% and the number of wells by 14%, although the latter is less robust. The estimate for the number of wells is 99% in this sample.

Table 2.10: Depth of drilling

	(1)	(2)	(3)
Dependent variable →	N x $Depth$	N	$Depth$
D = 1 rhs	18.933***	4.901***	0.314*
	(3.289)	(1.176)	(0.175)
$\bar{N}_{left}$	11.00	4.95	2.21
$\hat{\tau}/\bar{N}_{left}$	1.72	0.99	0.14
Observations	1143	1143	1143
Countries	29	29	29
Neighbours	40	40	40
R-sq	0.14	0.16	0.07

Notes: Dependent variable indicated in column heading (N: Number of wells). $Depth$ is observed for each well, measured in km along the drilling path, and included as the average per unit of observation. Restricted to wells for which we observe the depth. Controls for country size and landlocked status are included. Otherwise as described by the note of table 2.1.

In table 2.11 we investigate robustness of our inference to spatial correlation across observations, and the results hold.

Table 2.11: Standard errors adjusted for spatial correlation

	(1)	(2)
Dependent variable →	N	N (0s)
D = 1 rhs	4.572*	0.298***
	(2.610)	(0.102)
$\bar{N}_{left}$	4.81	0.27
$\hat{\tau}/\bar{N}_{left}$	0.95	1.11
Observations	1197	11800
Countries	29	29
Neighbours	40	40
R-sq	0.43	0.08

Bin width 0.1 km. Dist cutoff: 10 km. Developing=1. Onshore=2. Polynomial order: 3.

Notes: Column one as column two in table 2.1 and column two as column four in table 2.4, 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, 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 2.12 presents the countries included in our samples. Figure 2.6 shows a map of geological basins, where oil may be found, together with onshore and offshore national borders.

Table 2.12: 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			26	27	3	3	20	23
EAP	Indonesia	IDN	119	649	1267	2537	8	9	186	219
EAP	Malaysia	MYS	23	26	506	604	119	651	855	1904
EAP	Myanmar	MMR	7	7	131	156	48	77	312	440
EAP	Philippines	PHL	18	21	56	64	17	19	112	135
EAP	Thailand	THA	15	16	440	614	4	4	308	371
ECA	Turkey	TUR							4	4
HI	Australia	AUS							1	1
HI	Israel	ISR					2	2	2	2
HI	Italy	ITA					40	56	153	193
HI	Japan	JPN							26	27
HI	Saudi Arabia	SAU					4	4	14	14
HI	Singapore	SGP					14	16	509	745
HI	Spain	ESP					8	9	107	134
HI	United States	USA							1	1
LAC	Argentina	ARG	125	329	1564	3070	58	78	330	426
LAC	Bolivia	BOL	12	14	326	374	116	331	505	1213
LAC	Brazil	BRA	192	781	1303	3019	3	3	193	206
LAC	Chile	CHL	46	64	79	105	94	292	1381	2862
LAC	Colombia	COL	28	32	391	540	15	19	143	162
LAC	Ecuador	ECU	32	38	303	346	64	79	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	145	158				
LAC	Panama	PAN					9	10	54	59
LAC	Paraguay	PRY					104	484	864	1863
LAC	Peru	PER	46	58	225	242	18	20	249	277
LAC	Uruguay	URY					2	2	34	36
MENA	Algeria	DZA	35	37	749	832	56	70	479	547
MENA	Iraq	IRQ			5	5	2	3	3	4
MENA	Jordan	JOR	8	9	18	19				
MENA	Libya	LBY	17	18	476	520	38	48	410	459
MENA	Morocco	MAR	17	20	207	251			60	60
MENA	Tunisia	TUN	102	137	399	492	29	29	621	701
SA	India	IND	74	124	204	285	8	8	71	78
SA	Nepal	NPL			1	1			4	4
SA	Sri Lanka	LKA			2	2	35	56	59	88
SSA	Angola	AGO			5	5				
SSA	Benin	BEN	2	2	11	11	83	472	227	665
SSA	Cameroon	CMR	41	51	127	161	98	782	612	2206
SSA	Centr. Afr. Rep.	CAF							14	15
SSA	Chad	TCD			38	40	3	3	31	32
SSA	Ethiopia	ETH			7	7	1	1	7	7
SSA	Ghana	GHA	8	9	11	12			6	6
SSA	Kenya	KEN	2	2	23	24	8	8	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	19	20	11	19	18	27
SSA	Niger	NER	2	2	26	27			11	12
SSA	Nigeria	NGA	182	1255	836	2868	43	53	138	172
SSA	Sierra Leone	SLE			3	4			4	4
SSA	Somalia	SOM	2	2	9	9	2	2	24	25
SSA	South Africa	ZAF					3	3	12	13
SSA	Tanzania	TZA	18	26	33	42	2	2	4	4
SSA	Togo	TGO			4	4	8	9	12	13
SSA	Uganda	UGA							2	2
SSA	Zambia	ZMB			2	2			7	7
Total			**1197**	**3755**	**9995**	**17518**	**1197**	**3755**	**9995**	**17518**

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). 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). Europe & Central Asia (ECA) is excluded in this paper.

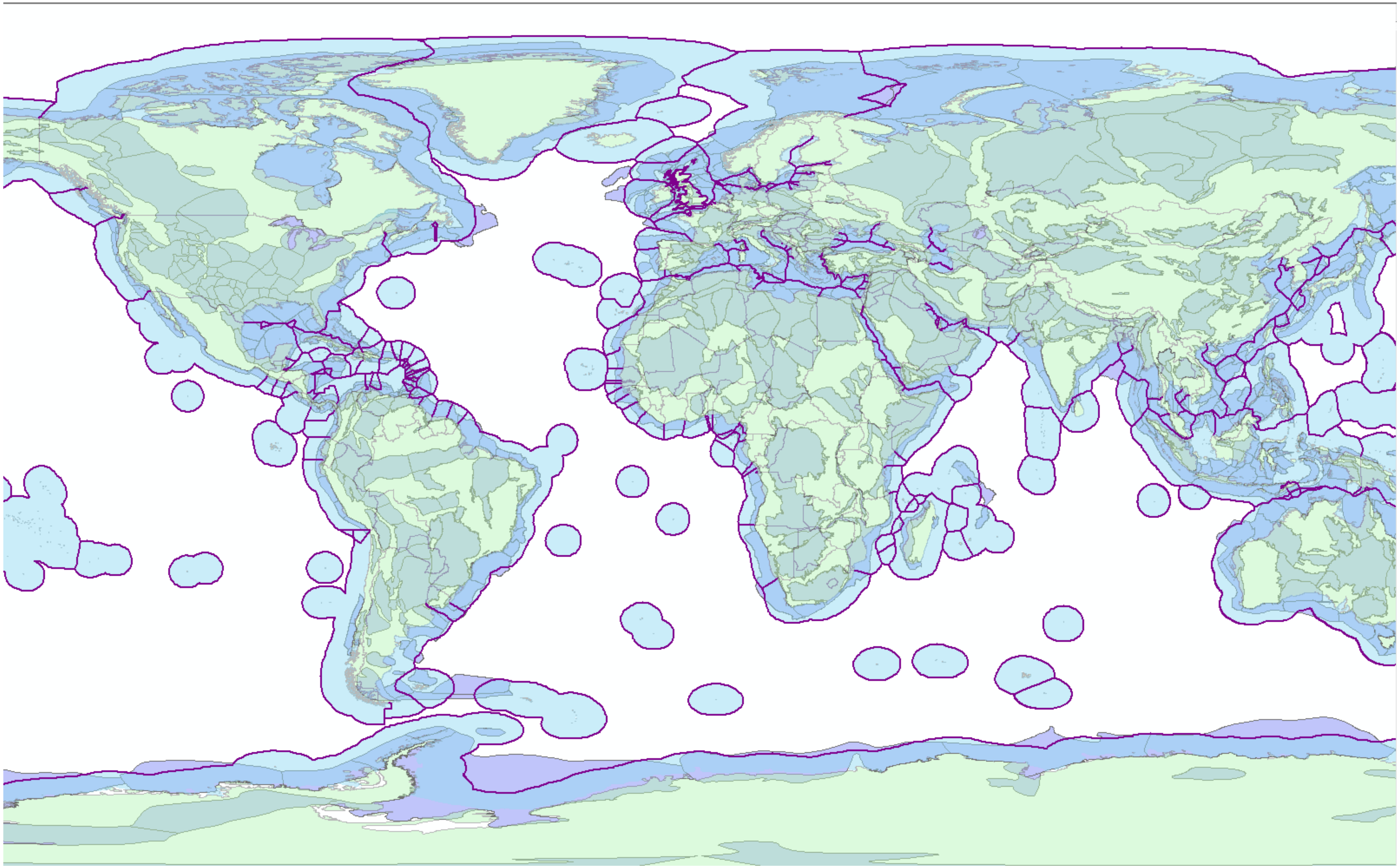

Figure 2.6: National borders and geological basins

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

Chapter 3

The Local Effects of Resource Extraction

James Cust

Abstract

This paper studies the economic effects of industrial mining on the local economy. Using a new dataset on mine locations across Indonesia, linked to household survey data and nighttime satellite data, estimates suggest mines have an important, though spatially concentrated footprint. By exploiting the location, timing and scale of mines, the causal effect of mining activity on the local economy can be identified. This work contributes to the growing literature on the economic linkages from natural resource extraction.

I provide estimates for the size and intensity of the economic footprint of mining. Results from nighttime lights estimates show the typical economic footprint from mining extends up to 15km. Estimates are presented for the effect of both mine opening and mine closure - with evidence for a roughly symmetric *boom and bust* effect.

Furthermore within this average 15km footprint the mine affects the structure of the labour market. Consistent with a local variant of the Dutch disease model developed by Corden and Neary (1982), and predicted by a simple model, results point to a shift of labour from the traded sector to the booming mining and non-traded sectors. This evidence suggests that mines have a local crowding-out effect on other traded goods sectors, while stimulating expansion of the services sector. I present estimates for how the local labour market effect varies with distance and between foreign-owned and domestic-owned mines. Results suggest a more pronounced effect from foreign-owned mines suggesting heterogeneity in local sourcing practices or other externality effects.

3.1 Introduction

This work explores the local economic effects and labour market responses to industrial mining activity in Indonesia. This research question falls within a wider literature examining the impact of natural resources on economic development in low income countries. It also complements a strand of literature seeking to understand local labour market responses to shocks, such as those arising from large investment projects.

The question of whether natural resources are a blessing or a curse has been subject to extensive debate amongst economists. Typically this work has employed aggregate time-series and cross-country data to explore various theoretical settings for the resource curse hypothesis (Sachs and Warner (1995), Auty (2001), Papyrakis and Gerlagh (2004)). Much of the literature on the resource curse identifies some combination of conflict, weak institutions, ineffective government or displacement of other productive sectors (as in case of the Dutch disease) as contributing factors in the resource curse hypothesis (Corden and Neary (1982), Collier and Hoeffler (2005), Ross (2001)). Van der Ploeg (2011) provides a recent survey of the literature.

The impact of individual mining projects on the local economy is only partially understood. While mining can be characterized as typical of *enclave development* - an isolated or extractive economic activity with only limited interaction with surrounding economy (Hirschman, 1958) - it is also known that mining projects can be large sources of wealth, infrastructure and employment. The extent to which such benefits accrue to local communities, the spatial distribution of those benefits and the nature of this process are important questions for policymakers and companies alike. The research presented in this chapter examines these issues via estimation of the economic footprint and labour market response to mine opening and closure. I estimate the size of the mining footprint using measures of nighttime light intensity close to mining sites (and across mine status and across time). This allows calibration of estimates for labour market spillover effects using distance cut-offs for mining footprint. I then estimate the labour market participation of communities and individuals proximate to mines. I test predictions from the simple model to see whether mines generate a reallocation of labour to the resource sector and the local non-traded sector. I find that while direct employment by the mine from the local community appears modest (measured as a shift to the mining sector), there is a more notable shift from the traded to the non-traded sector.[1] This conforms to the model and also the standard predictions of the Dutch disease hypothesis, but replicated on a micro scale.

3.1.1 Literature

The micro evidence for the economic dynamics of resource extraction in developing countries has only recently become the subject of investigation. A number of papers have examined the effect natural resource prices or revenue shocks to regional economies where extraction activity takes place (Caselli and Michaels 2013, Angrist and Kugler 2008). Recent work has examined district-level effects, in particular in relation to the effect of resource prices on conflict and violence (Dube and Vargas, 2006), while others examine the effects of resource extraction on local employment and welfare (Aragon and Rud, 2013) in a core-periphery framework for a single large mine. More recently work has expanded the scope to examine the mining impacts on women (Kotsadam and Tolonen, 2013) or the implications for conflict (Berman et al., 2014) and the implications for the negative externalities (Aragon and Rud, 2012).

The contribution of my work is to expand upon two areas of the literature. First I examine the spatial component of these effects- in particular the relationship of distance from extraction activity to local spillover effects - to estimate the economic footprint of individual mines. Second I examine the evidence for the labour market impacts of natural resource extraction at a localised level, using nationally representative data.

This work utilizes spatially-linked resource extraction project data, satellite data and household survey data to examine local spillover effects and implications for the regional economy from discrete mining projects. It is the first work, I am aware of, to link satellite data to multiple extraction projects. This allows examination of the economic activity in the locality of mines, while the household survey data facilitates examination of labour market dynamics. In partic-

[1] Where the traded sector can either be defined as the manufacturing sector, the agricultural sector or both combined

ular this approach allows identification of effects at much greater spatial resolution; previous work has typically exploited district-level variation or distance to urban centres, rather than the proximity to extraction activity itself. Further, effects can be estimated locally and their spatial and temporal dynamics tracked. The dataset contains location and production data for over 300 individual mining projects.

This paper follows most closely from the work of Aragon and Rud (2013). They examine the effects of a change in procurement policy and expansion of a large Peruvian gold mine. Their identification strategy allows the examination of both the localised labour market and welfare impact, but also the spatial disaggregation of effects along a distance continuum from nearby urban centres. However, unlike Aragon and Rud, this chapter draws upon both nationally representative household panel data and includes data on around three hundred individual and geographically dispersed mining projects, encompassing those operating, closed and under development. I examine the spatial relationship between the mines themselves, the economic activity in their locale and affected communities (rather than the urban centre and its hinterland). Further, I am able to examine the within effects at both the level of local communities and of individual workers. The study also complements more recent work by Aragon and Rud (2012) in Ghana who find evidence for declining agricultural productivity within 20km radius of polluting gold mines.

The work fits into a wider literature examining both the spatial dimensions of economic development and the effects of exogenous shocks such as large investments or government projects, particularly effects on the labour market. Notable contributions include the examination of labour market effect from large-scale construction projects (e.g. Alaskan oil pipeline, Carrington (1996)), coal boom and bust shocks in the US (Black et al., 2005) and the effects of military base closures in the US (Hooker and Knetter (2001)). Others have explored the effect of Enterprize Zones, where little positive employment spillovers were found (Bartik (2003)).

Moretti (2010), Carrington (1996), and Black et al. (2005) examine the impacts across traded and non-traded sectors. They estimate local multiplier effects from different industries, and different sources of external shocks. Moretti finds large long-run multipliers from manufacturing and high tech job creation. In contrast, Black et al. (2005) find relatively modest multiplier effects arising from coal mining activity in the north-eastern US. They find that each additional mining job generates 0.17 jobs in the non-traded sector. Interestingly, the estimated effect is not symmetric. The loss of a mining job, associated with mine closure and the decline of the industry, resulted in the loss of 0.34 non-tradable jobs.

A closely related and active area of the literature uses new datasets to investigate causal and structural effects of infrastructure on regional economies. The structural approach has been favoured in recent work on trade and transportation infrastructure such as that of Michaels (2008), which looks at the effect of then-new US interstate highway construction in the 1950s and Donaldson (2010) looking at the effect of railroads in colonial India. A second strand of this literature examines the reduced-form effects of infrastructure on the local economy including Duflo and Pande (2007) who examine the distributional effects of dam construction in India, Dinkelman (2011) who evaluates the effect of extension of rural electricity infrastructure in South Africa and Banerjee et al. (2012) who examine the welfare-raising effects of improved transport infrastructure in China. These approaches use a combination of difference-in-difference estimation and instrumental variables.

Further, this work is related to and informed by a section of the literature concerned with the spatial dimension of specialisation and structural differentiation. In particular, we know that economic specialisation is driven in part by distance to centres of economic activity, where transportation is costly. The von Thunen model predicts concentric circles of specialisation in agriculture surrounding cities. More recent empirical work has sought to estimate specialisation with respect to distances from urban centres. For example, Fafchamps and Shilpi (2003) examine the role of distance and connectivity in rural developing countries as a driver of specialisation. My analysis takes mines, not urban centres, as the focus of economic activity and the source of external stimulus to rural economy. However, some of the dynamics are comparable. In particular, investigation of the non-linear effect of distance is included in the estimation strategy and appears to matter for the main results. My approach examines the evidence for general equilibrium effects on the labour market through changes in quantities of labour allocated across different sectors (and the proportions thereof). I examine evidence for sector reallocation, including the extent to which mines employ local workers, the new job opportunities created by local extraction activity, and any shift from dependence on agricultural employment to other non-farm related activities. My estimation strategy allows me to calculate quantity effects through a shift in workers between sectors.

This paper is organised as follows. Section 3.2 outlines the the analytical framework. Section 3.3 presents the data and the nature of the variation in the dataset. Section 3.4 presents the empirical framework. Then, in Section 3.5 and Section 3.6 I present results and discussion. Section 3.7 concludes.

3.2 Analytical Framework

Modern mines are a significant economic phenomena for their locality, both by the virtue of their direct employment and local spending effects, but also by their typically rural setting in areas of low population density, which otherwise depend on agriculture.

A simple model provides a setting for an empirical examination of the localised labour-market effects of resource extraction. I want to examine how an increase in activity in the resource sector affects labour market allocation across other sectors. In particular, I want to contrast the impacts on sectors that are relatively traded (i.e. take world or regional prices as given) and relatively non-traded (i.e. local price makers). Standard trade model approaches such the Salter-Swan framework and the 1-2-3 general equilibrium model allow characterisation of traded and non-traded goods sectors. However these models typically lack explicit modelling of factor markets and households, and thus are limited in drawing insights regarding likely effects of labour market allocations and welfare, particularly in a localised setting.

3.2.1 Set-up

This model builds on a Salter-Swan two sector framework, following the example of models with a third *boom* sector and inclusion of specific factor shares and markets, such as the Dutch disease model developed by Corden and Neary (1982). In my model labour is allocated between the three production sectors (a resource sector, a traded goods sector - agriculture and manufacturing - and non-traded - services - sector).[2]

The stylized model of local labour markets comprises a representative local household who consumes locally produced goods and chooses how to allocate labour across three sectors. The three sectors demand labour for production, and allow for generation of positive profits in each sector with decreasing returns to scale. The model extends the Corden and Neary framework[3] to include an explicit measure of resource sector profits repatriated (e.g. to the capital city or rest of the world) and the amount spent locally. This can be treated either as a measure of local ownership, or as a measure of local procurement by the resource sector, affecting demand for both traded and non-traded goods locally. I am interested in empirical examination of both the first order effects of expansion in resource extraction (via labour markets) and second order effects via local procurement and spending (via the traded and non-traded goods sectors).

The setup assumes a competitive local economy, with a fixed supply of labour. The representative household (identical homothetic preferences) supplies labour and consumes both traded and non-traded goods, hence the two consumption goods are denoted c_N (non-traded service sector), and c_T (the traded good)[4]. The rest of the country also consumes the traded good and hence it may, in equilibrium, be a net importer or net exporter depending on local income and prices.

The production economy comprises three sectors, each producing a simple output using labour as a specific factor entering the Cobb-Douglas production functions, where output is denoted: R in the resource sector, T in the traded sector, and N in the non-traded service sector

[2]For simplicity I combine manufacturing and agriculture as a traded good sector for the model and discussion. In developed economies it is convention to treat agriculture as a traded goods sector, however for the lowest income economies agriculture may also be considered non-traded, particularly subsistence agriculture. I present estimates using a variety of specifications, and decomposing the traded goods sector. The results support the view that agriculture behaves similarly to the manufacturing sector in the data.

[3]Of primary interest is the movement of labour in response to the booming sector, and any profits generated. Since the focus is on the local economy, I treat capital stock as given.

[4]We can characterise the local economy as a small open economy trading with the rest of the country. As such I assume the relatively traded sector is import-competing, and hence takes prices as given, while the relatively non-traded sector can set local prices in equilibrium. In the case of this simple set-up I treat all sectors as either traded or non-traded. The resource sector is a price taker but modelled separately.

respectively. A denotes sector-specific productivity:

$$R = A_R L_R^{\kappa}, T = A_T L_T^{\gamma}, N = A_N L_N^{\beta}$$

I assume $\kappa < \gamma < \beta < 1$ such that the traded sector is less labour intensive in equilibrium than the non-traded service sector, and the resource sector is the least labour intensive; and that all sectors face decreasing returns to labour.

Here, each sector generates positive profits. These profits accrue to capitalists assumed to be located outside the local economy, with the exception of the resource sector, where some portion θ of profits is spent locally.

There are factor markets for labour, whereby the representative household allocates fixed labour supply across the three sectors (L_R, L_N, L_T). I assume full employment in equilibrium and by simplification, labour supply is assumed fixed:

$$L_N^S + L_T^S + L_R^S = \bar{L}$$

I also assume that the non-traded goods market must clear: $c_N = N$.

Prices in the traded sector and exporting resource sector are exogenously given. The price in the non-traded sector is determined in equilibrium. Here, price is determined by market clearing in the non-traded sector. Supply is given by:

$$N = A_N (\frac{w}{\beta p_N})^{\frac{\beta}{\beta - 1}}$$

So, equilibrium wage and price of non-traded goods will be found to satisfy these two market clearing conditions. Consequently the following equation will be satisfied:

$$Y = p_N c_N + p_T c_T = w(L_N + L_T + L_R) + \theta \pi_R$$

where c_N and c_T are domestic consumption quantities, and the θ parameter determines the proportion of resource-sector profits spent locally, versus those repatriated to shareholders in the city or rest of the world. This parameter can be thought of as either the extent of local ownership, or some measure of local procurement sourcing, determining some portion of local spillovers effects. This parameter can be useful for understanding the magnitude of a booming resource sector effect on the local economy.

3.2.2 Equilibrium

For equilibrium, there are two market clearing conditions and two prices: a single market clearing wage and a market clearing price of non-traded goods (the service sector). Since the prices of the resource good and the traded good are exogenously given, the market equilibrium is achieved by adjustment of the two endogenously determined prices.

In the labour market wages are determined by the relative productivity of the three production sectors, labour supply, and the factor intensity of production:

$$(\frac{w}{\beta p_N A_N})^{\frac{1}{\beta-1}} + (\frac{w}{\gamma p_T A_T})^{\frac{1}{\gamma-1}} + (\frac{w}{\kappa p_R A_R})^{\frac{1}{\kappa-1}} = \bar{L} \tag{3.1}$$

Prices in the non-traded goods sector are a function of non-traded goods share of domestic (local) income, the market wage, and the productivity and factor intensity of the non-traded goods sector:

$$p_N = \frac{\alpha Y}{A_N(\frac{w}{\beta p_N})^{\frac{\beta}{\beta-1}}} \tag{3.2}$$

Local household income is determined by wages, labour supply and income retained locally from the resource sector:

$$Y = w\bar{L} + \theta(p_R R - wL_R) \tag{3.3}$$

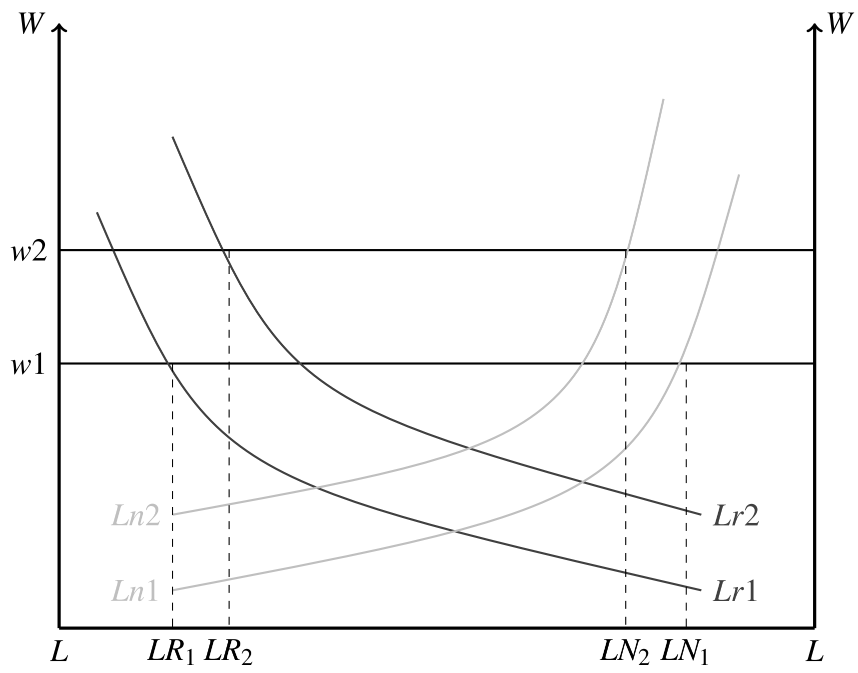

Figure 3.1: Illustrative labour market response (resources *LR* and non-traded *LN* sectors). (The traded sector - not shown - contracts proportionately to the expansion of the expanding sectors, since labour supply is fixed at L)

3.2.3 Effects of a productivity shock in resource sector

We can examine the effects of a resource sector boom which can be modelled as a permanent shock to resource sector productivity. The shock is applied to productivity in the resource sector, increasing from A_R to A'_R, such that $A'_R > A_R$.

The productivity shock creates a resource sector boom, encouraging expansion of the sector in the local economy. This expansion has two effects via two channels - the **resource movement effect** via the labour market, and the **spending effect** via household income and the goods markets.[5]

First, we can examine the direct effects on the resource sector. The overall supply function is increasing in A_R. Further, the labour demand function shifts right with an expansion in A_R. This effect is depicted in Figure 3.1, showing the simulated results of the model. The increased productivity in the resource sector shifts labour demand to the right $(Lr2)$. At the new level of labour demand from the resource sector, wages must rise - thus reducing labour supply in the

[5] I follow the terminology of Corden and Neary (1982)

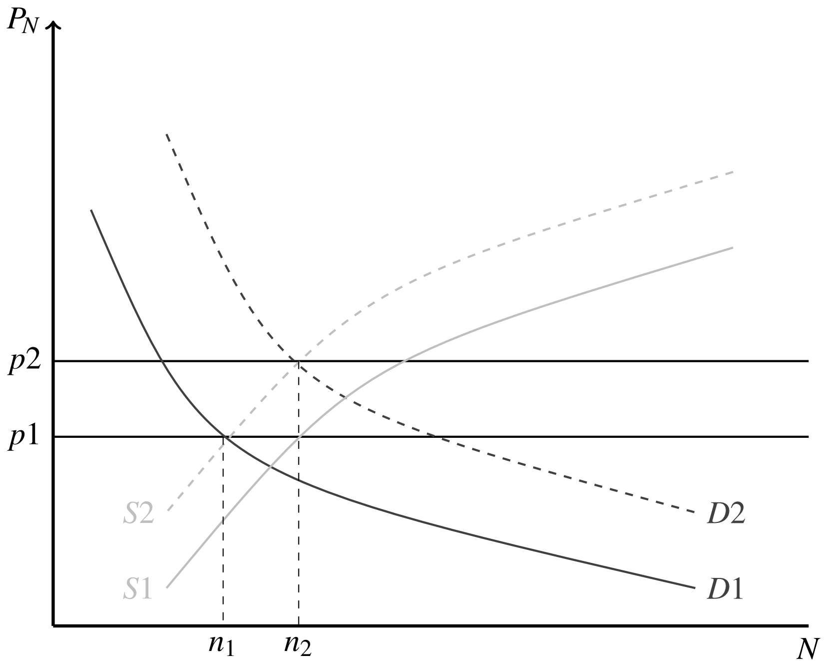

Figure 3.2: Illustrative effect on the market for non-traded goods and services (N)

non-traded and traded goods sectors, *ceteris paribus*.

The labour market effect from the resource sector translates into a bidding up of wages in equilibrium in response to an outward shift the labour demand.

$$\bar{L} = (\frac{w}{\beta p_N A_N})^{\frac{1}{\beta-1}} + (\frac{w}{\gamma p_T A_T})^{\frac{1}{\gamma-1}} + (\frac{w}{\kappa p_R \underset{(\uparrow)}{A_R}})^{\frac{1}{\kappa-1}}$$

This translates into increased income via increased market clearing wages and a rise in resource sector profits, as a function of resource sector productivity A_R. Hence spending on both traded and non-traded goods rises, leading to increased prices in the non-traded goods sector. This feeds back to the labour demand from the non-traded goods sector, shifting outward the labour demand curve. Labour used in the production of non-traded service goods increases from $LN1$ to $LN2$.

Finally, labour demand in the traded sector depends only on the wage level and exogenously determined variables. We know that labour demand in the traded sector is decreasing in wages, hence the model predicts a contraction in labour employed and an overall contraction in local supply. This is depicted by a movement along the labour demand curve shown in Figure 3.1.

So, the equilibrium conditions imply that labour employed in the resource sector and the non-traded sector rises, while labour employed in the traded sector falls. Similarly, output in the resource and non-traded sector rises, while output in the traded sector falls. In equilibrium both the price of non-traded goods and the labour market clearing wage have risen.

3.2.4 Role of local spending spillovers in local Dutch disease effects

So far I have assumed a value $\theta > 0$, thus imposing some positive spillover effect from increased resource sector profits. This is analogous to the spending effect of the Corden and Neary model.

However, for this setting it is of interest to turn this effect off (setting $\theta = 0$). This assumes all resource sector profits leave the local area, thus domestic demand for traded and non-traded goods is affected only via wage effects on household income. The key reason for this difference in result is the assumption that profits accrue to owners (rather than workers), and that the extent to which these profits accrue locally is determined by the parameter θ. In the Corden and Neary macro setting, both capital and labour returns accrue domestically, thus ensuring the domestic economy sees the boom effect. I allow for a relaxation of this assumption to reflect the ownership structures in the regional setting.

Here, I get the same first order effects - demand for labour in resource sector rises with rising productivity. This raises market clearing wages, and thus moves along the labour demand curve for traded and non-traded goods. However, the offsetting effect from increased consumption of non-traded goods is now lower (the spending effect), being driven only by increased local wages, and not by increased local profits to the resource sector. Thus, in equilibrium, the resource sector expands and the traded sector contracts. However, the overall effect on the non-traded sector is ambiguous.

Thus, in the local setting, where consumption depends upon wage income and profits from the resource sector, wages alone may be insufficient to generate a non-traded sector expansion. In contrast, where some profits remain in the local economy, the combination of labour market and spending effects ensures the standard Corden and Neary result.

The magnitude of the multiplier for the non-traded sector will depend on three factors. First, it depends on the extent to which the mining project demands local services and other inputs (given by the local profits parameter θ). This may in turn depend on the scale of the mine, its proximity to the community, and the nature of procurement policy.

Recent literature suggests where mines do not explicitly seek to source locally, this demand for local inputs may be modest. Many mines utilize existing supply chains that might extend to nearby cities or even overseas. Second, the size of the labour market effects depends on the type of jobs created by the mine- skilled and semi-skilled jobs are know to have a larger local multiplier than unskilled jobs (Moretti (2010), Moretti (2011)). Third, there are offsetting general equilibrium effects on wages and prices. In the model these depend upon the elasticity of substitution across consumption and the production functions of the three sectors. However, these general equilibrium effects may also include a spatial consideration, such as the elasticity of local labour supply (determined by distance from nearby cities, local population density and overall labour mobility). If local labour supply is highly elastic, this crowding out effect is more limited and the increase in labour costs small, making the overall multiplier larger. I do not explicitly model or estimate labour supply responses in this paper; however, migration responses constitute an important effect in other resource-rich settings, and estimates are presented in the robustness section.

3.2.5 Research Hypotheses

The opening of a new mine (or mine expansion), within a given proximity of a community is expected to have multiple effects on the composition of the local labour market.

My research hypotheses take the predictions of the stylized general equilibrium model above, applied to the empirical setting of mine opening and mine expansion in a local labour market: this predicts an increase in reported employment in the resource sector (conditional on mining activity being sufficiently proximate). The theoretical priors suggest the expected sign of effect on participation in the mining sector is positive. Therefore one would expect to see an increase in L_R in response to mine opening or expansion (positive coefficient where $Mine = 1$ in equation 3.4 or $\delta NearProd > 0$ in Equation 3.5).

In contrast, I expect decreased employment in traded sectors (agriculture and manufacturing). Where mining projects are present, there is an increase in demand for both mining employees and local non-traded services. Thus I expect a shift away from agricultural employment (and other relatively tradable sectors such as manufacturing) relative to comparison communities. This effect will be mediated by the relative labour market slackness. I would expect to see a decrease in L_T in response to mine opening or expansion (where $Mine = 1$ in equation 3.4 or $\delta NearProd > 0$ in equation 3.5).

Increased employment in non-traded sector supplying inputs to extraction activity (service sectors) relative to declines in other sectors. Here, the positive shock to the mining sector R implies an increase in demand for non-traded inputs. Furthermore increased direct employment will have localised multiplier effects via consumption by mine employees. Conditioning on the degree of labour market slackness and labour mobility, I would expect to observe an increase in overall employment in the non-traded sector (and potentially an increase in the overall *proportion* employed in the non-traded sector). I would expect to see an increase in L_N in response to mine opening or expansion (where $Mine = 1$ in equation 3.4 or $\delta NearProd > 0$ in equation 3.5).

The discussed sectoral employment effects are likely to be effected by the relative proximity of extraction activity. Effects of proximity and scale on labour market adjustment are expected to dissipate across space.[6] The magnitude of resource boom effect of sector shifts are expected to be lower the greater the community distance from the mine. Here I expect to observe a negative distance coefficient for the growing sectors (Resource sector and Non traded sector). Likewise I would expect a positive coefficient on the measure of distance from mine, where the dependent variable is employment in the Traded goods. I also test this relationship using the

[6]I do not explicitly model this process but can test for it in the data.

specification detailed in Equation 3.5.

The ideal test of parameter θ would be to split the sample of mines based on local ownership and local profit sharing. Local ownership in this setting would imply within the distance radius of the calibrated distance footprint of the mine. Instead I only observe foreign versus domestic ownership of mine. By testing for heterogeneous labour market responses depending on ownership I can examine whether there is evidence for a more pronounced local impact predicted in cases where $\theta > 0$.

3.3 Data

The Indonesia dataset is constructed using information about mining projects (drawn from both the Raw Materials Dataset and Wood Mackenzie's coal mine database (Wood Mackenzie, 2011)) combined with household and community panel survey data from the RAND Corporation. I supplement this with data on light intensity measured by nighttime satellite imagery (NOAA 2013).

The mine project data includes a location, status as well as a range of production-related variables. The survey data is drawn from the large-scale household and community panel survey for Indonesia (the Indonesia Family Living Survey) conducted in 1993, 1997, 2000 and 2007. The initial wave is representative of 83 percent of the Indonesian population across 13 of Indonesia's 27 provinces.[7] Indonesia is one of the most resource-rich economies of southeast Asia and has one of the largest mining sectors in the world, accounting for over 10% of GDP. The sample contains data on over 300 individual resource extraction projects comprising industrial mines and accounting for 95% of total production in Indonesia.

The dataset incorporates geo-coded data allowing me to spatially link individual extraction projects, nighttime lights and community clusters in the survey data. The distance from the community clusters to the location of extraction projects allows me to estimate the local effect of extraction activity on surrounding communities. Further, the variation in proximity

[7] 27 Provinces in 1999. Since decentralization Indonesia has seen a proliferation of new provinces and administrative districts, taking the total in 2009 to 33.

to extraction activity allows me to identify heterogeneous spatial effects of extraction on local populations.

The longitudinal nature of the dataset makes it possible to examine the dynamics of individual and household behavior in response to extraction shocks, including labour market decisions.

3.3.1 Data Description

The mining data has coverage across all of Indonesia, indicating concentrations of mining activity in the resource rich provinces of Java, Sumatra and South Kalimantan. These coincide with the location of many of the survey communities. Figure 3.3 illustrates the location of the mines and surveyed communities.

The dataset contains information on over three hundred individual mining projects (operating, closed and in development). This composite dataset, in addition to providing detailed location data on mining projects, also has coverage of historical production data, changes in ownership, investment levels, and other key metrics. Of particular interest to this work is the status of the project (open or closed), mine ownership, the changes in production over time, and the proximity of the mines to the surveyed communities.

The Indonesian Family Life Survey (IFLS) is an on-going longitudinal household survey in Indonesia. The sample is representative of about 83 percent of the Indonesian population and contains over 30,000 original individuals living in 13 of the 27 provinces in the country. The first wave of the IFLS (IFLS1) was conducted in 1993/94 by RAND in collaboration with Lembaga Demografi, University of Indonesia. IFLS2 and IFLS2+ were conducted in 1997 and 1998, respectively, by the same collaboration with UCLA. IFLS3, which was fielded in 2000 and covered the full sample, was conducted by RAND in collaboration with the Population Research center, University of Gadjah Mada. The fourth wave of the IFLS (IFLS4), fielded in 2007/2008 covering the full sample, was conducted by RAND, the Center for Population and Policy Studies (CPPS) of the University of Gadjah Mada and Survey METRE.

There were 7,730 households in the original target sample (IFLS 1 in 1993). Of those, interviews were completed with 7,224 households, and detailed individual-level data were collected from over 22,000 people. IFLS2 sought to re-interview all 7,224 households with 6% attrition. In addition to the approximate 6,750 origin households, over 850 split off households were interviewed in IFLS2 (1997). Thus, IFLS contains in total 8,116 households. In IFLS3 the re-contact rate was 95.3% of IFLS1 households. Of the baseline respondents who were re-interviewed in the fourth wave in 2007, over one-third had moved away from the community in which they were interviewed at baseline. Nearly 90 percent of IFLS1 households are complete panel households in that they were interviewed in all waves. These re-contact rates are as high as or higher than most longitudinal surveys in the United States and Europe.

In total, the IFLS covers 311 communities and the survey questionnaire, although primarily concerned with household health outcomes includes sections on consumption, employment, education, farming and business. The survey has coverage of the main islands comprising Indonesia and has particularly good coverage in key mining regions including the Iron Sands near Yogakarta, the coal fields of South Kalimantan. Out of the total sample approximately 52.4% of the households were located in rural areas.

In order to examine the forward linkage effects of extraction projects through labour market channels I use surveyed employment history for all adults (18 or over in IFLS3, 2000) in the sample. I construct both an individual and community level panel dataset using reported activity in 1997, 2000 and 2007. The survey contains information on individual characteristics, employment history, type of activity, sector of employment, salary as well as enterprise activity and unearned income.[8]

Table 3.1 summarizes the coverage of the dataset in terms of surveyed individuals, communities and years. In addition it includes a summary of migrating individuals.

Table 3.1: Summary of panel observations

Year	**Community clusters**	**Total sample indvs.**	**Non-migrating indvs.**
1997	303	38,250	34,457
2000	303	48,475	37,547
2007	303	59,013	40,208

[8]The first round, 1993, is excluded from the analysis due to lack of data on employment sectors.

Table 3.2: Sector employment shares, by community, by year

	Variable	Mean	Std. Dev.
Year=1997	Ag Sector	0.289	0.301
	Mining Sector	0.007	0.034
	Manu Sector	0.164	0.141
	Serv Sector	0.54	0.263
	Not in work/unknown	0.001	0.007
	N	303	
	Variable	**Mean**	**Std. Dev.**
Year=2000	Ag Sector	0.299	0.298
	Mining Sector	0.006	0.03
	Manu Sector	0.14	0.118
	Serv Sector	0.556	0.261
	Not in work/unknown	0.002	0.009
	N	303	
	Variable	**Mean**	**Std. Dev.**
Year=2007	Ag Sector	0.26	0.273
	Mining Sector	0.006	0.026
	Manu Sector	0.137	0.109
	Serv Sector	0.597	0.25
	Not in work/unknown	0.005	0.002
	N	303	

Table 3.2 provides an overview of sectoral composition by year for the surveyed communities. The sample sees a general upward trend in service sector employment, with modest declines in agriculture and manufacturing. Mining employment remains roughly constant as a proportion of reported community employment across the entire sample.

The IFLS household data can be summarized for those communities close to operating mining projects and those not. As illustrated in Table 3.3 and 3.4 below, there is a degree of variation between those districts with some mining activity and those without. As we would expect, I observe higher reported employment in the mining sector relative to other sectors. The proportion employed in agriculture is somewhat higher, whereas employment in manufacturing and the service sector is lower. However, communities close to mines are less urban and located further from provincial capitals than comparison communities on average.

Refer to Figure 3.3 for the complete mapping of the locations of the mines and surveyed communities. Figure 3.4 gives a higher resolution snapshot of a typical mine locality. The concentric circles represent the distance bands where I empirically test for economic effects over distance. Of particular importance to the identification strategy is the proximity of the communities to mines, and the high degree of spatial and temporal variation amongst the survey communities (also summarized in Figure 3.3). This can be thought of in terms of variation in the intensity of treatment.

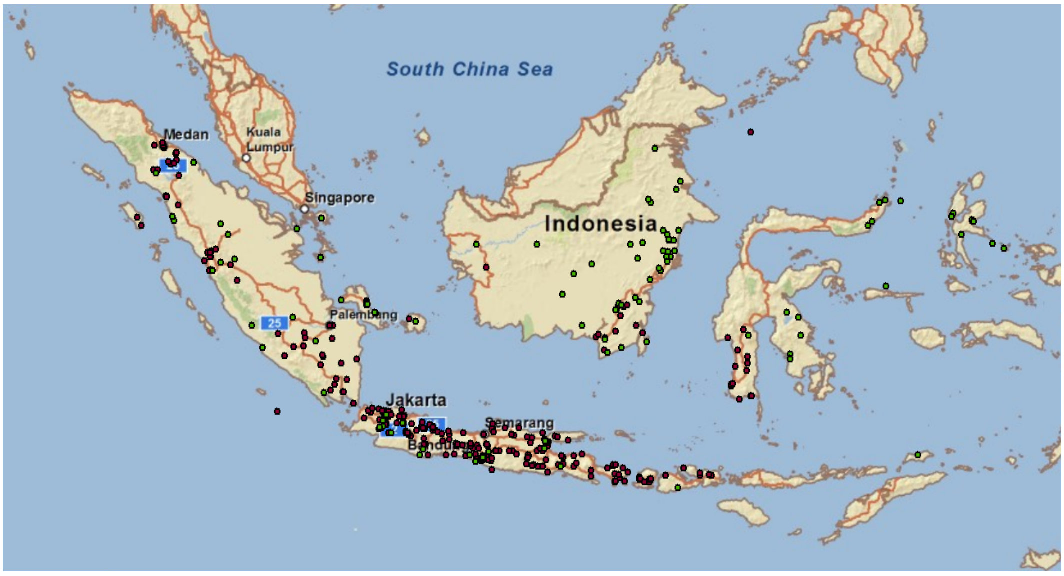

Figure 3.3: Map of mines and communities (green/grey=mines, red/dark=survey communities)

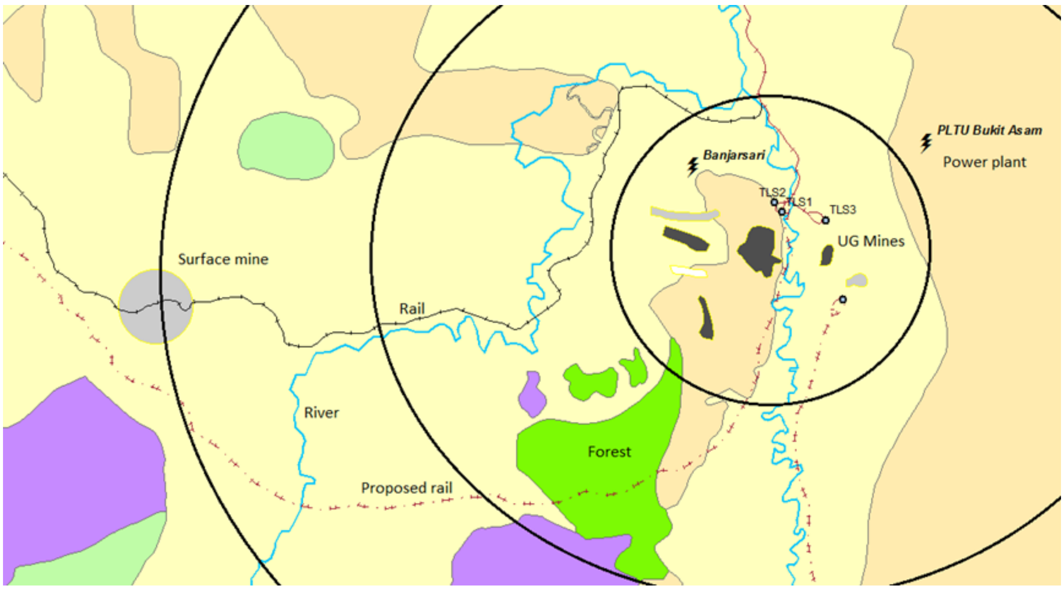

Figure 3.4: Mapping layers of a illustrative coal production area (with distance bands, 5km, 10km and 15km shown in black)

Table 3.3: Descriptive Statistics for non-mining comparison communities. IFLS waves 2, 3 and 4, years 1997, 2000 and 2007

Mine dummy (open and within 20km of community) = 0

Variable	Mean	Std. Dev.	N
Nearby mining production (MT ore)	0	0	825
Age	29.872	3.427	825
% completed highschool	0.62	0.8	825
% graduated	0.05	0.09	825
% female	0.513	0.043	825
% in Agricultural Employment	0.278	0.29	825
% in Mining (and Quarrying) Employment	0.005	0.022	825
% in Manufacturing Sector	0.148	0.125	825
% in Service Sector	0.569	0.259	825
Urban Dummy	0.576	0.477	825
% hhs with elec	0.92	0.197	825
Time to district cap (hrs)	0.824	2.228	825
Time to provincial cap (hrs)	2.375	2.856	825

Table 3.4: Descriptive Statistics for mining communities. IFLS waves 2, 3 and 4, years 1997, 2000 and 2007

Mine dummy (open and within 20km of community) = 1

Variable	Mean	Std. Dev.	N
Nearby mining production (MT ore)	170.917	352.6	83
Age	30.497	3.96	83
% completed highschool	0.633	0.72	83
% graduated	0.061	0.1	83
% female	0.524	0.044	83
% in Agricultural Employment	0.33	0.299	83
% in Mining (and Quarrying) Employment	0.024	0.071	83
% in Manufacturing Sector	0.132	0.105	83
% in Service sector	0.514	0.251	83
Urban Dummy	0.542	0.48	83
% hhs with elec	0.921	0.189	83
Time to district cap (hrs)	0.763	1.181	83
Time to provincial cap (hrs)	2.815	2.952	83
Distance to closest mine (km)	16.571	5.591	83

3.3.2 Construction of treatment variables

The mining dataset contains two key sources of variation. The first is the mine output, determined by the operation of a mine, measured in output volume (in MT of ore output). The second source of variation is the distance measure, defined by the distance between the community in question and its nearest mine (in km).

Table 3.5 presents descriptive statistics of the mine production volumes: the number of observations, the mean, the standard deviation. It shows the degree of variation across mines, and the changing production patterns over time. Broadly speaking we see increases in production in around 60% of the sample of mines with the remaining 40% either remaining constant or seeing a contraction in production. Volume for a single mine ranges from exploratory production of less that 1MT per annum, up to 250MT for the largest mines proximate to the surveyed communities.

Table 3.5: Mining Projects Descriptive Statistics

Mines		
Variable	**Mean**	**Std Dev**
Mine production 1997 (by volume)	7.052	43.165
Mine production 2000 (by volume)	9.069	54.457
Mine production 2007 (by volume)	14.805	63.765

I characterize mine treatment to a community in several different ways; exploiting a combination of both the spatial and temporal variation in the panel. The direction of results are robust to different configurations in the analysis and discussed in section 4.6.

Summary of ’treatment’ variables:

- The ‘open mine’ dummy - A binary dummy taking the value of 1 for at least one mine within a given distance of the community, and the mine is open in the year of survey. Taking the value of 0 if no mine is present or if the mine is closed in the year of survey within the given radius. I calibrate the distance radius using nighttime lights data. I also present alternative distance bands for the treatment in section 4.6. Figure 3.6 illustrates the variation in open mines in the sample.

- Mine production variable - A continuous measure of mine output. It is the sum of production for the 5 closest mines that are also with a given radius of the community. The variable takes a value > 0 where at least one mine lies within a given distance of the community, and the mine is operational in the year of survey. Otherwise the variable takes the value of 0 if no mine is present or if the mine is non-operational in the year of survey.

- $Distance_{jt}$ A continuous measure of distance from the community to the nearest operational mine. Measured in km.

- Spatial bands. These indicator variables take a value of 1 if at least one mine lies within the distance radius band, for example for 15km band the closest mine must lie between 10km and 15km from the community. Likewise the 10km band*production with 10km takes the production volume of the five closest mines to the community, conditional on those mines falling within the relevant distance term.

In the simplest specification $Mine_{jt}$ characterizes the group that is exposed to the treatment (an open mine) in the relevant period. It treats mines of different sizes equally. The control group, where $Mine_{jt} = 0$, is not exposed to the treatment in the given period. By adding a full set of time dummies to the question, this assumes the policy change has the same effect in every year. Since the treatment is applied at the community level rather than the individual level, potentially invalidating the assumption of zero conditional correlation across observations, Bertrand et al. (2004) recommends clustering standard errors at the level of treatment.

3.4 Empirical Methodology

The identification strategy exploits two sources of variation in the data. First, I use the opening or expansion of mines, relative to one another (for estimation of mining footprint) and within the locale of the surveyed communities. Conditional on the mine opening and being proximate to the surveyed community, I can use the change in activity over time (controlling for local fixed effects) to identify the link between mine activity and economic footprint in addition to local labour market composition. Where mine expansion draws resources from other sectors (resource movement effect) and also increases local real income (via wages and profits - the spending effect), I can test the predictions derived from the stylized model.

Second, I observe the sample of mines and their varying spatial distribution of economic footprint and variation in proximity to the survey communities. This relationship allows me to estimate correlations between mine proximity, economic activity, and labour market composition. For example, by spatially linking the household and community level data with the mining projects, I can examine employment composition associated with a resource sector boom using variation in mine proximity. Here I can compare the magnitude of labour market composition effects, with increasing and decreasing distance to mines. The distance measure is calculated using orthodromic distance formulae for point-to-point calculations.[9]

3.4.1 Identification Strategy

The estimation strategy follows from the literature examining the localised labour market dynamics in response to external shocks (Carrington (1996), Notowidigdo (2010) and Moretti (2010)). The placement of mines is largely a private decision, based on the nature of the underlying geology of a region and the extent to which that geology is known. However, companies must also acquire the necessary concessions from governing authorities. While these features mentioned do not directly relate to the economic conditions in a given community, the private placement decision may be influenced by local economic and political characteristics. In particular mines may be more easily placed in areas of low population density or outside built-up areas. This may be due to licensing restrictions or challenges associated with the potential social and environmental effects of industrial mining, as well as the cost of land. In contrast, mine owners may favour regions with ready availability of labour, therefore seeking populated areas, or those with good infrastructure. Transportation infrastructure is also likely to affect the cost of getting minerals to market. The direction of this placement selection is ambiguous,

[9] Stata command *nearstat*, Jeanty (2012).

but without addressing potential influencing factors, it may bias any cross-sectional estimation strategy.

The authority to grant mining concessions resided in the centre prior to 2001 (i.e. Federal authority) and has been nominally decentralized to the provincial level post-2001.[10] This administrative discretion is likely to be a driver of the timing and scale of mining in the dataset as it determines the underlying royalty and tax structure faced by the private mining companies when choosing whether or not to undertake large capital investments. The underlying royalty and tax structure (somewhat time-invariant per concession) will in turn determine the rate of extraction and relative profitability of the mine, which will effect operations on a year-by-year basis. These factors are likely to affect the dynamics of any labour market spillovers we might expect to observe, particularly with regards their temporal characteristics. I deal with these concerns via the use of province-level fixed effects.

The location of mines will also be affected by administrative heterogeneity and discretion. In particular, where concession granting resides with provincial authorities we may expect province-level factors to influence the rate at which concessions may be granted. If I do not account for these province-level (time invariant) effects, the results would suffer from bias introduced by omission of key province-level unobservables (such as variation in provincial government support for mining).

In addition, my priors suggest mine placement and scale of activity, although driven largely by geology, contractual terms, and government discretion, may also be influenced to a lesser extent by the economic characteristics of local communities. As such we might expect some degree of endogenous placement of mines driven by community variation. I argue however, that if present, any such endogeneity is likely to be driven by the underlying time-invariant characteristics of a locale, rather than its economic trajectory over time. If we are able to assume placement and activity is based only on time-invariant characteristics, the fixed effects estimator also allows me to control for these community-level effects that would otherwise bias the estimated coefficients.

[10] In practice the centre retained significant powers. The new 2009 mine law has sought to amend this.

Finally, given the longitudinal nature of the data I include year dummies to account for any common time effects across the surveyed communities. In particular we want to isolate common shocks such as the East Asian Financial Crisis of 1997 or changes in world mineral prices.

The panel can be characterized as containing two groups of communities or individuals, those within a given proximity to a mine project (opening, closing or expanding) during the survey period and those not. We can think of these two groups as the treated and the control. The simplest treatment variable is characterized by a dummy $Mine_{jt}$, which takes a value of either 0 or 1 depending on whether an operational mine is present within a given distance band of the community (in a given year). Mines open and close in the sample period.

Consider the reduced form specification for the binary treatment variable:

$$L_{jt}^{X} = \rho Mine_{jt} + B_{jt}'\beta + P_j'\theta + T_t' + (v_j + \omega_{jt}). \tag{3.4}$$

Where $Mine_{jt} = (0,1)$ for comparison and treated communities j in a given year t. B_{jt} denotes vector of community level characteristics. I determine whether communities are treated by a mine using the estimated footprint size, discussed in section 3.4.3. I can control for community-level and province-level time-invariant effects; $P_j'\theta$ indicates the inclusion of community fixed effects, I also include a vector of Province characteristics. ε_{ijt} gives the composite idiosyncratic errors. They can be decomposed as $\varepsilon_{jt} = v_j + \omega_{jt}$. The composite error therefore contains community time invariant effects in addition to a common ω_{jt} which is assumed to be i.i.d. over the sample. L_{jt}^{X} is the dependent variable of interest- we shall examine different choices for L_{jt}^{X} in section 3.4.2.

Identification depends upon the extent to which mine placement is exogenous after controlling for observables that might otherwise influence mine proximity to a community of interest as captured by the v_j error. As discussed, I expect province-level characteristics to be of importance for the dataset - due to the process of awarding public concessions at both the province (post-2001) and Federal-level (pre-2001). Furthermore, a range of provincial characteristics such as ports and inter-district transportation infrastructure will likely influence mine placement decisions and the resultant scale of operations. These unobserved characteristics will likely invalidate any assumptions of strict exogeneity. Instead I must apply province-level fixed effects to account for these unobserved effects.

I control for time effects and province-level fixed effects. Further I also include observable community characteristics (v_j) such as an urban dummy - if the community is located in a densely populated area, a dummy denoting the electrification status of a community, and a measure of distance to the provincial capital. These account for potential economic drivers of mine placement at the community-level that might otherwise bias the estimator. Where $E(Mine_{jt} \times \varepsilon_{jt}) = 0$, that is, the level of treatment is uncorrelated with the group and individual idiosyncratic errors, then OLS applied to Equation 3.4 is consistent. A further consideration is the presence of treatment at the community level, where inference is conducted at the individual level. Here I apply the clustering correction to standard errors at the community level.

The alternative specification, show in Equation 3.5, takes into account these considerations and additional introduces the continuous measure of 'treatment'. The $NearProd_{jt}$ variable contains information on the size (in terms of MT ore output) of the nearest mines to a community. This also allows a representation of the variation over time- all the operating mines either increase or decrease reported production volumes at some point during the surveyed period. Here the (changing) levels of output are a useful measure of variation in 'treatment' with time. Further, I also include a continuous distance measure from the community centre to the nearest mine. This captures a linear form of the spatial variation in the dataset, whereby I expect the magnitude of labour market spillovers to be mediated by distance.

Let $NearProd_{jt}$ denote the continuous treatment variable, which indicates the level of production for a sum of the five closest mines (within a given kilometre distance band) to a given community for a given year.

$$L^X_{jt} = \rho NearProd_{jt} + \delta Dist_j + B'_{jt}\beta + P'_j\theta + \varepsilon_{jt}. \tag{3.5}$$

Where $NearProd_{jt}$ gives a continuous measure of mine production volume (MT) for treated communities j in a given year t. $Dist_j$ gives a continuous of distance from the nearest mine to the community (in km) (this drops out when I include community level fixed effects since mine and community location are fixed).

An alternative approach to mine status and production level could be to use measures of reserves rather than production or mine status. However this presents several challenges. First the data on deposits or reserves is time-invariant. Where time-varying reserve measures are available, the reason they time vary is related to intensity of activity at the mine site, and thus another approximation of activity or production. Furthermore deposit or reserves are subject to the same concerns regarding the effect of external factors such as the issuing of exploration licenses or the characteristics of the region.

3.4.2 Estimating labour market dynamics

Thus far I have given little discussion to the dependent variable of interest L_{jt}^{X}. Drawing on the conceptual framework of local employment spillovers (Moretti, 2010) I am interested in testing the effect of the external shock on local labour market composition.

The dependent variable is the sector of employment, defined at both the individual and community levels. This follows similar approaches in the literature such as Deichmann et al. (2005). For the empirical analysis I define the dependent variable L_{jt}^{X} in two ways. First as employment in a single sector - taking a value of 1 if the individual is employed in that sector, and 0 otherwise. Second, I define the dependent variable and the employment status across the range of three sectors, plus non-employment. Specifically the dependent variable takes the value of 1 if the individual is employed in the resource sector (mining), 2 if employed in the traded sector (agriculture and manufacturing) and 3 in the non-traded services sector. The variable takes a value of 0 otherwise. Where the dependent variable includes binary responses I can use a standard linear probability estimation technique for analysis.

I also estimate the equivalent sector-by-sector specification at the community-cluster level. Here the sector values take a value between 0 and 1 indicating the proportion of people employed in each respective sector for the community and year in question. Employment in each respective sector is denoted L_{jt}^{R} for the Resource sector R, L_{jt}^{T} the traded sector T, and L_{jt}^{N} for the non-traded service sector N.

Large mines are relatively capital intensive and may not draw much of their workforce from the immediate vicinity. Instead the beneficiaries are typically urban dwellers, either shareholders in the mining parent company, employees in the mine, or beneficiaries of downstream industry and subsequent consumption goods. On the other hand, the potential for positive spillover effects is great- particularly in areas otherwise dependant upon agriculture. Large-scale resource extraction can increase the demand for local labour, in addition to increases in demand for locally-sourced goods and services. Where resource extraction is accompanied by capital investments, nearby communities may benefit from improved infrastructure such as road connectivity and the availability of electricity.

The model I use to test the effect on proportions of the community employed in the given sector is:

$$L_{ijt}^{X} = \alpha + \rho NearProd_{jt} + \delta Dist_{j} + B_{jt}'\beta + P_{j}'\theta + \varepsilon_{jt}. \tag{3.6}$$

for individual $i = 1,...,N$, community $j = 1,...,J$, and time $t = 1,2,3$ where $1 = 1997$, $2 = 2000$ and $3 = 2007$, and $X = (R,T,N)$. L_{ijt}^{X} is a binary dependent variable indicating employment in a given sector. $Dist_{jt}$ gives the measure of distance from the nearest mine to the community where the individual is located. This is the treatment variable, indicating the proximity effect of extraction projects on labour market outcomes for individuals. The second, $NearProd_{jt}$ indicates the level of production for up to five mines within a given distance (in kilometres) of the community. The ρ parameter captures the effects of local resource extraction activity on labour market composition. B_{jt} denotes vectors of community level characteristics and $P_{j}'\theta$ indicates the inclusion of Province fixed effects.

I estimate the above specifications at both the individual and community levels. At the individual level, sector of employment takes a value of zero or one indicating current job sector.

For the community specifications this measure is aggregated at the community level and takes a proportionate measure of sector composition for surveyed individuals in the community.

3.4.3 Estimating Mine Footprint

I calibrate the distance cutoff for treatment via estimating the economic footprint of mines in the sample. The model makes predictions based on the assumption that the mine is located in the locale of the surveyed labour market, and is therefore subjected to the treatment effect. However, the dataset contains survey villages and mines at varying degrees of proximity and sizes. The use of a first-difference estimator allows estimation of the effect of mine opening and mine closure over time. This accounts for time-invariant unobservables.

$$L_{jt}^{Distance} = \alpha + \gamma MineDummy_j + B'_{jt}\beta + P'_j\theta + \varepsilon_{jt}. \tag{3.7}$$

I estimate a specification using various distance cut-offs in the dependent variable - the sum of lights in a given distance radius around the mines (eq 3.7). I run specification tests on a variety of distance bands to determine the likely footprint of mines in the sample. This allows me to calibrate the treatment areas for the labour market analysis. I can then introduce distance-based treatment into the main specifications. Furthermore, estimating the specification this way offers the potential to capture non-linear effects of distance, in contrast to the continuous distance variable in the labour market estimates. Such distance band estimates are common in the literature of rural specialisation and labour market composition.

3.5 Results

This section provides the results for economic footprint and estimated effects on the local labour market. I first estimate and calibrate the size of an estimated economic footprint of mines in Indonesia- this process then allows me to analyse the impact of mines on the structure of the labour market. I am interested in both the magnitude and spatial scale of these effects. I present results from the panel data estimates at both the individual and community level.

3.5.1 Estimating size and intensity of the mining footprint

The dataset contains 338 individual projects, including operating mines, mines under construction, planned mines and exploration plots. Of these over 150 were operational mines at some point during the period of interest. Combining the location of mining projects with spatially precise estimates of economic activity allows us to examine the specific footprint of resource projects.

I am able to exploit the exogenous variation in mine opening and closure timing, as well as changes in production levels to examine the economic footprint effects. In particular changes in mine operations, driven by geological constraints, in addition to centralized concession licensing isolate the mine activity variation from being driven by local economic conditions. The isolated enclave nature of mining activities in Indonesia imply that expansion is driven by external factors, but also that expansion can have profound effect on local economic conditions.

An extreme example of this is the largest mine in the dataset. Batu Hijau, since its construction has seen a transformation of its surrounding region. Previously a heavily forested, sparsely populated coastal province, the mine has brought with it the development of two ports, several large processing factories and a range of mine-related economic activities.

The Batu Hijau mine is an open-pit copper and gold mine. It is operated by Newmont Mining Corporation and is the one of the largest mines in the world. The mine is the result of a ten-year exploration and construction program, with the site being acquired in 1990 and discovery of the copper deposits announced in 1999. Production started in 2000. By 2005 the mine was producing 325,500 tonnes of copper and 719,000 ounces of gold. It is set to continue production through to 2025. Batu Hijau contributes an estimated 1.5% of total global production by value.

From the unique dataset I am able to construct estimates of the variation over space and time as new mines open and expand. This allow estimate of the direct footprint effects, as well as the role of nearby communities, ports, infrastructure and other economic linkages. In the case of Batu Hijau, the activity located to the South West of the mine is a large ore processing facility. Further to the West and North are two ports, which have seen dramatic expansion since mine construction began, shown in Figure 3.7.

Table 3.6: Estimation of mine footprint - light number/km^2

	(1) Light 3km	(2) Light 5km	(3) Light 10km	(4) Light 15km	(5) Light 20km	(6) Light 25km
Mine dummy	0.412*	0.482***	0.267**	0.193**	0.093	0.019
	(0.223)	(0.173)	(0.111)	(0.086)	(0.074)	(0.069)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Panel	Fixed effects	Fixed effects	Fixed effects	Fixed effects	Fixed effects	Fixed effects
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Obs	6327	6327	6327	6327	6327	6327
F-Test	8.01e-27	7.03e-28	3.33e-29	1.06e-30	1.88e-32	1.02e-35
Adj. R-squared	0.185	0.167	0.198	0.221	0.230	0.232

Controls include year dummies and province fixed effects. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Estimation of mine footprint- change in light number/km^2

	(1) Δ Light 3km	(2) Δ Light 5km	(3) Δ Light 10km	(4) Δ Light 15km	(5) Δ Light 20km	(6) Δ Light 25km
Δ Mine open dum	0.221*	0.214**	0.123**	0.079*	0.011	-0.020
	(0.114)	(0.100)	(0.058)	(0.047)	(0.044)	(0.042)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Panel	First Diff.	First Diff.	First Diff.	First Diff.	First Diff.	First Diff.
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5994	5994	5994	5994	5994	5994
F-Test	0.000	0.000	0.000	0.000	0.000	0.000

Controls include year dummies. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

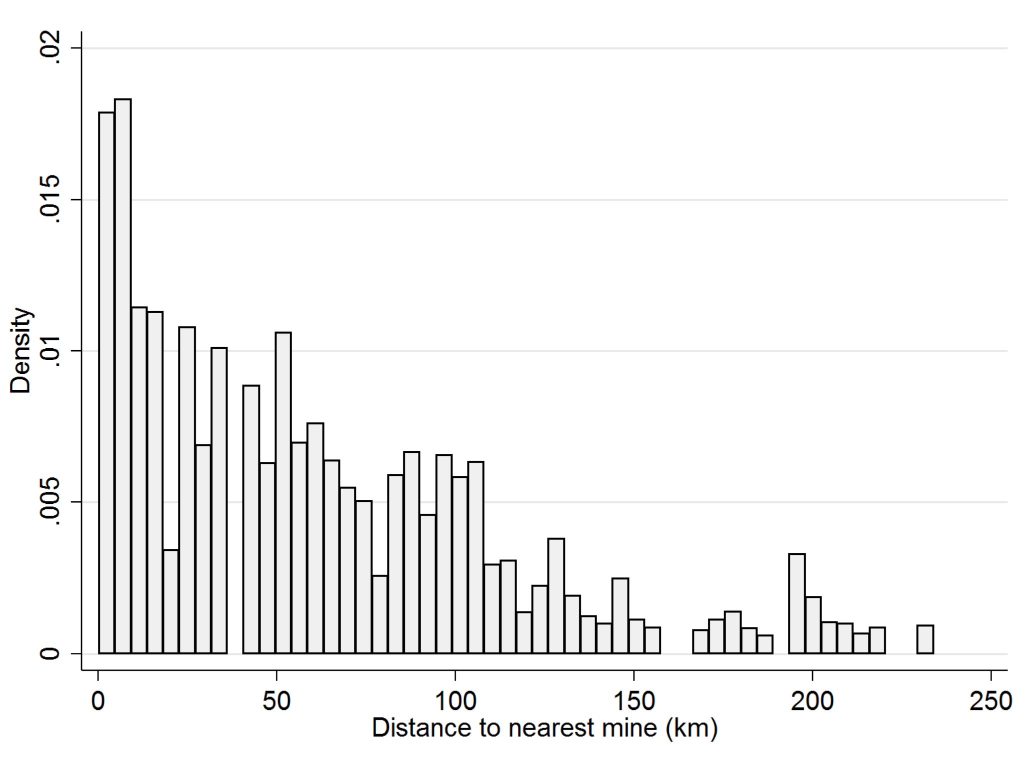

Figure 3.5: Plot of the distribution of observations in terms of distance (in km) from the nearest open mine

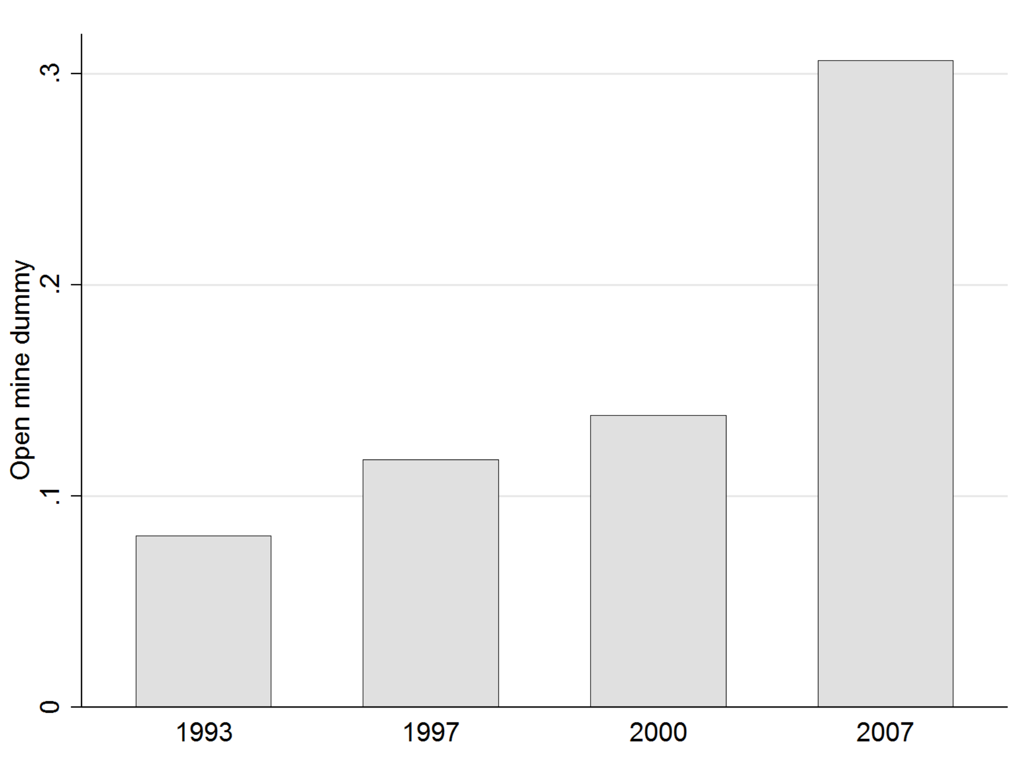

Figure 3.6: Plot of the proportion of mines open within the 25km distance radius of the community, by year

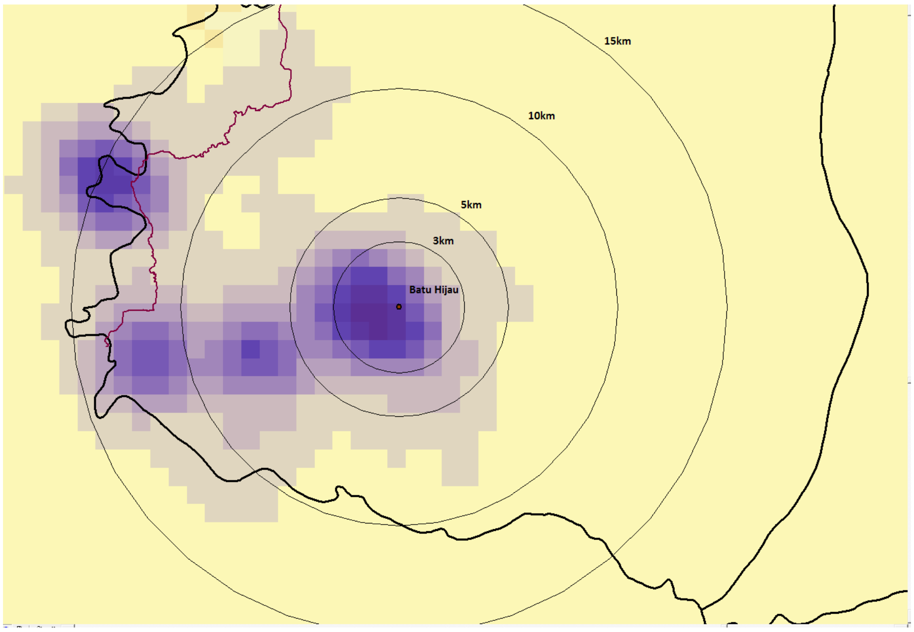

Figure 3.7: Batu Hijau Copper and Gold mine. Estimating the economic footprint of mining activity: total change 1995-2010. Distance bands shown in black (in km). The two left-most areas of effect are two ports, while the right-most is the mine itself. The middle concentration of light represents a mining-related settlement.

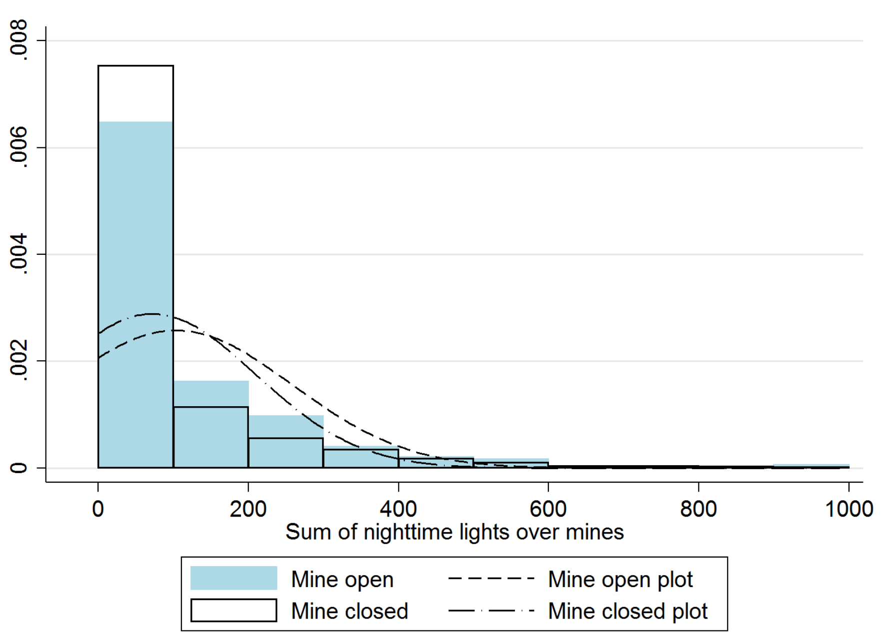

Figure 3.8: Plot of density of nighttime lights for open and closed mines. Within 15km band. Y-axis shows the density, X-axis the digital light value within the distance band.

I first examine the economic footprint of industrial mines using measured nighttime night intensity as a proxy for economic activity. The nighttime lights data allows me to conduct a range of specification tests to calibrate the distance cutoffs around mine sites over time. From this I can evaluate the size and magnitude of the economic footprint in the lights data. Table 3.6 presents the main results, which shows the effect of mine activity on nighttime light intensity within a given distance band. This effect is strong and positive within 3km, 5km, 10km and 15km.[11] Complementing this finding, Table 3.7 presents the same specification in first-differences. This now measures the change in lights arising from a contemporaneous change in mine status- the effect of mine opening. Again columns 1 to 4 provide evidence for a strong, though decreasing positive effect measured in light intensity around the mine site up to 15km.

Beyond this level the effect in both estimations appears to decay and is no longer significantly different from zero. While sign and significance are important clues, magnitude estimates should be treated cautiously for reasons discussed below. First, we cannot separate the magnitude effect of light emissions from the mine site and the nearby communities; instead must rely on the spatial band estimates to confirm whether this effect extends beyond the immediate vicinity of the mine (e.g. 3km band). Second, there are challenges in converting digital light measures into a reasonable approximation for local GDP. What does the light represent? Increased access to electricity due to a nearby mine providing anchor load? Or could it represent increased incomes and therefore expenditure on light consumption? Does it show a flourishing of public good provision in the form of street lights, or increased shops and other private activities? These channels are not identified in the paper and thus we must remain cautious about what economic interpretation we apply to these findings. Further, the problem of converting light measures into an approximation for economic activity is especially severe at the micro-level where low light radiance makes bottom coding a challenge (Chen and Nordhaus, 2011).

With these caveats in mind, we can be reassured however that this does constitute basic

[11] Recall the distance bands are concentric circles around the mine site, and divided by areas. Therefore the 5km variable is the estimated effect in the band extending 3km-5km in radius around the mine. To estimate the total effect, the bands would need to be summed.

evidence for a positive economic effect beyond the mine perimeter. And a clear estimate for how far beyond the mine this extends. This contrasts arguments dating back to Hirschman (1958) of mining following an enclave development model, and more recent findings by Aragon and Rud (2012) that some groups close to mine sites have been excluded from the development gains and suffer from negative pollution externalities.

This analysis serves us with a guide to the likely economic footprint observed in the household survey data. However, because the sample size is much smaller (only 301 survey clusters, only 83 of which have operational mines within 20km), it is likely the precision of the labour market estimates may be lower than the light-based measures.

3.5.2 Mine closure: estimating the bust versus the boom

In contrast, some mines have contracted or even ceased operations during the sample period. The Kelian goldmine, formerly operated by Rio Tinto, was closed in 2005 after 13 years of operation. It was ranked 875th largest of world mines by value of production, employing around 1500 workers at its peak. The closure of mines allows me to estimate a corresponding bust effects. To what extent do the positive impacts of mine opening or expansion persist beyond the mine lifetime? And how are these effects mediated by infrastructure investment, proximity to population or other factors? In the case of Kelian goldmine, the dataset allows estimation of the absence of legacy impacts, suggesting a negative rebound effect associated with the withdrawal of capital, labour and production activity- see Figure 3.9 and Tables 3.8 and 3.9.

Consistent with the estimates for the mine-opening effect, Table 3.8 shows a modest distance radius where effects are measured and significant. This extends only 10km in the data. The estimate magnitudes are comparable with the mine opening effect within this radius. The point estimates in Table 3.8 are slightly larger close to mine sites ($\hat{\beta} = -0.501$) for the 3km band for example, though drop off faster and are no longer distinguishable from zero at 15km. Results are not robust to first difference estimation however and thus must be treated cautiously. This may indicate a somewhat more persistent positive effect, before the bust sets it. Recall Table 3.9 estimates contemporaneous effects on light measures of mine closure- in that year. Thus if the bust effect propagates more slowly that the boom effect, our preferred estimates may be presented in Table 3.8.

In contrast to Black et al. (2005) who examine the coal boom and bust in north-eastern United States, I do not find strong evidence for an asymmetric magnitude of effect. Instead the nighttime lights based estimates provide tentative estimates for a roughly symmetric boom to bust. The point estimates for the *dimming* effect of mine closure should be interpreted with extreme caution however. They likely constitute, at best, a lower bound estimate for the size of the negative effect. Notably measurement of light intensity takes only non-negative values, ranging from zero to sixty three. Therefore any estimates of a dimming effect are vulnerable to bottom coding, or truncation at zero. This is a particular concern in remote, dimly lit regions of the world, such as rural Indonesia. This is a corollary to the top-coding challenge discussed in reference to GDP estimates from lights data in Chen and Nordhaus (2011).

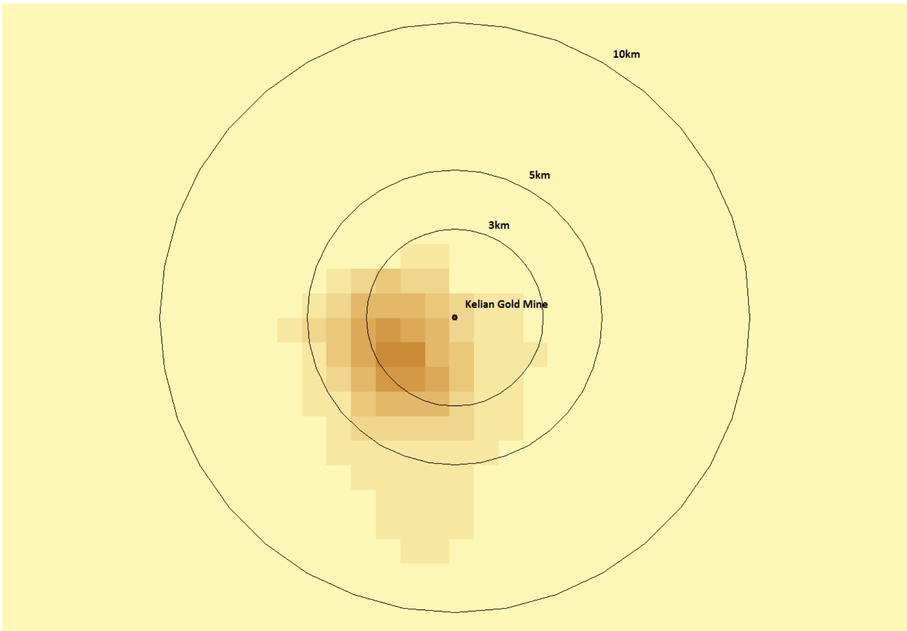

Figure 3.9: Estimating the economic footprint of mining activity: The closing of Kelian gold mine, pre- versus post- 1995 (date of mine closure)

Table 3.8: Estimation of mine footprint: proxied by nighttime lights

	(1) Light 3km	(2) Light 5km	(3) Light 10km	(4) Light 15km	(5) Light 20km	(6) Light 25km
Mine close dum	-0.501**	-0.481**	-0.218**	-0.095	-0.061	-0.042
	(0.206)	(0.186)	(0.107)	(0.079)	(0.066)	(0.059)
Constant	2.215***	2.656***	2.440***	2.211***	1.957***	1.755***
	(0.173)	(0.156)	(0.116)	(0.105)	(0.108)	(0.115)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Panel	Fixed effects	Fixed effects	Fixed effects	Fixed effects	Fixed effects	Fixed effects
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Obs	6327.000	6327.000	6327.000	6327.000	6327.000	6327.000
F-Test	0.000	0.000	0.000	0.000	0.000	0.000
Adj. R-squared	0.184	0.165	0.197	0.219	0.230	0.232

Controls include year dummies and province fixed effects. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.9: Estimation of mine footprint: proxied by change in nighttime lights

	(1) Δ Light 3km	(2) Δ Light 5km	(3) Δ Light 10km	(4) Δ Light 15km	(5) Δ Light 20km	(6) Δ Light 25km
Δ Mine close dum	-0.037	0.052	0.027	0.039	0.012	-0.026
	(0.131)	(0.100)	(0.060)	(0.054)	(0.042)	(0.035)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Panel	First Diff.	First Diff.	First Diff.	First Diff.	First Diff.	First Diff.
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5994	5994	5994	5994	5994	5994
F-Test	0.000	0.000	0.000	0.000	0.000	0.000

Controls include year dummies. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

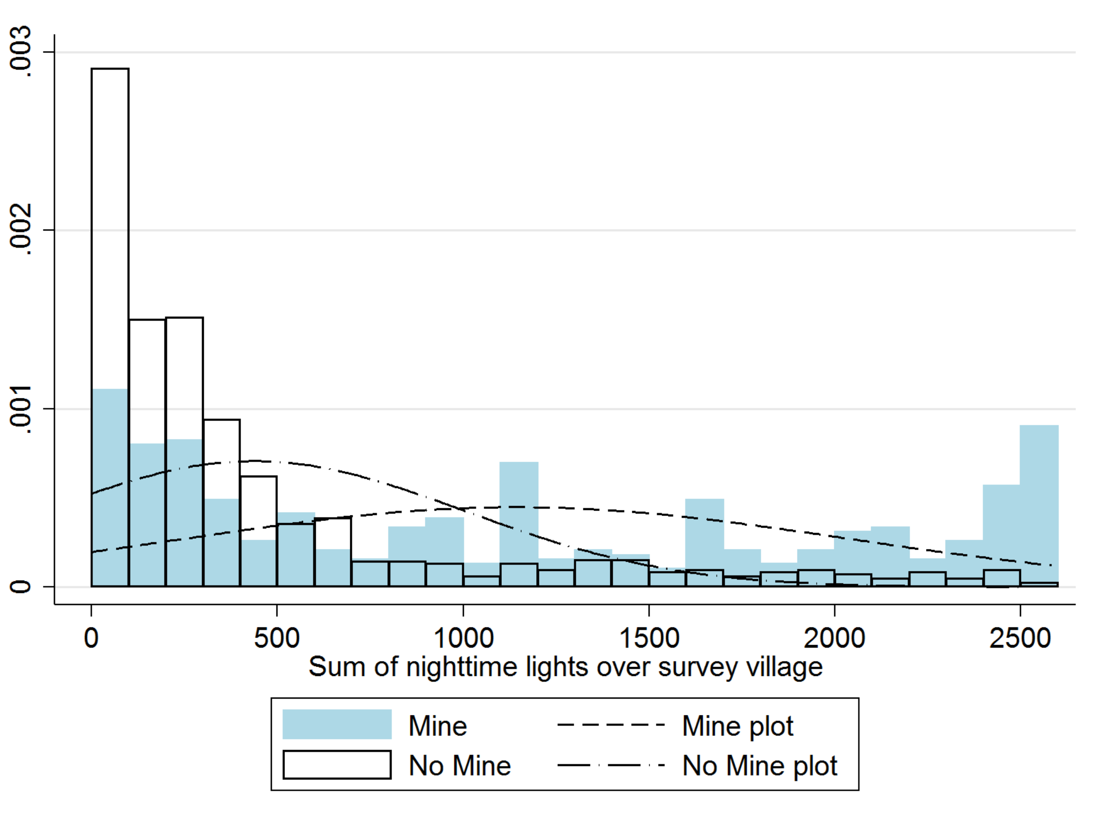

Figure 3.10: Plot of density of nighttime lights for mine and non-mining communities, defined by 15km band. Y-axis shows the density, X-axis the digital light value within the distance band.

3.5.3 Estimating the labour market effects

The main labour market results are presented in Table 3.10, as well as alternative specifications found in Tables 3.13 and 3.12. Here I explore the relationship between the treatment variables - the proximity of an open mine to the survey cluster, and whether a mine is open in a given year[12] - and the labour market effects of interest. Here I define the production treatment for all those mines that fall within a given distance cut-off of the community. To determine the right distance cut-off I use the nighttime lights calibration discussed above.

I present estimates for the local labour market effect on the various sectors individually. As predicted by the model I find a negative and significant effect on labour employed in the traded sectors ($\hat{\beta} = -0.007$), versus a positive effect on the non-traded sectors ($\hat{\beta} = 0.008$), with the fixed effects estimation. In contrast labour employed in the resource sector does not differ significantly from zero in the estimates. This supports my priors that the effect of direct employment in the resource sector may be quite modest for industrial mines in Indonesia. However it also supports the view that local spending effects can induce a shift of labour market activity from traded to non-traded sectors.

Where I include a continuous measure of distance I am able to generate estimates for how this effect dissipates across distance from the mine. However, since the distance measure does not time-vary, I cannot apply fixed effects here. Columns 1, 3 and 5 in table 3.10 presents the estimates. They support the main fixed effects estimates. A producing mine with a given distance band is associated with a positive effect on the non-traded sector, but this effect decreases over distance ($\hat{\beta} = -0.017$). In contrast the effect for the traded sector is negative with a given distance band, but increasing with distance from the mine ($\hat{\beta} = 0.019$). These effects are roughly symmetric, since the only available alternative labour market outcomes are unemployment (including job search) or working in the resource sector.

A caveat should come with interpretation of these results. Taken together the light footprint estimates, plus these estimated structural changes in the labour market, constitute evidence pointing towards a local boom. However, recent work by Aragon and Rud (2012) has indicated

[12]Tables 3.13 and 3.12 also present estimates using mine production volume

that mining, particularly polluting mining, may impose a negative externality on agricultural productivity. While I cannot reject this hypothesis for Indonesia using these tests alone, the labour market shift does not seem to be explained by declining agricultural productivity alone. We discuss this further in section 4.6. I consider it an important avenue for further investigation to examine incomes and welfare outcomes to see if the costs and benefits from mining accrue unevenly in the local area.

Table 3.10: Estimation of mine proximity on sectoral composition of employment: all sectors - community level

	(1) Traded sector	(2) Traded sector	(3) Resource sector	(4) Resource sector	(5) Nontraded sector	(6) Nontraded sector
Log dist nearest mine	0.019***		-0.001		-0.017**	
	(0.006)		(0.001)		(0.006)	
Log mine prod MT, 5km	-0.009*	-0.007*	-0.000	-0.001	0.009*	0.008*
	(0.004)	(0.003)	(0.000)	(0.001)	(0.004)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community F.E.		Yes		Yes		Yes
Obs	908	908	908	908	908	908
F-Test	0.000	0.000	0.328	0.509	0.000	0.000
Adj. R-squared		0.026		-0.001		0.026

Controls: Urban dummy, travel time to provincial capital and community electricity dummy. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Distance to nearest mine should be interpreted as treatment being a decreasing function of distance (the opposite of mine presence dummies.

3.5.4 Estimating the effect of foreign versus domestic ownership

Table 3.11 presents the results for foreign-owned mines only. Contrast these to the magnitudes estimated in Table 3.10. Here I observe point estimates for the traded and non-traded sectors roughly double those for the whole sample estimates. For example, column 6 puts the effect of log Mine production (in MT) at 0.013 on non-traded sector employment, contrasted to the earlier estimate of 0.008 in Table 3.10 for the full mine sample. This suggests a stronger resource movement and spending effect in the presence of foreign-owned mines. According to the model presented this heightened effect would be associated with larger values of parameter θ. Recall parameter θ determines the extent of value and profits retained locally. This can be thought of as the returns to capital accruing domestically in a macro setting, however in the micro setting this might be analogous to increased local procurement of goods and services. Therefore, Table 3.11 provides evidence to support the view that foreign-owned mines may have a stronger local impact than domestic-owned mines.

Table 3.11: Estimation of foreign-operated mine proximity on sectoral composition of employment: all sectors - community level

	(1) Traded sector	(2) Traded sector	(3) Resource sector	(4) Resource sector	(5) Nontraded sector	(6) Nontraded sector
Log nearest foreign	0.002		0.000		-0.003	
	(0.007)		(0.001)		(0.007)	
Log foreign prod., 5km	-0.015***	-0.013***	-0.000	0.000	0.015***	0.013***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community F.E.		Yes		Yes		Yes
Obs	908	908	908	908	908	908
F-Test	0	908	8.21e-12	908	0	908
Adj. R-squared		0.025		-0.002		0.024

Controls: Urban dummy, travel time to provincial capital and community electricity dummy. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6 Robustness tests

All reported standard errors are heteroscedasticity-consistent and are corrected for clustering in the sampling (at the community or PSU level). Where we have a mix of individual and community-level characteristics observed, but with estimation at the individual-level (as in Table 3.12), uncorrected standard errors will be biased downwards. This is because estimation at the individual level overstates the sample size for community characteristics. Deaton (1997) has emphasized the importance of correcting the standard errors when including aggregated community-level variables. This is of particular importance in this context where both forms of treatment: mine size and mine distance are defined relative to the community, rather than at the individual level. Likewise we would expect factor reallocation in response to these shocks to occur via local labour markets- again determined at the community rather than individual level. Clustering of standard errors at the community level imposes quite severe conditions on the data. However, its absence suggests a greater degree of variation in treatment than is present in the data. The results are robust to these corrections.

I also face an important measurement error challenge. One measure of treatment deployed is the mine output volume. This may be problematic as it may not be closely tied to the timing or magnitude of local employment or sourcing by the mine. Indeed, periods of construction and development, when mine status is not yet classified as open may constitute significant periods of local spending. I cannot observe this directly, neither the local employment by the mines, nor their sourcing policies. Therefore production volume as a proxy must be used.

3.6.1 Alternative specifications

Tables 3.13 and 3.12 present alternative specifications for the main labour market estimates. Here estimates are made at the level of individuals, rather than community aggregates. However it is important to note that treatment is still applied (via distance to mine site) at the community location level. Therefore clustered standard errors must also be applied at this level.

First I conduct separate estimates for the traded and non-traded sectors. In each of these

Table 3.12: Estimation of mine proximity on employment in the non-traded sector- individual level

OLS:						
	(1)	(2)	(3)	(4)	(5)	(6)
	Non-Traded	Non-Traded	Non-Traded	Non-Traded	Non-Traded	Non-Traded
Mineral prod. in 5km				0.022***	0.011***	0.011***
				(0.000)	(0.000)	(0.000)
Dist. nearest mine				-0.001	0.004	0.022*
				(0.000)	(0.000)	(0.000)
Mine dummy 5km	0.047***	0.021**	0.016**			
	(0.015)	(0.011)	(0.010)			
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies		Yes	Yes		Yes	Yes
Province F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Community F.E.			Yes			Yes
Obs	121022	121022	121022	121022	121022	121022
F-Test	0.000	0.000	0.000	0.003	0.000	0.000
Adj. R-squared	.0022			.000463		

Controls: Urban dummy, travel time to provincial capital and community electricity dummy. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Distance to nearest mine should be interpreted as treatment being a decreasing function of distance (the opposite of mine presence dummies.

specifications I use the original treatment variables- whether a mine is open within a given distance band, and a continuous measure of distance from the survey cluster to the nearest operational mine. I supplement these with the continuous (and time-varying) measure of mine output in each of the sample years. The estimates are consistent with the main findings presented in Table 3.10. In Table 3.12 I find a positive coefficient on the open mine dummy ($\hat{\beta} = 0.0016$) and mine production variable ($\hat{\beta} = 0.011$) for the non-traded sector. For Table 3.13 I estimate the effect of treatment on participation in the traded good sector. Here I find the opposite, consistent with the model predications: a negative response to a mine being open ($\hat{\beta} = -0.031$), or the log of volume of output ($\hat{\beta} = -0.017$), within a given distance band. The continuous distance effect is not distinguishable from zero once we take account of mineral production levels.

Table 3.13: Estimation of mine proximity on employment in the traded sector- individual level

	(1) Traded	(2) Traded	(3) Traded	(4) Traded	(5) Traded	(6) Traded
Mineral prod. in 5km				-0.025***	-0.011	-0.017***
				(0.000)	(0.000)	(0.000)
Dist. nearest mine				0.003	-0.007	-0.017
				(0.000)	(0.000)	(0.000)
Mine dummy 5km	-0.066***	-0.028***	-0.031***			
	(0.018)	(0.013)	(0.014)			
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies		Yes	Yes		Yes	Yes
Province F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Community F.E.			Yes			Yes
Obs	121022	121022	121022	121022	121022	121022
F-Test	0.000	0.000	.	0.002	0.000	0.000
Adj. R-squared	.00437			.000626		

Controls: Urban dummy, travel time to provincial capital and community electricity dummy. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Distance to nearest mine should be interpreted as treatment being a decreasing function of distance (the opposite of mine presence dummies.

Table 3.14: Estimation of mine proximity on traded sectors: manufacturing and agriculture - community level

	(1) Traded sector	(2) Traded sector	(3) Manu sector	(4) Manu sector	(5) Agricultural sector	(6) Agricultural sector
Log dist nearest mine	0.019***		-0.003		0.025***	
	(0.006)		(0.004)		(0.006)	
Log mine prod MT, 5km	-0.009*	-0.007*	-0.010***	-0.008**	0.001	0.001
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Community F.E.		Yes		Yes		Yes
Obs	908	908	908	908	908	908
F-Test	0.000	0.000	0.000	0.001	0.000	0.025
Adj. R-squared		0.026		0.019		0.023

Controls: Urban dummy, travel time to provincial capital and community electricity dummy. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.15: Agriculture sector - individual level

	OLS:					
	(1) Agriculture	(2) Agriculture	(3) Agriculture	(4) Agriculture	(5) Agriculture	(6) Agriculture
Mineral prod. in 5km				-0.025***	-0.007	-0.014***
				(0.000)	(0.000)	(0.000)
Dist. nearest mine				0.011	-0.002	-0.034*
				(0.000)	(0.000)	(0.000)
Mine dummy 5km	-0.077***	-0.026***	-0.029***			
	(0.018)	(0.012)	(0.012)			
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies		Yes	Yes		Yes	Yes
Province F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Community F.E.			Yes			Yes
Obs	121022	121022	121022	121022	121022	121022
F-Test	0.000	0.000	0.000	0.000	0.000	0.000
Adj. R-squared	.00595			.000784		

Controls: Urban dummy, travel time to provincial capital and community electricity dummy. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Earlier we discussed the correct categorisation of the traded sector. For a largely rural Indonesia, is it right to think of agriculture as facing world or regional prices, versus serving local markets and subsistence uses only? We can test this be decomposing out measure of traded sector employment into manufacturing versus agriculture. When I separate the measures of traded sector employment the overall direction of effects holds. Tables 3.14, 3.15 and 3.16 present these estimates, indicating consistent direction and magnitude of effect. For manufacturing the shift of labour away from the sector appears dependent on the size of mine, given by Columns 4,5 and 6. In contrast, for agriculture shown in Table 3.15, results indicate labour is sensitive to scale of mine (Mineral prod in 5km), distance to the nearest mine and the dummy indicating the presence of a mine, regardless of scale, within the distance band. These results taken together support the earlier assumption that both manufacturing and agriculture be treated as traded goods sectors, from which labour would shift in favour of the booming resource sector and non-traded service sector.

Table 3.16: Manufacturing sector - individual level

	OLS:					
	(1) Manufact.	(2) Manufact.	(3) Manufact.	(4) Manufact.	(5) Manufact.	(6) Manufact.
Mineral prod. in 5km				-0.005	-0.008*	-0.007*
				(0.000)	(0.000)	(0.000)
Dist. nearest mine				-0.010	-0.007	0.022
				(0.000)	(0.000)	(0.000)
Mine dummy 5km	0.002	-0.009	-0.010			
	(0.006)	(0.005)	(0.006)			
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies		Yes	Yes		Yes	Yes
Province F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Community F.E.			Yes			Yes
Obs	121022	121022	121022	121022	121022	121022
F-Test	0.815	0.000	.	0.539	0.000	0.000
Adj. R-squared	-5.01e-06			.0000985		

Controls: Urban dummy, travel time to provincial capital and community electricity dummy. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6.2 Employment and out-migration

Table 3.17 presents two important checks for the main results. First it estimates the effect of mines on overall community-level employment. There is no evidence to suggest mining activity decreased overall employment, and instead suggests a small, but statistically significant decrease in reported unemployment, conditional on remaining within the community ($\hat{\beta} = -0.008$). This challenges the earlier simplifying assumption of full employment both prior to and after the resource boom. This evidence does however support the view that the boom has some positive spillovers on nearby communities (and for example, the labour market adjustment does not reflect falling agricultural productivity, as found in Ghana by Aragon and Rud (2012).

Columns 3 and 4 in Table 3.17 present estimates for out-migration. Here I test for whether communities proximate to mining activity see increased levels of out-migration compared to the non-treated communities. Here I find no evidence to support this. This is consistent with the assumption of fixed labour supply for the communities of interest. Note however I am not able to test for in-migration, except into surveyed households that already form part of the panel in 1993. Therefore this data cannot be used to effectively assess whether mines induce in-migration or draw labour from beyond local communities. This is an important caveat since any such in-migration will play and important role in setting equilibrium prices and wages.

Table 3.17: Estimation of proportion employed and proportion out-migrating

	(1) Unemployed	(2) Unemployed	(3) Out-migrating	(4) Out-migrating
Log dist nearest mine	-0.003		-0.000	
	(0.002)		(0.002)	
Log mine prod MT, 5km	-0.007***	-0.008**	0.004	0.005
	(0.002)	(0.003)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes
Community F.E.		Yes		Yes
Obs	909	909	909	909
F-Test	0.000	0.000	0.000	0.000
Adj. R-squared		0.910		0.820

Controls: Urban dummy, travel time to provincial capital and community electricity dummy. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.7 Conclusion

Using a new dataset on mine location in Indonesia linked to household survey data and night-time satellite data, estimates suggest that mines have an important spatially-concentrated economic footprint. By exploiting the location, timing and scale of mines, the causal effect of mining activity on the local economy can be identified. This work contributes to the growing literature on linkages from natural resource extraction.

The paper provides estimates for the size, intensity and distance gradient of the economic footprint of mining. Using a dataset of over three hundred mines and survey communities, combined with nighttime lights satellite imaging, we can apply a fine resolution of spatial analysis. Results show the typical economic footprint from mining extends up to around 15km, with evidence for an economic boom beyond the mine perimeter fence. This result is strongest close to mines (presumably driven, in part, by light from mining operations), however extends far enough that we are confident of backward linkage effects. Our estimates for mine closure suggest little evidence for a contemporaneous bust- i.e. switching off the lights. However, in our fixed effects estimation we do find a roughly symmetric bust effect to our previous boom effects from mine opening. This time extending only 10km around the mine site however.

Furthermore within the estimated 15km mine footprint the mine has strong effects on the structure of the labour market. Consistent with a local variant of the Dutch disease model developed by Corden and Neary (1982), and predicted by a simple adapted model, results point to a shift of labour from the traded to the non-traded sector. However there is no evidence to suggest employment effects in the booming resource sector among the local survey communities. This evidence suggests that mines have a crowding-out effect on other traded goods sectors, while stimulating a compensating expansion of the services sector. This includes both agriculture and manufacturing. There is no evidence to suggest increased out-migration from communities proximate to mines, and some evidence to suggest a small increase in overall reported employment.

When the sample is split between foreign-owned and domestic-owned mines, I observe a strong labour market effect from foreign mines for a given mine size. This suggests evidence to support the view that foreign-mines have more extensive externality effects. This raises intriguing policy questions. Is it better sourcing practices by foreign-firms that increase this labour market shift? Or stronger crowding-out of otherwise productive traded activities? Furthermore, Indonesia in its recent Mining Law (4/2009) has sought to curtail foreign ownership of mines, limiting them to 49% ownership after a staged period of divestment. Furthermore this Law seeks to limit export of raw ore in favour of promoting domestic value addition. The differential role and effect of foreign-owned mines is therefore a highly contentious debate, worthy of deeper examination.

Chapter 4

Disentangling the effects of resource extraction: quantifying local government and project multipliers

James Cust[1]

[1]This chapter constitutes joint work with Ridwan Rusli, Nanyang Technological University.

Abstract

How does an oil and gas boom affect the regional economy? We examine the economic consequences of resource extraction and associated revenue windfalls in Indonesia, measured at the subnational level. Our analysis focuses on disentangling the effects of variation in petroleum sector activity across districts and municipalities in terms of district GDP. Two important channels are identified: direct spillover effects from extraction investments and production; and fiscal spillovers from local government spending associated with revenue windfalls from extraction. We exploit Indonesia's fiscal sharing rules to isolate these two channels by application of an instrumental variable and spatial identification.

We show that petroleum production is associated with increases in district GDP. The main economic gains are estimated to accrue via transfers to local government having a multiplier effect of around 0.9 for every unit of additional revenue. In contrast, oil sector output appears to have a small though positive effect - estimated at closer to 0.3, but in line with the literature. However, to identify project spillovers separately from revenue effects we focus on districts adjacent to petroleum production. Here we find an attenuated effect. For every dollar of petroleum output in producing districts, neighbouring districts experience no discernable increase in economic output during our sample period. There is however evidence that project expenditures rather than output may contribute positively (an effect of roughly 0.3) while capital expenditures may operate in reverse (-0.2).

We are able to separate estimates for the spillover effects for oil and gas. While estimates for oil cannot be distinguished from zero, the gas spillover effect to neighbouring districts is estimated at 0.1 for every dollar of output in adjacent districts.

4.1 Introduction

Oil and gas booms, like the one experienced in Indonesia between 1999 and 2009, can affect different regional economies unevenly.[2] A question of particular interest is how the differential effects may be propagated. Indonesia is a country that has undergone significant fiscal decentralization, motivated in part to return a greater share of resource revenues to resource producing regions and islands. It is in this setting that we examine the economic consequences of resource extraction and associated revenue windfalls, measured at the subnational level. Our analysis focuses on disentangling the effects from variation in petroleum production across Indonesian districts and municipalities in terms of district GDP. Here we deploy a variety of identification strategies to attempt to disentangle the regional effects of the boom, measured in terms of district GDP. We estimate effects arising from transfers of revenue to local government. Using an instrumental variable approach we isolate the fiscal channel from resource projects to estimate the effect of increased local government revenues on district GDP over the boom decade. Separately we estimate the effects on districts bordering oil producing districts and examine how this differs from non-adjacent districts. This approach allows us to estimate a spillover effect from project expenditures separately from fiscal revenues.

The paper contributes to our understanding of the ways resource extraction affects the subnational economy. First, government revenues, instrumented for by offshore petroleum output, contribute positively to economic output, demonstrating a positive multiplier effect in coastal districts. Second, there is evidence to suggest petroleum extraction projects should not be viewed as purely enclave activities. We show, for example, that project investment spillovers to neighbouring districts can be positive and significant in the short-run. Third, consistent with the wider literature on the Dutch Disease we also find a modest crowding out of manufacturing in favour of a boom in service sector output.

The approach follows closely the analysis of Brazilian oil windfalls by Caselli and Michaels (2013). We extend this analysis to Indonesia, replicating their instrumental variable technique

[2]Throughout the paper we use oil as shorthand for all petroleum sector activity unless explicitly stated, for example when we examine gas production or downstream processing separately

to isolate measures of the impact of government revenue windfalls. In contrast to Caselli and Michaels' (2013) study we find evidence for positive and significant effects of government revenue windfalls on economic output across the sample period. Furthermore this result is fairly large- implying a fiscal multiplier between 0.8 and 0.9 over the decade. We then move beyond the fiscal analysis to attempt to identify project spillovers, in particular those accruing in districts neighbouring petroleum producers. This allows us to observe modest effects via project expenditures on capital and operating expenditures.

Indonesia provides a fascinating setting to attempt to disentangle the effects of fiscal windfalls and increased economic activity in the resource sector. Since 2002 the country has followed a revenue sharing formula which redistributes central government oil and gas revenues back to local government. This has led to significant budgetary windfalls for Indonesian district authorities. Furthermore, Indonesia is a large, diverse and geographically dispersed country. Thus the inflows of capital associated with oil and gas production can have varied and potentially dramatic effects on the structure of the local economy and regional growth trajectories.

Our data suggest that over the decade 1999-2009 local government budgets saw year-on-year increases from oil and gas revenue of as much as 262 percent. Furthermore, inflows of capital during the oil and gas boom reached as much as 17 percent of local district GDP. The overall consequences of the oil and gas boom saw local district GDP rise by 18 percent more over the decade (in current GDP) compared to non-producing districts. Our panel estimates suggest that oil production within a given district is associated with an average 4 percent increase in local GDP across the decade.

Since oil-producing districts saw a boom in government revenues at the same time as a resource sector boom, attribution of causal effects is challenging. We must therefore identify an exogenous source of revenue variation, in the form of an instrumental variable for offshore oil and gas production. This identification strategy exploits the Indonesia Law 25 which established the rules by which coastal districts receive a fixed share of oil and gas revenues from proximate offshore production. This allows us to isolate the effects of oil production on local GDP via the fiscal channel, as opposed to onshore producing regions where production and revenue windfalls occur concurrently.[3]

We also estimate associated effects of oil sector activity on local GDP. Here we are able to use a variety of observed measures of oil sector activity including the value of production, capital and operational expenditures. Again we need to isolate exogenous changes in oil project activity from oil revenue windfalls. We can attempt to identify the spillover effects of resource windfalls by exploiting the spatial discontinuity of fiscal transfers; since producing districts receive a higher fixed share of oil revenues, we can examine the effects on neighbouring regions, while controlling for home districts. Thus, conditioning on the assumption that private-sector activity is more likely to spill across district boundaries than local government spending is, we can estimate an alternative measure of sector effects. Our results suggest there is a small effect on neighbouring districts; a positive spillover in the case of operating expenditures and a negative one in the case of capital expenditures.

The paper is structured as follows: section 4.1.1 provides background on Indonesia while section 4.2 discusses the conceptual framework. Section 4.3 describes the data and section 4.4 the identification strategy. Results are presented in section 4.5, the discussion of robustness measures in section 4.6, and section 4.7 concludes.

4.1.1 Oil and gas in Indonesia

Indonesia has a long-established petroleum sector which has been active for more than 125 years, since the first oil discovery in North Sumatra in 1885. The country now holds proven oil reserves of 4 billion barrels ranking twenty-first among world oil producers, and, eighth among world gas producers (PWC, 2012). Indonesia is a net importer of crude oil and net exporter of natural gas. Oil production peaked in around 2004 at 400 million barrels, falling to around 330 million barrels by 2011. In contrast, gas has steadily risen during the decade, peaking in 2010 at around 9500 mmscf. The output of both rose sharply during the decade 1999-2009, in part due to exogenous price rises in world market, even as production volume in the oil sector fell.

[3] This approach follows closely the study by Caselli and Michaels (2013) who also exploit fiscal sharing rules linking offshore production to coastal municipalities, allowing them to disentangle fiscal revenue effects. The Indonesian rule is based on straight-line proximity.

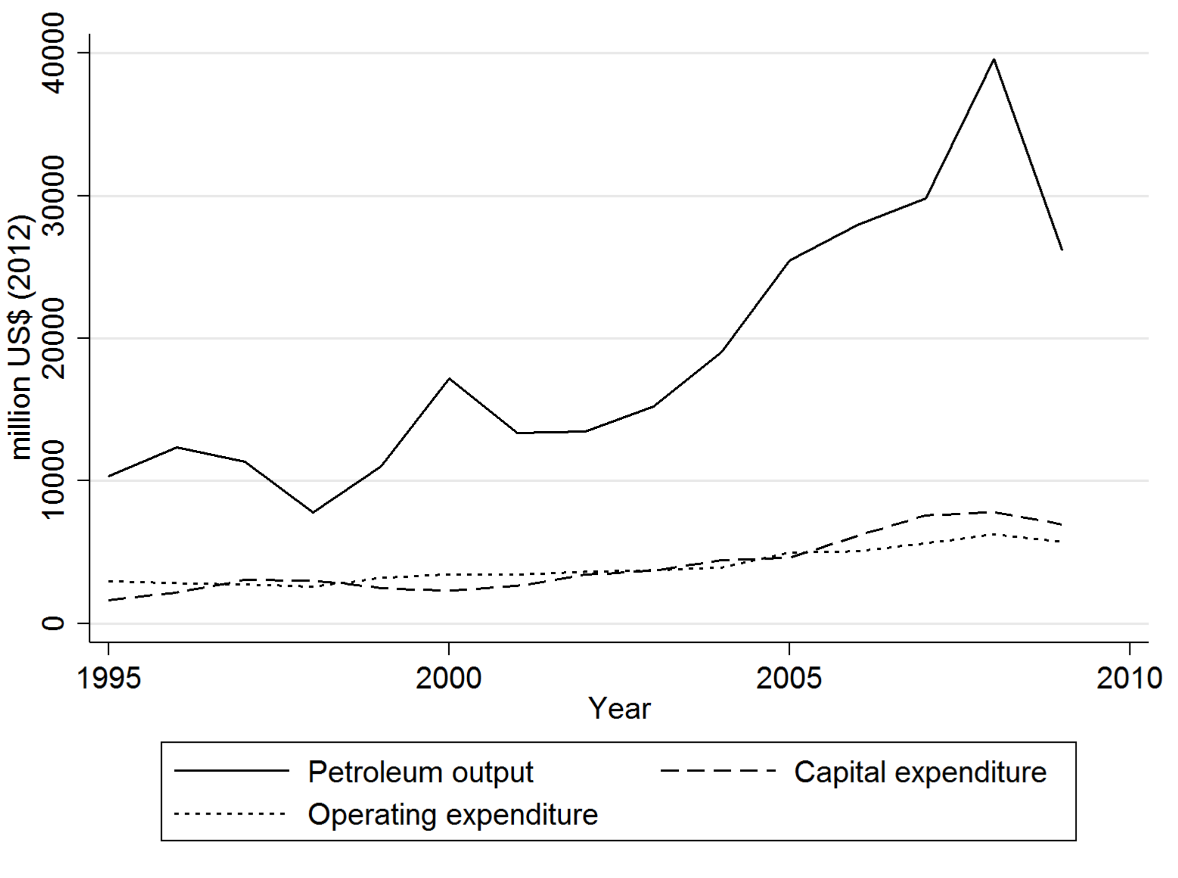

Figure 4.1: Production and investment in the oil and gas sector

Nevertheless oil and gas production has declined between 2011 and 2014. Figure 4.1 shows the patten of oil and gas production and investment in Indonesia since 1995.

The petroleum industry has become an important source of government revenues; it is the single largest contributor by sector (see Figure 4.2). By 2011 over 20% of domestic revenue was derived from the oil and gas sector. Mineral resources are also important in Indonesia. The contribution to GDP from mining has risen from around 4 percent in 2001 to almost 7 percent in 2010. Some regions, such as East Kalimantan and Papua, are particularly dependent on mineral revenues.

The natural resource sectors formed a critical part of the negotiations concerning to Indonesia's democratic transition, beginning in 1998. To understand their contribution to the Indonesia economy after 1998 one must also examine how they fit with the radical decentralization policies pursued over this period.

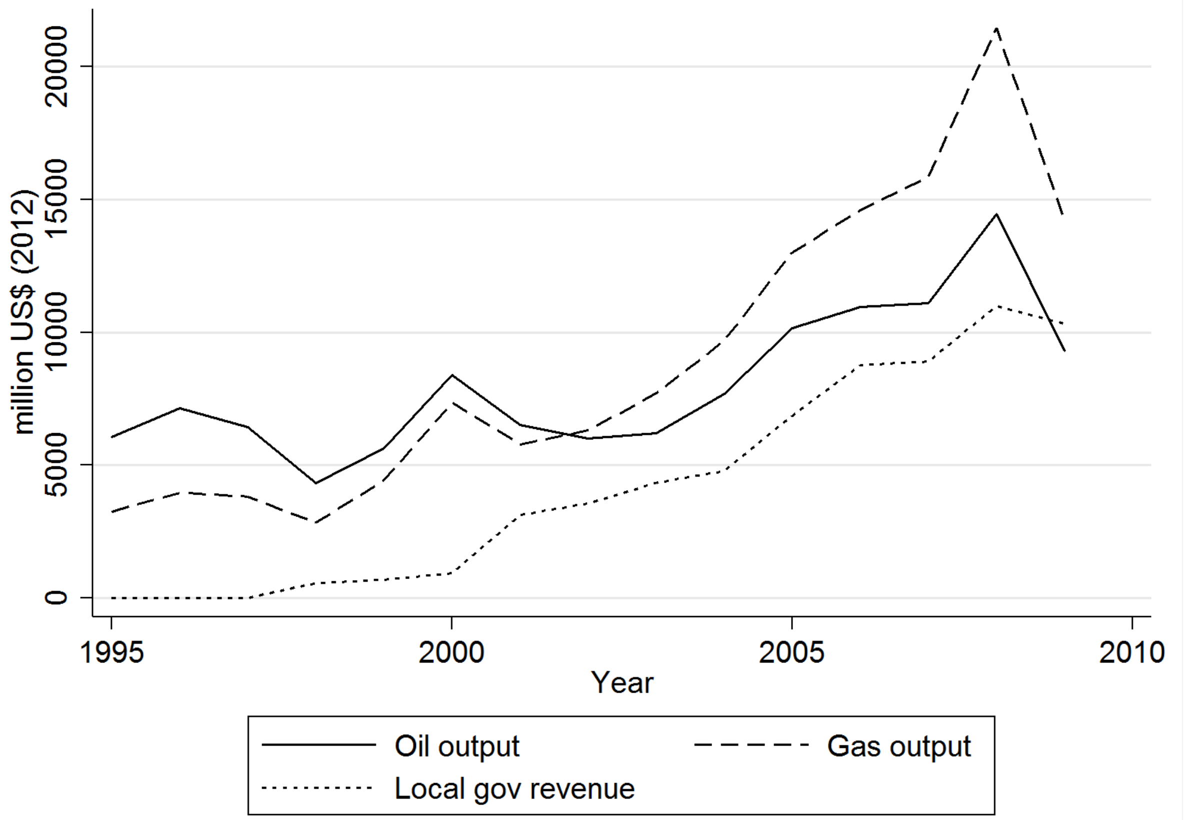

Figure 4.2: Petroleum output value and total subnational government revenues, including transfers

4.1.2 Indonesia's decentralization and government revenue sharing scheme

A major pillar of Indonesia's transition to democracy has been rapid and radical decentralization of power. The resignation of President Suharto in May 1998 was followed by the passage of the 1998 democratic reforms and the 1999 decentralization laws which led to the direct elections of the country's president as well as of local government heads. The country's Law 22 of 1999 on local autonomy and Law 25 of 1999 on fiscal balance granted local governments, in particular the second-tier districts and municipalities (*Kecamatan* and *Kota* respectively), greater autonomy to arrange their own economic, administrative and social policies, including local budgets.[4]

Prior to democratization and decentralization Indonesia's fiscal system was heavily centralized, with about 80% of district spending originating from central government funds, especially subsidies of autonomous regions- so called SDO (*Subsidi Daerah Otonom*)- and presidential instructions, INPRES (*Instruksi Presiden*). The demands of regional stakeholders for greater shares in fiscal revenues associated with natural resource extraction activities in their jurisdictions were a significant driver of the country's decentralization (Barr, 2006).

The general legislative framework for fiscal sharing was provided by Law 25 of 1999. Law 25 allows provincial, district and municipal governments to finance decentralization and their regional budgets from four main sources. The Balancing Funds (*Dana Perimbangan*) comprise the general allocation fund DAU (*Dana Alokasi Umum*), the special allocation fund DAK (*Dana Alokasi Khusus*), as well as natural resource and land tax sharing (*Dana Bagi Hasil*). The core of Indonesia's new system of inter-governmental transfers is the general allocation fund DAU. Every year the central government reserves about 25% of its national budget for the DAU.[5] DAU (and DAK) are meant to close the fiscal gap for each subnational level of government. The natural resource revenue sharing scheme defines how oil and gas, timber and mining royalties are to be re-divided among Indonesia's central, provincial and district governments

[4] See Ford and Fitz (2000), Azis (2008), Duek and Rusli (2010).

[5] The proportion of the special allocation fund DAK, which covers case-specific needs, is generally smaller than DAU.

under regional autonomy.[6] Besides the Balancing Funds, the three other sources of district revenues include local taxes, levies and regional government-owned firms; regional borrowing from domestic sources; and other legal sources.

Greater fiscal autonomy, together with the devolution of power and authority from the central to local governments, has had unexpected consequences. The uneven distribution of resource endowments and revenues across the country has resulted in increased interregional disparity. Compared to historically much lower shares of government receipts since 1999, individual districts home to resource extraction have become better financed, receiving approximately 6%, 12% and up to 32% shares of government revenues for oil, gas, and mining output, respectively. Together with a variety of new extraction-related local government revenue sources and the country's unique post decentralization fiscal allocation scheme, this has led to large windfalls in resource-rich districts prompting unprecedented local spending increases.[7]

As a result, Indonesia's decentralization has brought both benefits and challenges. On the political side decision making now involves local communities, for example through direct local elections. However, while direct elections promote local government accountability and reforms (Asia Foundation, 2008), they also increase the prevalence of local patronage and government-business collusion on the subnational level (Duek et al., 2010). Moreover, Brata (2008) points out that the proliferation of new districts, being one consequence of decentralization, has resulted in inefficient local governments, increased administrative cost as well as higher inequality in income and lower human development index score (HDI).

On the fiscal side higher revenue sharing and budgetary autonomy increase local spending flexibility. Nevertheless, the challenge for the government is to address the weaknesses in the design and implementation of the fiscal sharing program and remain vigilant to potential disparities driven by unbalanced allocations (see Hofman et al. 2006). We examine the limited success of efforts for "fiscal equalization" in section 4.6.2.

[6] While Law 25 extends across a variety of natural resource revenues, we focus our analysis on petroleum production only.

[7] See Barr (2006) and Duek et al. (2010).

4.1.3 Oil and gas contracts and government revenue sharing

Most petroleum companies now operating in Indonesia do so under the terms of a Production Sharing Contract (PSC).[8] The first PSCs were signed in the mid 1960s. Various changes were made over the years. The production sharing agreement is between the contractor and the government, and lasts about 30 years. Day-to-day operations of projects are then managed under the auspices of a regulator, SKK Migas, which is part of the Ministry of Energy and Mining Resources (MEMR).[9]

Signature and production bonuses are levied for the exploration and production periods respectively. The terms of the PSC typically include the government's share of revenues from oil and gas production, which the government receives both in form of monetary and in-kind benefits: bonuses, First Tranche Petroleum (FTP), Domestic Supply Obligation (DMO) for oil, and more recently, gas, as well as a fraction of profits, income and withholding taxes. The government typically receives about 65%-85% of an extraction project's lifecycle revenues.[10]

The portions of oil and gas revenues retained by the government are then allocated between central and regional governments. The oil and gas revenue shares are split and redistributed to the district of origin, in which the oil and gas production takes place (if onshore). Exceptions to this include Special Autonomous Regions who are governed by special resource revenue allocation rules, and offshore production which follows an additional proximity based rule. Additional shares are grants to non-producing districts and municipalities in the province of origin, and to the provincial government. If a contract area straddles several districts the home district allocation is calculated based on the proportion of reserves or production. Table 4.1 summarizes the fiscal revenue distribution that was amended following the Decentralization Law 25 of 1999:

[8]Other existing contracts include Technical Assistance Contracts (TAC) and Cooperation Contracts (KSO) between companies and the national oil company, Pertamina.

[9]The former upstream oil and gas regulator, BP Migas, formed by the 2001 Law 22 governing the oil and gas industry, was recently renamed SKK Migas and integrated back into the MEMR. See Appendix A. When the contractor is successful in exploration and makes a commercial discovery, a plan of development must be submitted and approved by the regulator. Upon start of production the government, through the regulator reimburses the contractor for approved capital and operating costs.

[10]For an overview of Indonesia's oil and gas sector and it's regulatory and fiscal regime see Wood Mackenzie Indonesia Country Overview Wood Mackenzie (2011).

Table 4.1: Fiscal sharing rules relating to resource revenues, since 1999

Revenue Source	Old sharing arrangement	Major change	New sharing arrangement
Oil[a,b]	100% centre	Assignment of revenues after tax deduction to regional governments	85% centre, 3% province of origin, 6% district/municipality (*kecamatan/kota*) of origin, 6% to other districts/municipalities in same province
Natural Gas[a,b,c]	100% centre	Assignment of revenues after tax deduction to regional governments	70% centre 6% province of origin 12% district/municipality of origin 12% other districts/municipalities in the province of origin
Mining land rents	65% centre, 19% province, 16% district/municipality	Continued with new sharing arrangement	20% centre, 16% province, 64% district of origin
Mining royalties	30% centre, 56% province, 14% district/municipality	Continued with new sharing arrangement for districts/municipalities in province of origin	20% centre 16% province, 32% district/municipality of origin, 32% other districts/municipalities in the province of origin

Distribution of oil, gas and mining revenues (Source: Ford and Fitz, 2000, p. 26, updated by authors)

Notes: (a) Additional shares of oil and gas revenues for the provinces of Nanggroe Aceh Darussalam (formerly Aceh) and Papua are stipulated in the Special Autonomy Laws 18/2001 and 21/2001, respectively. In particular, under Law 18/2001, provincial shares of natural resource revenues in Aceh include 70 percent of gas and oil revenues, and 80 percent of revenues from forestry, fisheries, and general mining. Increased oil and gas revenues are reduced to 50% after eight years. Under Law 21/2001, Papua also receives 70 percent of natural oil and LNG taxes and 80 percent of forestry, fishery and general mining. Oil and gas revenues, however, are decreased to 50 percent after 25 years. (b) According to Law 33/2004 (which replaced Law 25/1999), the share of oil and liquid natural gas revenues has been changed so that the central government (the ”Centre” above) would receive 84.5 percent and 69.5 percent of oil and gas revenues, respectively. The share of the regions remains the same. The difference of 0.5 percent of the oil and gas revenues will be allocated for basic education, from which provinces will receive 0.1 percent, originating districts will receive 0.2 percent, and other districts within the provinces will get 0.2 percent. (c) Including Liquid Natural Gas.

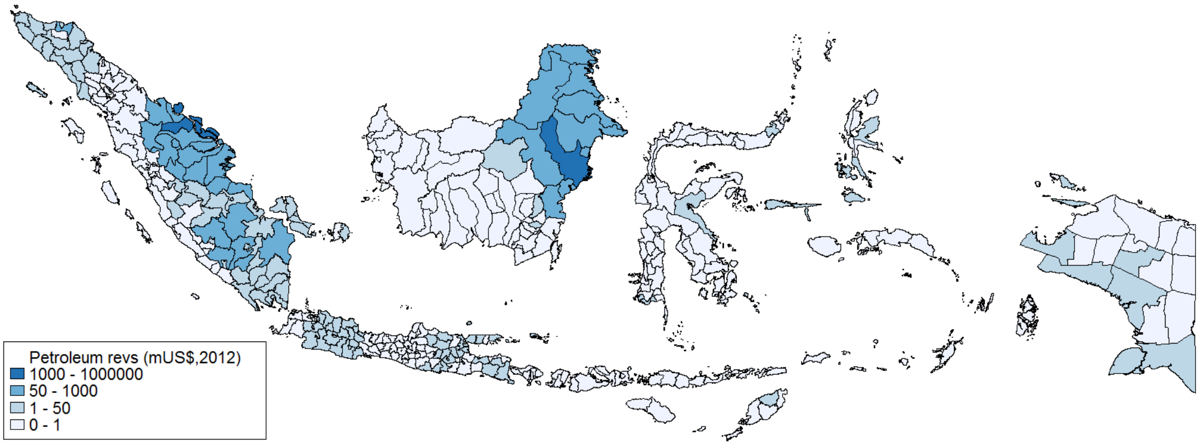

Figure 4.3: Distribution of government petroleum revenue transfers

Resource revenue sharing rules drive the allocation of oil and gas, mining and other royalties, taxes and bonuses across the different districts and provinces. This data is collected by BPS, with input from the relevant ministries e.g. the Ministry of Mines and Energy. Individual districts home to resource extraction have become better-financed, receiving approximately 6%, 12% and up to 32% shares of government revenues for oil, gas, and mining output, respectively.[11] Together with other districts and municipalities in the same province, the provinces of origin receive up to 9%, 18% and 48% of oil, gas and mining-derived government revenues to divide among themselves.

Figure 4.3 shows that the resource and therefore resource revenue-rich regions are located primarily in the islands of Sumatra, Java, the Eastern regions of Kalimantan and West Papua. For our analysis we focus on the relationship between petroleum production (both oil and gas), revenues derived from petroleum production, and economic activity. We therefore ignore the activity and revenues associated with mining, instead controlling for these in our analysis. While the percentage shares associated with mining revenues are much larger than for petroleum (32% of revenue going to district governments versus 6% and 12% for oil and gas respectively), the absolute amount of revenue shared is typically smaller.

[11]While local distribution of mining revenue accounts for a larger percentage share of government take, it accounts for a smaller absolute level of revenue across our sample districts. While not the subject of focus of this paper, we control for mining activity throughout our estimation.

4.2 Conceptual framework

Resource wealth has been associated with both rapid economic development and a range of negative economic phenomena collectively known as the 'resource curse' (van der Ploeg, 2011). Much of the research investigating these dynamics has focused on cross-country comparisons at the national aggregate level; recently studies have begun to examine the effects of resource wealth within countries, and in particular at the sub-national level.

The production of oil and gas can generate sizable government revenues and spur significant inward investment. Resource-rich countries face the combined challenge of managing these windfalls while ensuring that the economy is not adversely effected by the booming sector. This challenge has an important sub-national component, since investments are often spatially concentrated and many countries now distribute significant proportions of petroleum revenues to local government.

Given the spatially concentrated characteristics of oil and gas - the location of which is ultimately dictated by geology - the challenges of harnessing the potential of resource wealth can sometimes be held in the hands of regional or local government. This is particularly true where countries distribute resource revenues to the sub-national level. Many of the regions in receipt of additional revenues may also be host to resource extraction, and the associated capital inflows and economic activities. This paper seeks to disentangle the economic effects of the boom into separate components: the sub-national revenue transfers and the spillover effects from extraction activity.

Studies about the way these factors operate within-country and at the subnational level have become the subject of recent investigation (e.g. Monteiro and Ferraz 2010, Caselli and Michaels 2013, Aragon and Rud 2013 and Beine et al. 2012). While resource rents accruing to government have historically been centrally concentrated, the regional picture matters for several reasons.

First, we are concerned with the drivers of positive or negative economic performance associated with resource wealth. Examination of the subnational and regional impacts of resource windfalls may help us identify the factors associated with the 'resource curse' and its observed causes. Previous work by Michaels (2011) for example, suggests that regional resource abundance can have long term and persistent effects.

Second, extraction activity is spatially concentrated. Where project related spillovers are associated with resource extraction we would expect these to accrue disproportionately to resource rich regions, their neighboring regions and the general locality of extraction. This is due to several factors including extraction activity, the bulk weight of a resource, transportation costs, local spending effects and inward investments. Recent work suggests neighbouring regions can be significant beneficiaries of resource booms via shifts in export patterns (Vermeulen, 2013).

Third, resource revenues are substantial, and thus the spending choices associated with resource windfalls are likely to be an important transmission channel. Where resource revenues are distributed regionally (such as via fiscal sharing rules), these windfalls will affect the regional economic picture. Many countries are now moving to a greater degree of fiscal decentralization, transferring resource revenues back to home districts of extraction activity, and increasingly shifting spending discretion to subnational government units (see Zhang and Zou, 1998, Raveh, 2013 and Arzaghi and Henderson, 2005). Understanding how these units respond to increases in budget and resource windfalls in particular can help us better understand the challenges associated with this trend. Recent work suggests the subnational dynamics are important for understanding whether a country is resource cursed or not (Perez-Sebastian and Raveh, 2012). This study argues that resource booms can lead to structural adjustment and potentially misallocation of labour and capital, depending on the degree of fiscal decentralization.

We seek to examine this in the case of Indonesia. Such analysis requires a greater degree of spatial disaggregation than has typically been feasible with publicly available data. Recent work in this areas follow that of Caselli and Michaels (2013) who exploited spatial variation in offshore oil to look at the effect of revenues accruing to municipal authorities in Brazil.[12] They trace the impact on socio-economic indicators (or lack of impact) from increased public

[12]The parliament and government in Brazil have been debating an increase in, and a more equal redistribution of, oil royalties across municipalities, including those that are not host to the oil production.

expenditure and inward investments. They find no significant impact on economic outcomes, but do detect evidence of GDP composition changes.

This paper extends their approach to disentangle the separate channels of fiscal windfalls and project spillovers arising from a petroleum boom.[13] First, we seek to estimate the fiscal channel effects for Indonesia, complementing the approach used in Brazil. Second, we examine the spillover effects of resource investment and production. This study is made possible by detailed field-level information regarding capital and operating expenditures as well as petroleum output.

This work also contributes to a wider public finance literature that seeks to examine the spending multipliers associated with local government spending and the economic consequences of exogenous economic shocks to a region. A common challenge faced in the literature is in isolating exogenous sources of variation in government spending in order to estimate fiscal multipliers (for example see discussion in Kraay (2012) of a novel method to estimate the effect using World Bank project spending). Likewise, estimating the direct consequences of inward investment projects has occupied a range of authors. There is, for example, some evidence that local projects can have harmful effects on their locality; the study of Enterprise Zones is one example where a variety of authors have documented neutral and even negative effects on local firms (Bartik, 2003). Other studies of local economic damage include a study of the effects of military base closures in the US (Hooker and Knetter, 2001). More recent developments have provided theoretical foundations for this work, such as work on local multiplier effects in Moretti (2010). A broader empirical literature has pioneered innovative identification strategies and use of spatial data in other settings (for the effect of dams see Duflo and Pande 2007; for rural electrification Dinkelman 2011; and for rail privatization Lowe 2014).

[13] Oil and gas extraction have three main beneficiaries. The first are the shareholders, typically private individuals and institutions located in urban centers at a distance from extraction or overseas. The second group are the government, who via equity stakes, contractual terms, and taxation can be recipients large quantities of rents generated by extraction. Countries like Norway are able to capture over 80% of total revenues through their state owned company and taxes. This revenue can accrue to the center, or, as in the case of Indonesia, be shared with subnational administrative units following fiscal sharing rules. Such transfers can increase economic activity at the subnational level through spending effects. Third are the regions where extraction activity is located, who benefit through direct capital outlays, employment and wider external effects such as agglomeration benefits arising from transportation, downstream processing of extracted resources, or by use of the resource as an input in production.

4.3 Data

This paper uses a new dataset, bringing together spatially disaggregated data spanning resource extraction and investment, economic statistics, government revenues, transfers and expenditures. Our units of analysis are local administrative districts, *Kabupaten*, the third administrative level in Indonesia, below the central and provincial governments. We use data from the Indonesian Statistical Bureau BPS (*Badan Pusat Statistik*) on district-level economic activity, population, government revenues and expenditures, combined with our oil and gas production and investment information. We obtain data for the 515 districts and municipalities in 1999. Subsequent to 1999 there is extensive splitting and creation of new districts. We therefore rely on the original (non-splitting) district boundaries. This avoids problems of endogeneity created by the underlying reasons for new districts post-1999, some of which were likely motivated by the opportunity to capture resource revenues.

Our oil and gas data is drawn from the Wood Mackenzie PathFinder database (2011). This dataset contains records on over 1200 individual Indonesian oil and gas wells and fields (fields typically contain multiple well sites). The commercial database collects investment and production data for all major assets in Indonesia across state-owned and privately-owned operations. We utilise the location of investment and production data and associate them with the administrative districts they lie within, or proximate too (see Figure 4.4).

Around thirty-five percent of our total sample of production fields are located offshore in Indonesian waters. This data can be linked to the most proximate coastline and linked to its home district as defined by the fiscal sharing rules. Linking offshore production (and associated revenues) to home regions allows us to employ this measure as an instrumental variable for government disbursements to the district otherwise isolated from direct project-related investments.

We link oil production to districts using spatial joins (see Figure 4.4 for locations). An offshore oil and gas field is linked to a home district for revenue sharing purposes if it is located within a distance of 3 miles to the closest coastline. If the field is between 3 and 12 miles from

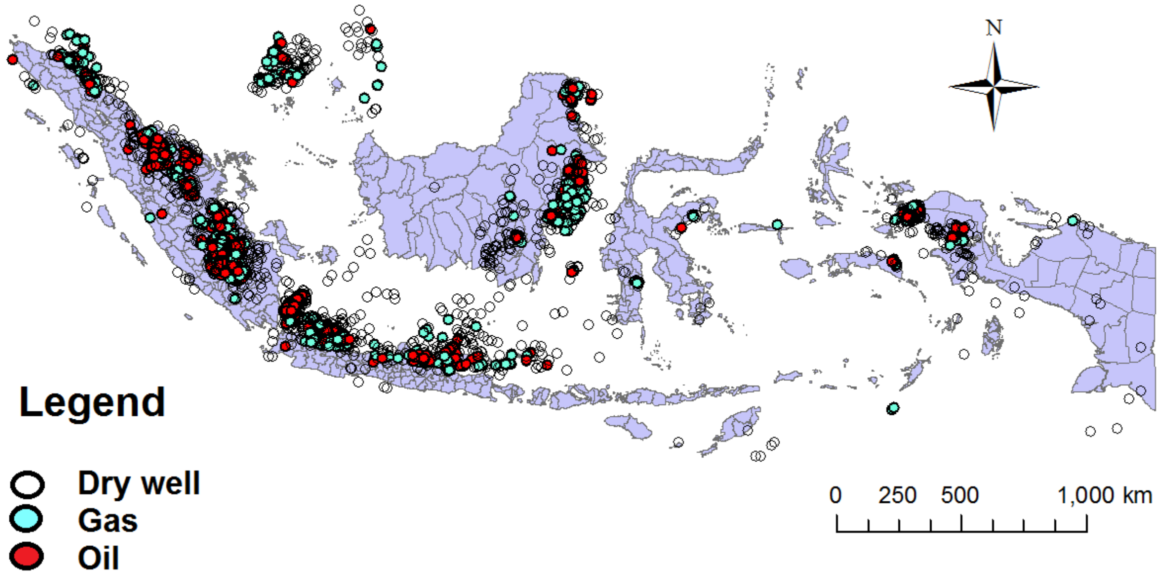

Figure 4.4: The location of oil and gas wells in Indonesia

the closest coastline, the nearby province (rather than district) is designated as home province and recipient of funds. Beyond 12 miles the entire project revenues are allocated to the central government. We therefore create a linking dummy for those fields that fall within the various distance parameters laid out in the formula. This then allows us to construct measures of oil output that contribute to revenues destined for specific coastal districts. This forms the basis for our instrumental variable.

Similarly, for our analysis of project spillovers to neighboring districts, we must first associate adjacent districts with oil producers. Here again we create spatial joins in the ArcGIS software package, allowing us to identify those district borders which lie in adjacent to oil producers, versus districts in the same province but which are not adjacent. This distinction across space is central to our identification strategy, since all districts in a oil-producing province (except the home district are treated equally), whereas we hypothesise that adjacent districts are more likely to see project spillover effects.

Around forty percent of the fields (474 out of 1280) are discovered after 1991, around twenty percent of the fields started production post-2000, and around half of those fields onshore and half offshore. This degree of exogenous variation allows for an identification strategy using a fixed effects estimator to examine the temporal variation in oil output and therefore government spending, once we have controlled for total government budgets.

4.3.1 Gross Regional Domestic Product (GRDP)

Our data on economic performance is drawn from the BPS. Following the United Nation's SNA 1993/2008 approach, BPS compiles GRDP data across districts, municipalities and provinces as well as national GDP figures using both the production and the expenditure approach.

Using the production approach, these figures are calculated based on the value added generated by productive factors of labor, land and capital. In terms of sector output the data is divided into agriculture, livestock, forestry and fishery; mining and quarrying (which includes onshore oil, gas and mining), manufacturing; electricity, gas and water; construction; trade, hotel and restaurant; transportation and communication; finance, real estate and business services; and services. On the expenditure side GRDP data comprise private and government consumption, gross fixed capital formation as well as export and import data.[14] Sectoral GRDP data, including for onshore oil and gas is based on reported value-added.[15] We use annualized GRDP data from BPS, reported at the level of districts and municipalities for the years 1998-2009. Our annualized GRDP is calculated using the latter value-added method.

4.3.2 Government revenues

Our data on government revenues is drawn from official records held by BPS. This data contains reported resource revenue shares for government districts from the period 1999-2010, as well as district government revenues (including local and shared taxes, general and special allocation funds). Government receipts are broken down by source and separately identify non-tax receipts including those revenues drawn from resource extraction.

[14] See also notes on GDP and export-import statistics in Strategic Data, 2010/2011, National Statistics Agency (BPS), Jakarta, pp. 72-82.

[15] McCulloch and Malesky (2011) discuss other approaches using district-level GRDP for Indonesia.

In Table 4.2 petroleum sector dummies indicate the presence in a district of oil, gas or any petroleum production. The offshore production dummy denotes production outside but associated with a particular district. The remaining variables present measures of petroleum production volumes and output, as well as investment and expenditure associated with particular oil fields, all linked to a specific district: they either fall within it, or proximate to it (where oil is offshore).

Table 4.2: Summary petroleum sector statistics, 2009

	(1) Obs.	Mean	Std. Dev.
Petroleum production dummy	515	.101	.302
Oil production dummy	515	.0874	.283
Gas production dummy	515	.0641	.245
Neighbour is a petroleum producer dummy	515	.449	.498
Offshore petroleum dummy	515	.0447	.207
Total Production- onshore (thousand bpd)	515	5812	56252
Total Production- offshore (thousand bpd)	515	1902	20048
Value of onshore production (mUS)	515	383605	3712617
Value of onshore production (mUS)	515	125547	1323161
Capital expenditure (real 2000 million US)- onshore	515	11215	101850
Operating expenditure (real 2000 million US)- onshore	515	9116	104519
Capital expenditure (real 2000 million US)- offshore	515	2316	30712
Operating expenditure (real 2000 million US)- offshore	515	2044	21822

In Table 4.3 we distinguish between oil-producing and non oil-producing districts. We show district GRDP converted to US$ and presented in per capita terms. We also present population levels, a dummy for whether a district is designated a city or not. Further we present information on the levels of district government revenues in total, as well as those derived from oil and gas output.

Table 4.3: Summary producer and non-producer district statistics, 2009

	(1) Obs.	Mean	Std. Dev.
GRDP production approach, million US$2012	463	1821	6801
GRDP per capita, US$2012	463	167	248
Population, interpolated, all years	459	935064	3216166
Govt oil revenue, million US$2012	463	3.55	34.1
Govt gas revenue, million US$2012	463	2.32	27
Total govt revenue, million US$2012	463	17.9	77
Neighbour is a petroleum producer dummy	463	.406	.492
City dummy for specific districts	463	.192	.394

Petroleum production=0

	(1) Obs.	Mean	Std. Dev.
GRDP production approach, million US$2012	52	1893	2473
GRDP per capita, US$2012	52	419	640
Population, interpolated, all years	52	592931	541197
Govt oil revenue, million US$2012	52	13.9	27
Govt gas revenue, million US$2012	52	9.51	31.2
Total govt revenue, million US$2012	52	39.2	55.1
Neighbour is a petroleum producer dummy	52	.827	.382
City dummy for specific districts	52	.135	.345

Petroleum production=1

2012 exchange rate: 9390 IDR to 1 US dollar: http://www.xe.com

4.4 Empirical strategy

We investigate the economic effects of oil extraction through the linkages associated with project investments and through government revenues disbursed at the subnational level. We build an aggregate district-level panel dataset comprising data on economic performance, petroleum output and investment, and government revenues from various sources. This allows us to retain spatial disaggregation in order to analyze the economic effects of oil windfalls at a sub-national level across space and time.

Our hypothesis posits that oil and gas production affects the regional economy through a variety of channels. In the context of a booming resource sector, high world prices and newly introduced revenue sharing rules, investigation of these channels, and evaluation of their relative magnitudes can shed light on the dynamics of resource wealth at the sub-national level.

Are oil-related investments growth enhancing? Are local government windfalls well spent? And does distribution of oil revenues to oil-rich districts exacerbate local Dutch disease effects? Our approach is to employ identification strategies to isolate these various channels and explore their relative importance. We exploit the offshore characteristics of around 35% of our oil production activity to instrument for district oil revenues in those coastal districts defined as home regions to the offshore oil according to Indonesian fiscal sharing rules.

4.4.1 Identification of fiscal channel effects

A common challenge in identifying the size of the fiscal multiplier lies in isolating exogenous sources of variation in government spending (e.g. Kraay 2012). Fiscal revenues and expenditures may be connected to other economic shocks in a region via automatic stabilizers and other government responses. Making causal inference on changes in government spending is therefore challenging given this (sometimes mechanical) level of endogeneity. By contrast, our variation in oil production levels across time and space, and the connection between project revenues and government budgets, allows us to deploy an alternative approach.

Since 2002, oil production has been directly tied to local government revenues via fiscal sharing rules outlined in section 4.3. Under conditions of exogenous variation in oil output, the fiscal sharing rules can be exploited to predict an exogenous source of government revenues. The sharing formula prescribe that districts with oil production receive a disproportionate share of government revenues associated with the value of oil and gas. Therefore, for the purposes of estimating the effects of government spending on local economic activity alone, we must find a way to disentangle the fiscal effects of windfall revenues from any direct effects from oil production and project expenditures.

This approach follows the one developed by Caselli and Michaels (2013) for estimating the impact of oil revenue windfalls in Brazilian municipalities. Here we exploit variation in offshore oil production and the specific features of Indonesian fiscal sharing formulae connecting these projects to proximate coastal districts. This allows us to instrument for revenue windfalls at the level of coastal districts using the offshore oil fields most proximate to each district.

For our instrumenting strategy to be valid we depend on the distribution of offshore oil production to be exogenous to economic conditions in designated home regions and that sharing rules are sufficiently arbitrary that offshore oil benefits those regions only through government expenditures and not other channels. As in Brazil, Indonesian fiscal sharing rules display a degree of arbitrariness, connecting offshore rents to home districts via proximity rules (3 miles of nearest coastal point), rather than landing offtake points, refineries or transportation hubs.

The instrumental variable (IV) estimation is shown below, whereby we estimate the effect of government oil-revenues, instrumented by offshore oil production, on district-level outcomes. Here we distinguish between onshore and offshore oil and gas production Oil_{it}^{ON} and Oil_{it}^{OFF}:

$$
\begin{aligned}
Y_{it} &= \beta_1 \widehat{Revenue}_{it} + \beta_2 Oil_{it}^{ON} \\
&\quad + X_{it}' \beta_X + \xi_t + \varepsilon_i + \kappa_{it}
\end{aligned}
\tag{4.1}
$$

And the first stage using the instrument of offshore oil revenues:

$$
Revenue_{it} = \gamma_1 Oil_{it}^{ON} + X_{it}' \gamma_X + \nu t + \alpha_i + \eta_{it} \tag{4.2}
$$

Here we can interpret β_1 as the coefficient capturing exogenously determined government revenue and its impact on district-level economic performance. Using offshore oil production and revenues shared with district-level governments, we are able to isolate the government revenue effect from the oil production externalities, thus obtaining an estimate for the impact of the marginal dollar of government oil windfall spending. We present our estimates in two forms, levels specifications, taking 2009 data, and in the changes over the decade (1999-2009). For our levels estimation we use data drawn from the specified year unless missing, in which case we replace with data from adjacent years (for example population data is measured accurately only for 2000 and 2010 due to the completion of national censuses).[16]

[16]For our panel data first-differenced estimation, we construct two time periods 1999 and 2009 using three year averages (i.e. for 1999 this takes the form of $Y = (Y_{1998} + Y_{1999} + Y_{2000})/3$). This is done to help mitigate concerns of measurement error associated with the variables of interest, in particular district GDP. In our extensions we also provide estimates using the annual panel data. Our main findings are robust to a variety of approaches.

$$\Delta Y_i = \beta_1 \Delta \widehat{Revenue}_i + \beta_2 \Delta Oil_{it}^{ON} \qquad + \Delta X_{it}' \beta_X + \xi_t + \kappa_{it} \tag{4.3}$$

And the first stage using the instrument of offshore oil revenues:

$$\Delta Revenue_{it} = \gamma_1 \Delta Oil_{it}^{ON} + \Delta X_{it}' \gamma_X + \nu t + \eta_{it} \tag{4.4}$$

While we can be confident that the source of variation in oil-related government revenues is directly tied to oil extraction in the home district and province via fiscal sharing rules, there are several additional concerns we must address. First, we may worry that the level of oil extraction - and hence a portion of government revenues - is also associated with initial economic conditions of oil and non-oil regions prior to development, and other time-invariant unobserved differences across oil and non-oil districts. Our panel data allows us to account for time invariant district-level characteristics, which we do so in our first-difference estimation.

Additionally, the timing of discovery and production assists our strategy. Around 40% of the fields (474 out of 1280) were discovered after 1991 and around 20% of the fields started production post-2000, with half of those fields onshore and half offshore. Petroleum output experienced a sharp increase across the country during our sample period, in part due to rising world oil prices. This degree of exogenous variation allows for an identification strategy using fixed effects estimator to examine the temporal variation in oil output and therefore government spending, once we have controlled for other factors. However we supplement both our levels and changes estimation with an instrumental variable approach.

Our instrumental variable allows an alternative identification of the pure fiscal revenue effect. Here we are able to isolate district government windfalls from project-related investments by exploiting variations in onshore and offshore oil extraction, and variations in the application of fiscal sharing rules.

According to Indonesia's fiscal sharing rules, only those wells with a given distance band

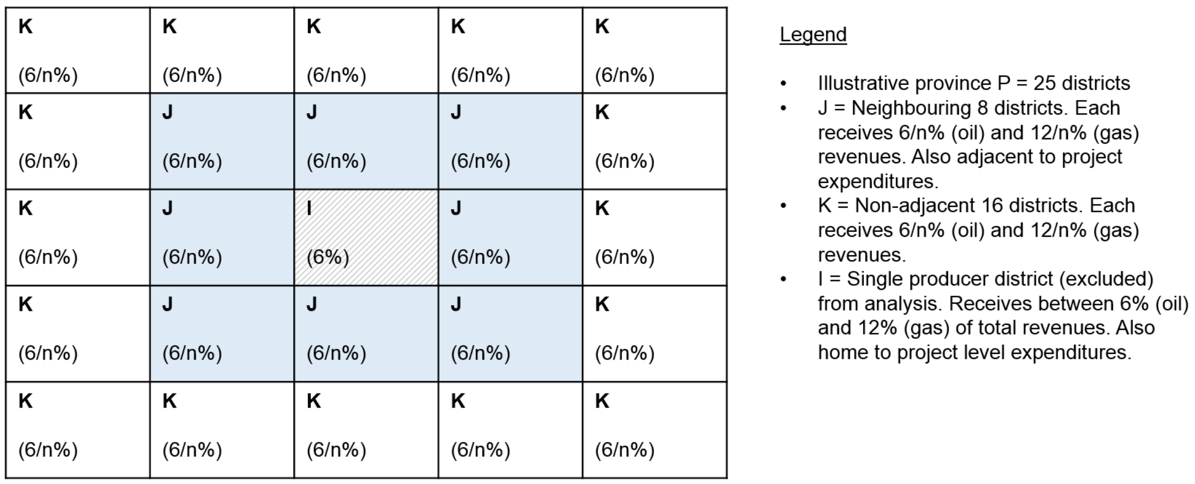

Figure 4.5: Neighbour identification strategy: Illustration of revenue sharing rules in petroleum producing province

(3 miles) are associated to coastal districts. Outside this band, revenues accrue to the province (3-12 miles) and the central government ($>$ 12 miles) respectively. Our $Revenue_{it}$ is therefore the measure of government revenues instrumented using offshore oil within 3 miles of the coast, whose predicted value $\widehat{Revenue}_{it}$ becomes the independent variable in the second stage regression.

4.4.2 Identification of project channel effects via neighbour spillovers

To estimate oil production spillovers, our ideal setting would allow us to observe districts with oil production and associated investments, but no changes in other conditions- in particular no increase in government revenues or expenditures. In the Indonesian setting, post-1999, all districts with oil or gas production also receive a share of government revenue receipts. We must therefore attempt to disentangle the government spending effects from any potential project spillovers.

A possible approach would be to exploit the features of the fiscal sharing rules. According to Law 25, there are three categories of local government transfers. To those districts hosting oil and gas production, between 6% and 12% of the revenues are allocated. To those districts in the same province of the oil production, but not hosting it, they share a remaining 6/n% to

12/n% of the revenues, where n is the number of districts in the province. Remaining districts, i.e. those outside of the oil producing province receive no special transfers, while the remaining share of government receipts accrue to the central government.

The difference between these categories can be exploited to attempt to isolate the project spillovers versus the government spending channel associated with oil production. By focusing on those districts *neigbouring* oil production, but not home to production themselves, we can estimate a spillover effect based on proximity to oil and gas production, but not driven by local spending. See Figure 4.5 for an illustration of the identification strategy.

We compare economic performance indicators in neighboring districts adjacent to districts with oil production with economic performance in other more remote districts within the same province. We hypothesize that neighboring districts, due to proximity, will benefit disproportionately more from direct economic spillover effects than other less proximate districts in the same province, even though all non-producer districts in the same province receive identical revenue shares. We exclude from our analysis all petroleum producing districts.

Therefore, for our unit of analysis, the non-producing district, any GDP effect from neighbouring district oil production can be interpreted as a spillover effect from proximity, rather than an effect of transfers and government spending. We use the following specification:

$$Y_{jt} = \gamma_1 Oil_{it} + \gamma_2 G_{it} + \mathbf{X}'_{it}\delta_1 + \alpha_t + \beta_j + \varepsilon_{jt}. \tag{4.5}$$

Here Y_{jt} gives a measure of economic output in the neighbour district, our new unit of interest (i.e. including those adjacent to our excluded oil-producing districts denote i). Our estimation is therefore defined at the district level j for year t, and measured as GDP in local currency units in terms of total district GRDP. Our Oil_{it} variables capture some time-varying measures of project related extraction activity and investment in the adjacent producing district(s), where our district j neighbours a producer district. Otherwise these variables take the value zero. G_{it} likewise denotes government revenue generated inside the producing district for a given year, and otherwise, revenues for neighbouring non-producing districts. $\mathbf{X}'_{it}$ is a vector of other time-varying district-level controls. We also examine time period dummies α_t and time invariant district fixed effects β_j, including unobserved characteristics of our district of interest.

The validity of this identification strategy also depends upon the structure of the spatial relationship between regional neighbours. Our key identifying assumption is that economic output in neighbouring districts reacts to the exogenous changes in oil sector activity, but that oil sector activity is not influenced in reverse, i.e. oil sector activity does not move across district borders in response to neighbourhood economic activity.[17] Unbiased and consistent estimation by OLS (and first differenced panel) requires that the errors be uncorrelated with unobservable in the neighbouring districts, such that $E[\varepsilon_{jt}|X_{it}] = 0$. This condition would be violated in the case of spatial sorting, whereby oil sector activity is driven by time-variant unobservables in neighbouring districts. Given the exogenous characteristics of oil investments (where location is largely driven by geology) and the world prices driving much of the increases in output observed, we argue these constitute plausible identifying assumptions, particularly in the context of our first-difference estimator.[18]

4.5 Results

4.5.1 Fiscal channel effects

Table 4.4 provides our baseline estimates for the effects of the oil and gas revenue windfalls on district GDP. Our estimates are presented for levels, measured in 2009, indicating the effects of a contemporaneous increase in government revenues from offshore oil and gas production on GDP. Our estimates are in log form- therefore we can interpret the coefficients as elasticities: a one percent increase in government revenue is associated with a 0.45% increase in GDP, while a one percent increase in local government revenues instrumented using offshore oil is associated

[17] This is equivalent to the spatial sorting assumption; examples of this might include parents moving children into neighbouring school districts to seek out better education quality.

[18] In section 4.6 we relax this assumption and include corrections for spatial correlation in our error terms. These corrected standard errors now yield broadly consistent results.

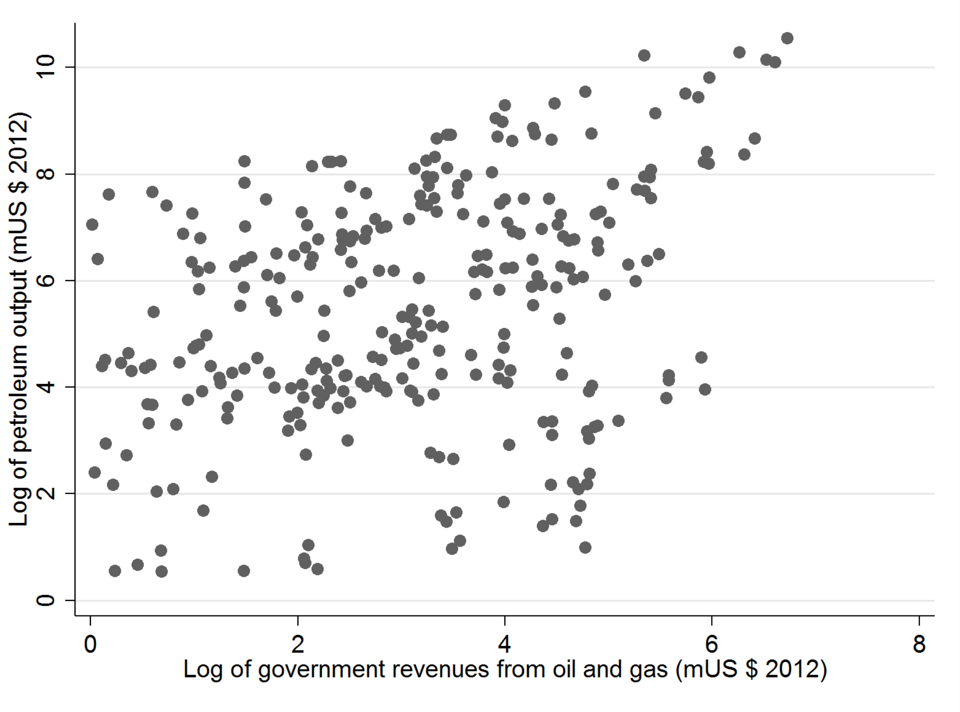

Figure 4.6: Oil production and government revenue

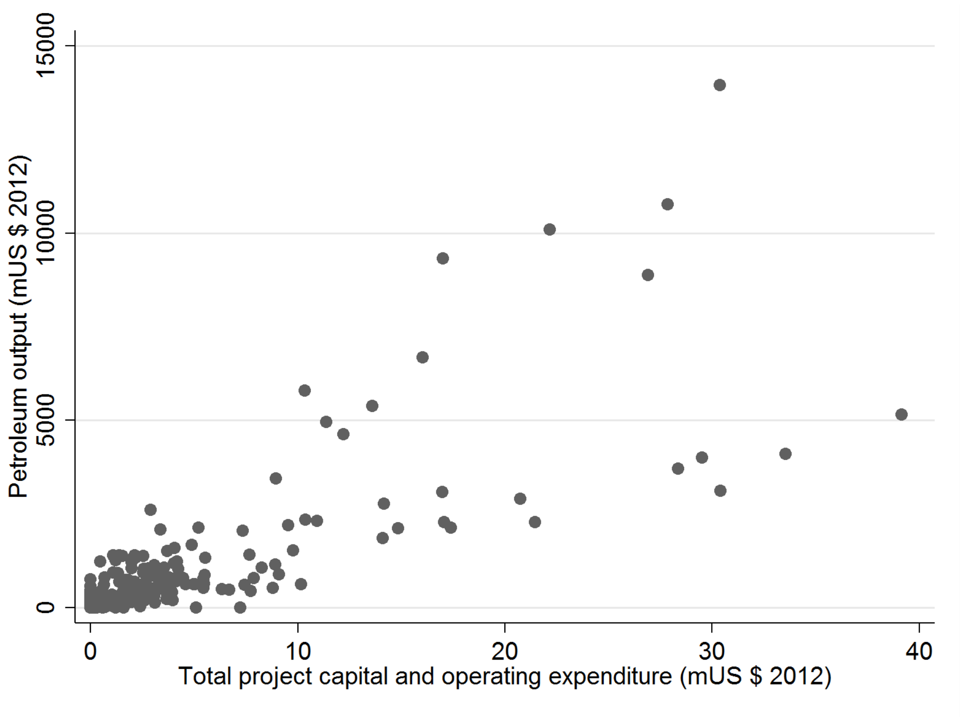

Figure 4.7: Oil field investment and production

Table 4.4: Levels

OLS:				
	(1)	(2)	(3)	(4)
	OLS Log GDP	IV Log GDP	OLS Log GDP	IV Log GDP
Log local govt revenue	0.452***	0.879**	0.426***	0.896
	(0.058)	(0.328)	(0.063)	(0.590)
Log onshore production value			0.025	-0.096
			(0.016)	(0.066)
Log capital exp onshore			-0.031	-0.087
			(0.027)	(0.050)
Log operating exp onshore			0.090*	0.141
			(0.029)	(0.050)
Controls	Yes	Yes	Yes	Yes
Obs	511	511	511	511
F-Test	0.00	.	0.00	.
Underid. test p		.027		.155
Adj. R-squared	.416	.277	.419	.265

Controls: Population, mining revenue, city dummy, Special autonomous and other province dummies. Clustered (at district level) robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

with a 0.87% increase in GDP. This provides initial evidence for a positive and significant fiscal multiplier associated with government spending that derives from petroleum activity.

However, these levels estimates should be interpreted carefully, since we are concerned about unobserved heterogeneity across districts. These time-invariant characteristics can be addressed via our second set of estimations in changes- using our first difference estimator. Here again we run both the specifications on our differenced data (measured between 1999 and 2009). The application of our difference estimation is our preferred specification for reasons discussed above.

Table 4.5 provides an estimated elasticity of around 0.8-0.9% GDP associated with a 1% increase in government revenues. These estimates appear to be stable across the different specifications and controls presented. It is also consistent, though larger than our levels estimation for contemporaneous revenue effects. Since we are able to directly observe oil sector output and expenditures, we include these as additional controls to further reassure ourselves that our identification is not picking up the effect of oil sector activities via our offshore oil instrument.

Table 4.5: Changes

	OLS:			
	(1)	(2)	(3)	(4)
	OLS GDP	IV GDP	OLS GDP	IV GDP
Log local govt revenue	0.799***	0.905***	0.796***	0.900***
	(0.052)	(0.299)	(0.052)	(0.319)
Log onshore oil output	-0.168***	-0.203***		
	(0.028)	(0.057)		
Log capital exp. onshore			-0.143**	-0.167**
			(0.066)	(0.093)
Log operating exp. onshore			0.065	0.057
			(0.090)	(0.094)
Controls	Yes	Yes	Yes	Yes
District F.E.	F.D.	F.D.	F.D.	F.D.
Obs	1000	978	1000	978
F-Test	0.00	.	0.00	.
Underid. test p		.138		.174
Adj. R-squared	0.76	0.49	0.75	0.49

Controls: Population, mining revenue included as controls. Time-invariant controls are excluded. Clustered (at district level) robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

It should be noted that these estimates are strictly local in two respects. First, for the levels estimation they measure only the contemporaneous effect of government receipts on GRDP at the district level, and do not capture regional consequences of spending choices. Second, both IV approaches should be thought of as local average treatment effects - i.e they capture only those receipts arising from proximate coastal oil production which is to say that which falls within three miles of a coastal district.[19] This therefore cannot be interpreted as an estimate for the public spending multiplier in general.

4.5.2 Project channel effects via neighbour spillovers

We directly observe petroleum production, capital expenditure and operating expenditure for each producing district. Our estimation strategy can therefore exploit the location of produc-

[19]Collectively, offshore petroleum accounts for over 30 percent of Indonesia's total oil production and revenues. Our instrument excludes revenues from all onshore oil production, and all offshore oil production beyond three miles from the coast.

tion relative to neighbouring non-producing districts. By estimating the effect on these non-producing districts which neighbour producing districts we can disentangle the spillover effect (from investment and production) from any potential government spending effect. This is because neighbouring districts receive no additional revenue transfers associated with oil and gas compared to other (non-neighbouring districts of the same province).

Table 4.6: Neighbourhood effects: levels

	(1) OLS Log GDP	(2) OLS Log GDP
Neighbour petroleum output	0.011 (0.011)	
Neighbour capital expenditure (onshore)		0.142** (0.022)
Neighbour operating expenditure (onshore)		-0.129*** (0.018)
Controls		
Obs	458	458
F-Test	0.00	0.00
Adj. R-squared	.409	.411

Controls: Population, mining revenue, city dummy, home government revenues and neighbour government revenues, special autonomous and other province dummies. Clustered (at district level) robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Our levels estimation in Table 4.6, for neighbouring districts, finds no evidence that oil output alone has a multiplier that differs significantly from zero. Instead it would seem that any positive spillover from neighbouring oil producers may come via operating expenditures rather than output. Taken together the neighbour spillover effect of capital and operating expenditures suggest contemporaneous effects operating in opposite directions.

For our changes estimation we take first differences, like for the fiscal channel estimate. Our estimates again show no clear relationship between oil output in neighbouring districts and economic activity in the district of interest. Instead there is some evidence for effects of oil sector expenditures. When we include measures of capital expenditure and operating expenditure associated with oil and gas fields, we find some evidence for a positive spillover

Table 4.7: Neighbourhood effects: changes

	(1) F.D. Log GDP	(2) F.D. Log GDP
Neighbour petroleum output	0.088	
	(0.052)	
Neighbour capital expenditure (onshore)		-0.220**
		(0.064)
Neighbour operating expenditure (onshore)		0.294*
		(0.092)
Controls	Yes	Yes
District F.E.	F.D.	F.D.
Obs	832	832
F-Test	832	832
Adj. R-squared	.826	.826

Controls: Population, mining revenue, local government revenue and neighbours revenue. Time-invariant controls are excluded. Clustered (at district level) robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

from operating expenditures (typically personnel and recurrent costs), whereas some evidence for a weakly negative effect for capital expenditures (plant and machinery). These results are robust to using onshore, offshore, or combined project level expenditures in neighbouring districts. For offshore expenditures (not presented) these do not differ significantly from zero which is consistent with our priors that offshore activity generates little economic footprint onshore.

Therefore our spillover effect from operating expenditures in a neighbouring district is estimated at around 0.3 for every $1 of petroleum spending. This is smaller than our estimates for fiscal channel effects in producer districts. However, this effect is stronger than we might expect for a form of economic activity associated with enclave production, being both spatially concentrated and highly capital intensive. Possible explanations could include the important role for transportation infrastructure and downstream processing such as oil refineries, both of which may be located in nearby districts. Over the decade under investigation the spillovers from capital spending appear to have operated in the opposite direction, yielding estimates of -0.2 over the decade. The net effect is not distinguishable from zero. Additionally it should be noted that capital and operating expenditures have opposite signs in the levels data compared to the changes. While we consider the latter our preferred estimates due to taking account of time-invariant factors, it is worth considering why these results may diverge. Recall Table 4.6 estimates the contemporaneous effects in a given year- 2009. Capital expenditure tends to be associated with development and construction phase of a new operation. Operating expenditures, in contrast, are more akin to ongoing costs associated with production and operations. Therefore over the decade, districts neighbouring those with greater *opex* may see larger spillovers for the same dollar increment relative to capital expenditure. On the other hand, capital expenditure may exert and strong short-term effect on neighbouring districts such as via demand for construction firms from the region. Evidence suggests this does not persist however. Further, projects that are relatively capital expenditure heavy, relative to operating expenditures, may have a different degree of linkages to the regional economy versus crowding out effect. Operating expenditures tend to include things like personnel and procurement which may be sourced in neighbouring districts.

4.6 Robustness and Validity of the identification strategy

4.6.1 Validity of the instrumental variable

The validity of our estimation strategy depends critically on the ability to identify exogenous increases in government funds and project-related expenditures. In the case of project related investment, this will depend on two types of decisions taken by private agents. First, the location of oil project investments and production will be determined by the presence of oil deposits. Here, while the distribution of oil *reserves* across Indonesia is likely to be random with respect to other economic characteristics, we may worry that the process of oil discovery is endogenous to province and district-level characteristics. We thus control for district fixed effects. Furthermore, subject to oil being discovered in the district, the decision to proceed with extraction-related investments may also be affected by regional characteristics, such as the investment environment and local taxes.[20]

Table 4.8 presents our first stage estimates for the main levels estimation for our fiscal channel effects. The upper panel presents the first stage indicating the relationship between offshore production, our instrument, and government revenue. The lower panel presents our IV estimation for the effect of government revenues on contemporaneous district GRDP. The first stage estimates support the informativeness principle of our chosen instrument. The upper panel shows that offshore production within the distance threshold (within 3 miles of the coastal district) is a good predictor of a portion of government revenues in 2009. This is also reflected in our first-differenced IV.

Table 4.8: Instrumental variable first stage estimates: levels

First stage:

	(1) 1st stage
Log offshore production value	0.121**
	(0.030)
Controls	Yes
District F.E.	
Obs	511.00
F-Test	5.13e-27
Adj. R-squared	.25

Robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Table 4.9 presents the results for our first-difference instrumental variable estimation. Again these provide support to the informativeness principle discussed above, whereby a change in offshore oil output between 1999 and 2009 is a strong predictor of a portion of increased government revenues in the same period. We use this prediction for our IV estimation in section 4.5.

[20]Exploration and production licenses are issued and controlled by central government in Indonesia. While we control for district fixed effects, this feature provides further reassurance of the exogeneity of oil output to unobserved district-specific characteristics.

Table 4.9: Instrumental variable first stage estimates: changes

First stage:	
	(1) 1st stage
Log offshore (3m) oil output	0.129***
	(0.063)
Controls	Yes
District F.E.	F.D.
Obs	1000.00
F-Test	6.0e-166
Adj. R-squared	.429

Robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

4.6.2 Fiscal equalization

Identification depends upon the increases in government receipts (and spending) associated with oil windfalls. The central government seeks to compensate non-oil rich districts using its general or special allocation funds. However, during the sample period this policy was not effectively implemented. Figure 4.8 illustrates the absence of correlation between the government receipts derived from oil revenues and the general allocation funds (DAU) received by district governments.

As mentioned above, this fund is intended for use as a top-up to help district governments meet their budgetary outlays. In fact we find that the DAU correlates strongly with the district population, the latter being a driver of local expenditures and therefore fiscal deficit as well as DAU allocation, especially in the early days of fiscal decentralization.[21] We thus conclude that in general government budgets are- at most- only weakly linked to oil extraction related local government resource revenues.

[21] See Fitrani et al. (2005).

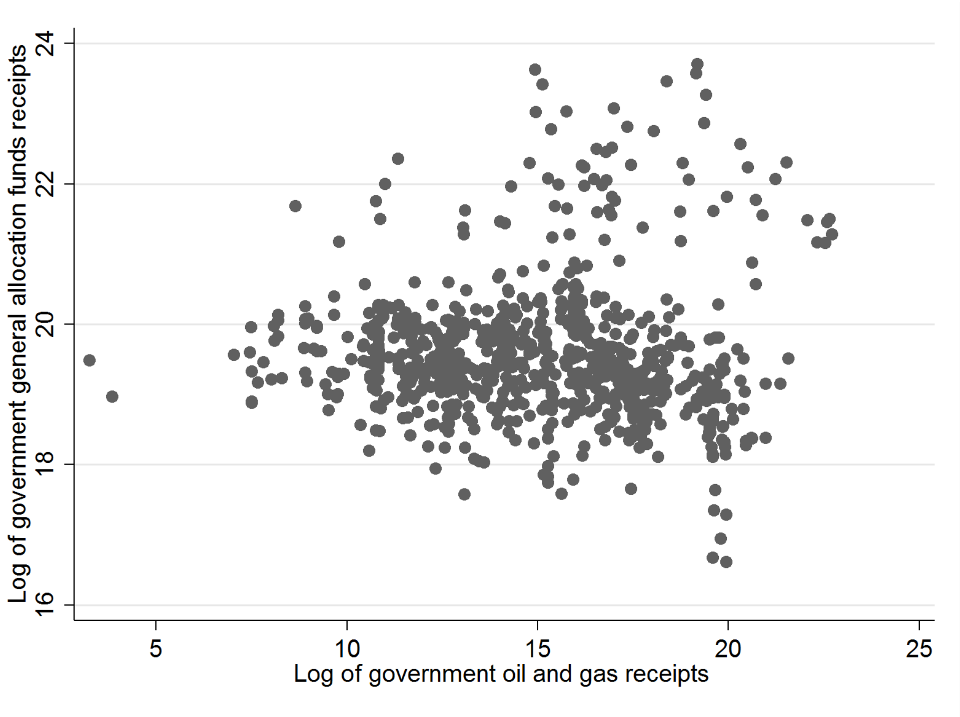

Figure 4.8: General allocation fund and oil revenues

4.6.3 Robustness checks and channels

In this section we present a additional investigations beyond our main specification and identification strategy.

For our neighbourhood spillover estimation we may worry about the structure of the error term with respect to neighbouring districts. As discussed in our results section we may worry about the prospects for spatial sorting, which would lead to inconsistent estimation of the neighbour spillover effect without correcting for this effect via the standard errors. We can therefore apply a spatial weighting matrix to account for spatial correlation. We apply the standard error adjustment of Hsiang (2010), following the method of Conley (1999). Our main results are broadly robust to this correction. Due to the time invariant spatial effects, the spatial correction presented below cannot be readily combined with other panel data techniques including our fixed effects estimator.[22]

[22] An alternative approach to the neighbourhood estimation would be to implement a full spatial panel approach. Examples include the Spatial Durbin model which imposes both spatial and temporal fixed effects. Unfortunately the number of islands (i.e. non-contingent districts and provinces) in Indonesia, as well as our limited time period (low T) make this approach challenging to implement.

Table 4.10: Neighbourhood effects: spatial

	(1) 2009	(2) 2009	(3) 99-09	(4) 99-09
Neighbour petroleum output	0.000901		-0.00578	
	(0.11)		(-1.07)	
Neighbour capital expenditure (onshore)		-0.0527		-0.101**
		(-1.05)		(-2.31)
Neighbour operating expenditure (onshore)		0.0184		0.0835**
		(0.43)		(2.33)
Controls	Yes	Yes	Yes	Yes
Obs	421	421	711	711

Specification 1 and 2: Levels in 2009 with distance cutoff: 250km, Specification 3 and 4: 1999,2009 with distance cutoff: 250km. Controls: Population, mining revenue, city dummy, home government revenues and neighbour government revenues, special autonomous province dummies. Spatially-corrected White-robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. Note however that our spatial correction estimator no longer uses first-differences or Clustered (at district level) robust standard errors.

4.6.4 Difference-in-difference estimation

We would like to estimate the effect of petroleum sector activity within the producing district. However, to do so we need to ensure that we are measuring the impact of investments directly, and not capturing the spillovers associated with increased government revenues, which are positively correlated with the amount of oil extraction taking place in the district (see Figure 4.6). To do so, our dataset allows estimates using accurate project-related investment (broken down by capital expenditure and operating expenditure), rather than simply the presence of a project or proxying for the size of projects by the amount of oil extracted (see Figure 4.7). Since we are interested in spillovers from project spending, this is likely to be our preferred measure, even when correlation with petroleum output is modest.

Using our panel data we can control for district-level fixed effects to handle sources of unobserved local time-invariant characteristics, which might potentially lead to endogeneity in the presence of oil related investments. This allows us to control for between-district variation and identify the impact of oil investments based on temporal variations in size and presence of these investments.

We investigate the effect of oil-related investments, and oil-related revenue windfalls, on economic output at the district level. We can express the relationship between oil and economic output as:

$$Y_{jt} = \gamma_1 Oil_{it} + \gamma_2 G_{it} + \mathbf{X}'_{it}\delta_1 + \alpha_t + \beta_j + \varepsilon_{jt}. \tag{4.6}$$

Our panel dataset allows us to exploit the time-variation in our oil and oil revenue windfall data. This allows us to rewrite the static specification in terms of short-run dynamics, and enables us to estimate the multiplier effects of oil investments and fiscal windfalls on economic output:

$$\frac{Y_{it} - Y_{it-1}}{Y_{it-1}} = \gamma_3 \frac{Oil_{it} - Oil_{it-1}}{Y_{it-1}} + \gamma_4 \frac{G_{it} - G_{it-1}}{Y_{it-1}} + \mathbf{X}'_{it}\delta_1 + \alpha_t + \xi_{it} \tag{4.7}$$

where Y_{it}, Oil_{it}, G_{it} and $\mathbf{X}'_{it}$ follow the same definitions as in (4.6). The district fixed effects are subtracted out. The key parameters of interest are now γ_3 and γ_4 which capture the impact of project investments and the government spending multiplier respectively. This specification allows us to estimate the change in output associated with oil-extraction, as a measure of the short-run cyclical impact of spending (see a similar approach employed by Kraay (2012)).

Table 4.11 presents estimates using a variety of treatment variables, namely onshore production, as well as capital and operating expenditures. These estimates are direct estimates of our channels of effect. Columns 1 shows the positive multiplier associated with districts seeing increased value of oil production during the sample period, on local GDP. In contrast, estimates in columns 2 suggest no clear direct effect of capital expenditures, but a positive effect from operating expenditures on district GDP. These estimates suggest a stronger home district effect of operating expenditure compared to our estimates for neighbouring districts, as one might expect if these effects fall with increasing distance.

This table should be interpreted with caution. First, these should be considered associations between the variables of interest rather than causal estimates. Second, the estimates do not isolate the investment or production value associated with specific districts using proximity or fiscal sharing rule formulae. It is therefore likely that these estimates are subject to correlation with important unobverables compared to our main instrumental variable results presented in section 4.5. This table is therefore intended to provide reference estimates to contrast with our identification strategy and main results. Furthermore, these present measures of the effects of oil output on district GDP, of which some portion is constituted from oil output. The is the same challenge documented by Caselli and Michaels (2013). Here the onshore oil output is measured in terms of gross value, while GDP calculates value addition. Like Caselli and Michaels (2013) it is possible out estimates merely reflect this value-addition rather than any spillover effect. The magnitudes of onshore production effect of roughly $\hat{\beta} = 0.3$ is comparable to the 0.4 observed for Brazil.

Table 4.11: Panel results

Reduced form estimates:

	(1) Log GDP	(2) Log GDP
Log onshore production value	0.281***	
	(0.015)	
Log capital exp onshore		0.089
		(0.043)
Log operating exp onshore		0.144**
		(0.034)
Controls	Yes	Yes
Year dummies	Yes	Yes
Province F.E.	Yes	Yes
District F.E.	Yes	Yes
Obs	837	837
F-Test	0.00	0.00
Underid. test p		
Adj. R-squared	.167	.158

Controls: Mining revenue, city dummy, Special autonomous and other province dummies, year dummies and fixed effects. Clustered (at district level) robust standard errors in parentheses. * p¡0.10, ** p¡0.05, *** p¡0.01

4.6.4.1 Effects on sector-specific GDP

We can use sector breakdown of district GDP to examine the channels through which these effects may be operating. The instrumental variable to isolate government revenues derived from offshore oil production can be applied to our sectoral measures of GDP. Table 4.12 presents estimates for the effect of these increased revenues on the service and manufacturing sector GDP. The estimates do not suggest any effect is detected that differs significantly from zero via government spending. In contrast, where we use direct measures of oil sector activity such as oil output and sector investment, while controlling for government revenues, we find two effects. For services, our first-difference estimator suggests a positive relationship between oil sector investments and sector GDP ($\hat{\beta} = 0.013$). Any effect on the manufacturing sector cannot be distinguished from zero. In contrast, for oil output, the manufacturing sector sees a small multiplier of -0.004 in GDP for every 1 percent increase in the value of output. This is tentative evidence for a form of regional Dutch disease (Corden and Neary, 1982), whereby oil sector activity appears to somewhat crowd out manufacturing (considered a relatively traded sector) while inducing growth in the service (considered a relatively non-traded) sector.

These estimations should be interpreted cautiously. First, we have no well-identified specification that isolates the channel of effect. Our instrumental variable estimation yields no measure significantly different from zero for government revenues on specific sector output. In contrast, we find some evidence for effects via the oil sector directly, but rely on the validity of our difference in difference strategy for interpretation as causal estimates. Furthermore these estimates are necessarily short-run and severely limited by data availability. Since we only have data on sectoral breakdown of GDP for years 2005 and 2006 we rely on these two periods for our levels and changes estimations.[23]

4.6.4.2 Estimating the neighbour spillover from gas versus that from oil

So far we have treated oil and gas equally, yet Indonesia during the 2000s provides a fascinating setting to compare the effects from oil and gas production. Both commodities see significant year-on-year price rises throughout the decade, contributing millions of additional dollars in company profits, capital expenditure, and local government revenues. However, the location and timing of oil and gas production varies significantly across the islands of Indonesia and our

[23] Such findings warrant further investigation and we are in the process of procuring additional years of sectoral breakdown of GDP data that would allow us to build a more robust picture of this relationship.

Table 4.12: Channels

Instrumental variable estimation:				
	(1)	(2)	(3)	(4)
	OLS Log services	IV Log services	OLS Log manu.	IV Log manu.
Log petroleum output	-0.004		-0.004*	
	(0.001)		(0.001)	
Log total sector investment	0.013***		0.006	
	(0.001)		(0.004)	
Log local govt revenue	0.007	-0.057	0.001	0.082
	(0.017)	(0.310)	(0.009)	(0.327)
Controls	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
District F.E.		F.E.		F.E.
Obs	895	892	893	890
F-Test	1.15e-94	.	4.88e-91	.
Underid. test p		0.58		0.59
Adj. R-squared	0.62	0.20	0.58	-0.13

Controls: Mining revenue, city dummy, Special autonomous and other province dummies, year dummies. Clustered (at district level) robust standard errors in parentheses. * p¡0.10, ** p¡0.05, *** p¡0.01

priors suggest that the extent of linkages to the regional economy also differ. We can exploit the variation across time and space to separately estimate the spillover effects of oil and gas.

Our results, presented in table 4.13 suggest a stronger (and positive) effect from gas sector activity versus oil sector activity on neighbouring districts. This is consistent with our prior that gas extraction and development tends to be biased towards greater regional consumption (due to higher transportation costs). Even where export is the preferred option, gas requires the construction of large and expensive liquification terminals to prepare natural gas for export. Oil can be exported more readily via shipping terminals (or via truck or rail movement).

4.7 Conclusions

In this paper we examine the relationship between an oil and gas revenues and local economic performance. The main research question looks at the effect of local government spending increases due to oil windfalls. In order to estimate this effect we isolate the direct (investment)

Table 4.13: Oil versus gas effects in neighbour districts, 99 and 09

	(1) Log GDP
Neighbour oil output	-0.007
	(0.011)
Neighbour gas output	0.103*
	(0.013)
Controls	Yes
Year dummies	Yes
Province F.E.	Yes
District F.E.	Yes
Obs	728
F-Test	0.00
Adj. R-squared	0.85

Controls: Mining revenue, city dummy, Special autonomous and other province dummies, year dummies. Clustered (at district level) robust standard errors in parentheses. * p¡0.10, ** p¡0.05, *** p¡0.01

channel from the indirect (government spending) channel, which we do so using an instrumental variable for those revenue windfalls arising from offshore oil production. Our results point to some key findings.

Petroleum production is associated with increases in district GDP. The main economic gains are estimated to accrue via transfers to local government having a multiplier effect of around 0.9 for every unit of additional revenue. In contrast, oil sector output appear to have a small though positive effect - estimated at closer to 0.3, but in line with the literature. However, to identify project spillovers separately from revenue effects we focus on districts adjacent to petroleum production. Here we find an attenuated effect. For every dollar of petroleum output in producing districts, neighbouring districts experience no discernable increase in economic output during our sample period. There is evidence that project expenditures on operating costs may contribute positively (an effect of roughly 0.3) while capital expenditures may operate in reverse (-0.2).

The paper contributes to our understanding of the ways resource extraction effects the subnational economy. First, extraction need not be viewed as a purely enclave activity. In particular, by producing estimates for revenue windfall effects in offshore petroleum-producing districts. Second, we show that project investment spillovers to neighbouring districts can be positive and significant in the short-run. Third, consistent with the wider literature on the Dutch disease we also find a modest crowding out of manufacturing in favour of a boom in service sector output.

The approach follows closely the analysis of Brazilian oil windfalls by Caselli and Michaels (2013). We extend this analysis to Indonesia, mimicking their instrumental variable technique to isolate measures of the impact of government revenue windfalls. In contrast to Caselli and Michaels (2013), for Indonesia we find evidence for positive and significant effects of government revenue windfalls on economic output across the sample period. Furthermore, this result is fairly large- implying a fiscal multiplier around 0.9 over the decade. We then move beyond the fiscal channel to attempt to identify project spillovers, in particular those accruing in districts neighbouring petroleum producers. This allows us to observe modest effects via project expenditures on capital versus operating expenditures. We are able to separate estimates for the spillover effects for oil and gas. While estimates for oil cannot be distinguished from zero, the gas spillover effect to neighbouring districts is estimated to be positive and significant, supporting the idea that potential economic linkages from gas exceed those from oil production.

Interpretation of our results, and particularly magnitudes depends crucially on the estimates of GDP. Where we estimate the impact on our full sample using measures of district-level GDP, we capture both the within-sector and non-oil sector impacts. This includes the net export revenues accruing to the oil sector associated with onshore extraction activity. As would be expected increased oil production and increased oil project investment raise the level of overall GDP in the district. This implies increases in the size of the oil sector, and the less than complete crowding-out of other activities.

Our analysis at present tells us little about how the spillover benefits accrue within the districts and across the oil and non-oil sectors, nor how fiscal windfalls drive spending choices. This approach can thus be extended to look at government expenditure data and spending patterns to try to explore exactly how this result is emerging. All these insights are important once we look beyond subnational income and economic growth towards the impact of local government income and spending behavior on the subnational-level human development measures such those concerning health and education.

4.8 Indonesia oil and gas sector background

The natural resource sector, together with geographic and ethnic factors, was among the main drivers of the decentralization process. It includes oil and gas, coal and metals mining, as well as forestry, fishery and agricultural production, and typically makes up more than one quarter of the country's GDP and half of the country's exports.

The oil, gas and mining industries are economically and politically significant, since they contribute three quarters of the natural resources GDP and two-thirds of natural resource sector exports. The associated industries play a special role in the political and fiscal decentralization dynamics of the country. In fact, resource-rich regions have historically been rather vocal about their ability and desire to gain more administrative and fiscal autonomy.

In Indonesia, natural resources are considered national assets and controlled by the government. The oil and gas industry was governed by a series of laws for more than four decades. Among them were Law 44 of 1960 on oil and gas exploration and production activities, Laws 2 and 15 of 1962 on domestic oil obligations, and Law 8 of 1971 regarding the state-owned oil and gas company Pertamina. In 2001, Indonesia's oil and gas Law 22 superseded a significant part of the historical legal regime in the industry.

Oil and gas Law 22 defines upstream business activity as exploration and exploitation, which include exploration of potential oil and gas reserves, drilling wells, the construction of transportation, storage and processing facilities and the processing of natural gas into liquefied natural gas. Indonesian and foreign companies may engage in upstream business activity as long as they fulfil financial, regulatory, technical and operational requisites. Law 22 established an upstream oil and gas sector regulatory body (*BP Migas*) to replace the national oil company's (Pertamina's) historical regulatory and supervisory function, as well as its role as government representative countersigning production sharing contracts. However *BP Migas*, the upstream regulatory body, was recently integrated back into the Ministry of Mines and Energy (MEMR), under the directorate general of oil and gas, and renamed *SKK Migas*. In contrast to the more decentralized mining industry regulation, the role of BP Migas in granting exploration and development licenses, counter-signing cooperation contracts and monitoring project expenditures remains largely under central government control.

Chapter 5

Conclusions

James Cust

5.1 Future research directions and research agenda

The work presented in this thesis raises perhaps more questions than it answers. These may constitute a starting point for several potentially fruitful avenues of investigation. The variety of new datasets and an explosion in new studies taking advantage of better identification strategies employing spatially identified data demonstrate the interest in this field. Furthermore, the past decade's commodity price boom has reinforced the need by policy-makers to understand the opportunities and threats presented by resource wealth, in addition to the policy tools they might have at their disposal to affect the outcomes.

5.1.1 Oil investment and governance

Chapter 2 could provide us with a new approach to estimating the magnitude of how much institutions matter for inward investment; and separating the estimate of the effect of institutions from other candidates for fundamental causes of economic development such as geography or culture. In the first instance this yields estimates for oil exploration, but potentially could be shown to have wider external validity. Beyond my paper co-authored with Torfinn Harding, these are several directions I would like to extend this work:

Unpacking the institutional black box: what drives exploration location? While Chapter 2 estimates the effect of broadly measured institutions (proxied by Freedom House and Polity IV measures of democratic rights and constraints on the executive) we expect that different components of good governance matter more for oil investors than others. Building on Acemoglu and Johnson (2005) I propose to investigate the role of different components of institutions, for example to compare the importance of property rights versus contracting institutions. Is it property rights institutions related to executive constraints and protection against expropriation that really matters? In contrast, is state capacity important, and in what ways? The identification strategy developed for Chapter 2 provides a potentially powerful setting to compare these different institutional configurations.

Furthermore, this approach could examine other competing forces beyond governance alone. Could fiscal incentives such as low royalty rates or the offer of tax holidays tip the balance in favour of investment, as presumed by many policy makers? The provision of public infrastructure, or levels of corruption might also matter. Furthermore, the supply of licenses might be an important channel for the headline result, as opposed to say, weak governance deterring investors on the demand side.

Endogenously determined resource wealth: The second avenue is to examine and potentially address endogeneity in resource exploration and therefore discovery; an issue that strikes at the heart of most empirical papers on the resource curse. If oil exploration and discovery is endogenous to institutions, then oil abundance is endogenous to institutions. If we treat observed resource abundance as exogenously given, studies are likely to be biased. This also means, amongst other things, that the returns to improving institutions may be even higher then we have previously thought. Unfortunately many measures of resource abundance are likely to be subject to a range of biases.

One of the challenges can arises through selection and timing of discovery. As discussed in Chapter 2, we find evidence that the exploration process is driven by both the quest for promising geology and the institutional environment prospecting takes place. At the margin, investors will drill on the better governed side of the border two times out of three. While the magnitude of this effect has decreased over the past four decades, there is no evidence that we are at the end of history in terms of the exploration process. Indeed the past decade has been characterised a series of major new petroleum discoveries such as offshore oil in Ghana, offshore gas in Mozambique and Tanzania and the pre-salt layer in Brazil. Furthermore our findings suggest the probability of success does not vary as one crosses the border.

Where we worry that the exploration process is subject to decisions about preferred countries to drill in, or actions by host government, similarly the timing of exploration and hence discovery is likely to be subject to supply and demand effects. In Chapter 2 we present evidence to suggest that the institutional setting can delay drilling than the counterfactual where timing choice is based on geology alone. While new datasets such as the giant oilfields offer potentially exogenous sources of timing of discovery variation, the prize remains to find a more broadly applicable measure or instrument for exogenously determined resource wealth.

5.1.2 Resource extraction and development

Many resource-rich countries face pressure to decentralize spending and distribute the proceeds of resource extraction disproportionately in favor of resource rich regions. It is therefore of great relevance to study the cases where developing countries have already undertaken local revenue sharing. The literature suggests they have had mixed success. Early evidence suggests that local government spending, in particular where the capacity of local government is limited, can crowd out private sector activity through a burgeoning public sector. More research on how sharing rules and spending policies should best be designed and administered, in addition to the capacity constraints and pre-requisites for effective administration could provide significant benefits for countries following this path.

Indonesia offers a promising setting to examine these questions in more detail. Looking at district government capacity and expenditures, as well as a broader measure of outcomes could hele shed light on the heterogeneity in effects and to decompose the channel for the effects identified in Chapter 4. Furthermore this may allow us to reconcile these results with equivalent though contrasting findings in Brazil.

5.1.3 Data

To support the new methodological approaches being applied in the resource literature, new datasets have been sought. These new datasets typically seek to combine information on resource extraction, or fiscal information associated with resources, with other socio-economic variables.

For petroleum a range of data sources are now being used. The first example are regional statistics, either reporting extraction activities like volumes or investments at the state or district level, and fiscal information such as government revenues and expenditures associated with resource wealth. This information is typically sourced from national governments, for example in Brazil by Caselli and Michaels (2013), or in Canada by Papyrakis and Raveh (2014). Such data can be costly to assemble. It may need conversion from paper or pdf format, translation, and its acquisition may be preceded by negotiation with government agencies and officials to obtain access for research. The global trend toward public disclosure of petroleum information, and in a digital format, is increasing but remains incomplete.

The second type of petroleum data is micro-data at the level of companies, project or even wellhead. Examples include the Horn (2004) dataset of giant oilfields, which includes information on location, historic production and discovery year, used by Lei and Michaels (2011). PRIO produce the Petrodata dataset for researchers (Lujala et al., 2007) which contains extensive field-level coverage for onshore and offshore activity. Less widely used a commercial data sources such as the Wood Mackenzie PathFinder database (Wood Mackenzie, 2011) (used for example in Chapter 2), which contains rich and close to comprehensive global data on exploration wells and field-level production. The Wood Mackenzie PathFinder database contains over 120,000 exploration wells and extensive data on field level production and investment. It represents the most extensive global coverage and accuracy used to date, as well as being regularly updated, access however remains limited to mostly commercial subscribers.

For solid minerals the sources of data are equally diverse. Aggregate government statistics at the subnational level are used in countries including for the US coal industry (Black et al., 2005). Micro-data at the company or mine-level includes the USGS database for the world, the Raw Materials Database (InterraRMG, 2013), and the PRIO dataset on diamonds (Gilmore et al., 2005). These data sources typically include information on mine location as GPS coordinates, historic production volumes and status, and levels of reserves. Challenges associated with these data are significant; the USGS dataset contains information on ore and geology, with very limited and patchy coverage on actual mining activity. Furthermore, the data quality can be low - the USGS dataset for example categorizes mines simply as small, medium, or large without defining the categories. Indeed these datasets share a sample selection process that is unclear and undocumented, introducing potential bias that may affect empirical work. The RMD is more comprehensive in coverage of mine projects, including production volumes and status, however coverage is patchy and location data can be very inaccurate. Since the primary function of the RMD has been as a database of ownership, its use for economic analysis at the mine level can be challenging. It does not, for example, have systematic data on inputs, costs, investment, prices, labor or capital. Researchers have typically relied on mine status, volume or inferred revenue for analysis. Estimating the rents however would require cost data that is typically absent. Other sources such as the SNL Minerals database, who recently acquired InterraRMG, offer a much larger and improved database for commercial subscribers. They claim this is a significant improvement in accuracy over the RMD dataset.[1]

Many studies use world price shocks, exogenous changes in policy, or simply changes in production or revenues in a panel data setting for identification. Endogeneity remains a perennial problem in micro-econometric studies as it has for cross-country analysis. In particular any extraction activity is conditional on knowledge of where resource deposits reside, and new research presented in this thesis suggests that the exploration process, at least in the petroleum sector, may be subject to strong effects from institutional quality (Chapter 2). If governance indeed plays a critical role in choice of exploration location, any inference based on the presence of extraction activity may also be subject to these same effects. The Horn (2004) database

[1] http://www.snl.com/Sectors/MetalsMining/Default.aspx

has the advantage over many of the other resource datasets survey, via its appeal to the size of oilfields used. By including only giant oilfields those exceeding ultimately recoverable reserves of 500 million barrels whose global economic importance is clear, authors such as Lei and Michaels (2011) argue that their discovery is a plausibly exogenous event compared to specific country characteristics. However this too may suffer from the influence of institutions and technology on choice of exploration location and therefore discovery location and timing.

The data landscape has changed dramatically in the last decade making studies of the sub-national effects of resource wealth possible. However the data remains patchy and few datasets used have been prepared and collected with researchers in mind. As the research agenda develops further, there will be important work to collate resource data more systematically, with better documentation and review to ensure quality and coherence. The global trend towards increased public disclosure of information about the mining and petroleum sectors, started by the Extractive Industries Transparency Initiative (EITI) in 2003, continues to gather pace. EITI now makes available 235 reports, for over 35 countries, with a total of 48 countries signed up to join the initiative. Reporting involves submitting detailed information on company payments made and received for oil, gas and mineral resources, and will soon include information disaggregated to individual projects.

Further there exists large amounts of extractive sector data available in the public domain though rarely collated into a useful format for research. Examples include company annual reports (the basis of many commercial databases) as well as corporate filings for publicly listed companies or those seeking project financing. New legislation such as the US Dodd-Frank Act 2010 include provisions for enhanced corporate filings by oil, gas and mining companies. Amendment 1504 when it comes into force (sometime in the next couple of years) will require project level disclosure of all tax and royalty payments by US-listed entities. The EU, Canada and Norway already have similar legislation to come into force next year. Furthermore new efforts like JODI- the Joint Oil Data Initiative- seeks to provide more comprehensive and timely petroleum data.[2] All of this adds up to a promising landscape for future research.

[2]JODI is a collaboration between six international organisations - APEC, Eurostat, IEA, OLADE, OPEC and the UNSD, and their member countries. More information can be found at https://www.jodidata.org/ There are

now over 70 participating countries, representing 90% of global oil supply and demand.

Bibliography

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*, Volume 1 of *Handbook of Economic Growth*, Chapter 6, pp. 385–472. 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, 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.

Allcott, H. and D. Keniston (2013). Dutch disease or agglomeration? the local economic effects of natural resource booms in modern America. *Preliminary First Draft*.

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.

Angrist, J. D. and A. D. Kugler (2008). Rural windfall or a new resource curse? Coca, income, and civil conflict in Colombia. *The Review of Economics and Statistics 90*(2), 191–215.

Aragon, F. and J. P. Rud (2012). Mining, pollution and agricultural productivity: Evidence from Ghana.

Aragon, F. M. and J. P. Rud (2013). Natural resources and local communities: Evidence from a Peruvian gold mine. *American Economic Journal: Economic Policy 5*(2), 1–25.

Arellano-Yanguas, J. (2011). Aggravating the resource curse: decentralisation, mining and conflict in Peru. *The Journal of Development Studies 47*(4), 617–638.

Aroca, P. and M. Atienza (2011). Economic implications of long distance commuting in the Chilean mining industry. *Resources Policy 36*(3), 196–203.

Arzaghi, M. and J. V. Henderson (2005). Why countries are fiscally decentralizing. *Journal of Public Economics 89*(7), 1157–1189.

Auty, R. M. (2001). *Resource abundance and economic development*. Oxford University Press.

Azis, I. J. (2008). Institutional constraints and multiple equilibria in decentralization. *Review of Urban & Regional Development Studies 20*(1), 22–33.

Banerjee, A., E. Duflo, and N. Qian (2012). On the road: Access to transportation infrastructure and economic growth in China. Technical report, National Bureau of Economic Research.

Barr, C. M. (2006). *Decentralization of Forest Administration in Indonesia: Implications for Forest Sustainability, Economic Development, and Community Livelihoods*. CIFOR.

Bartik, T. J. (2003). Local economic development policies.

Beine, M., S. Coulombe, and W. N. Vermeulen (2012). Dutch disease and the mitigation effect of migration: Evidence from Canadian provinces. Technical report, CESifo working paper: Resource and Environment Economics.

Berman, N., M. Couttenier, D. Rohner, and M. Thoenig (2014). This Mine is Mine! How minerals fuel conflicts in Africa.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

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

Black, D., T. McKinnish, and S. Sanders (2005). The economic impact of the coal boom and bust. *The Economic Journal 115*(503), 449–476.

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

Brata, A. G. (2008). Creating new regions, improving regional welfare equality? Technical report, University Library of Munich, Germany.

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

Carrington, W. J. (1996). The Alaskan labor market during the pipeline era. *Journal of Political Economy 104*(1), 186–218.

Caselli, F. and G. Michaels (2013). Do oil windfalls improve living standards? Evidence from Brazil. *American Economic Journal: Applied Economics 5*(1), 208–38.

Caselli, F., M. Morelli, and D. Rohner (2012). The geography of inter-state resource wars. Working paper, London School of Economics.

Chen, X. and W. D. Nordhaus (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences 108*(21), 8589–8594.

Chinitz, B. (1961). Contrasts in agglomeration: New York and Pittsburgh. *The American Economic Review*, 279–289.

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 B. Goderis (2009). Commodity prices, growth, and the natural resource curse: reconciling a conundrum. *Growth, and the Natural Resource Curse: Reconciling a Conundrum (June 5, 2008)*.

Collier, P. and A. Hoeffler (2005). Democracy and resource rents. *Department of Economics, University of Oxford*.

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.

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

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.

Dal Bó, E. and P. Dal Bó (2011). Workers, warriors, and criminals: Social conflict in general equilibrium. *Journal of the European Economic Association 9*(4), 646–677.

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

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

de Haas, R. and S. Poelhekke (2014). Business constraints in mining regions. *Working Paper*.

Deaton, A. (1997). *The analysis of household surveys: A microeconometric approach to development policy*. World Bank Publications.

Deichmann, U., K. Kaiser, S. V. Lall, and Z. Shalizi (2005). Agglomeration, transport, and regional development in Indonesia.

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

Dinkelman, T. (2011). The effects of rural electrification on employment: New evidence from South Africa. *The American Economic Review*, 3078–3108.

Donaldson, D. (2010). Railroads of the Raj: Estimating the impact of transportation infrastructure. Technical report, National Bureau of Economic Research.

Dube, O. and J. F. Vargas (2006). *Are All Resources Cursed?: Coffee, Oil and Armed Conflict in Columbia*. Center for International Development at Harvard University.

Duek, A., B. Brodjonegoro, and R. Rusli (2010). Reinterpreting social processes: How system theory can help to understand organizations and the example of Indonesia's decentralization. *Emergence: Complexity & Organization 12*(4).

Duek, A. and R. Rusli (2010). The natural resources industry in decentralised Indonesia: How has decentralisation impacted the mining, oil and gas industries? Technical report, Center for Research in Economic Analysis, University of Luxembourg.

Duflo, E. and R. Pande (2007). Dams. *The Quarterly Journal of Economics 122*(2), 601–646.

EIA (2013, September). Country analysis: Sudan and South Sudan.

Fafchamps, M. and F. Shilpi (2003). The spatial division of labour in Nepal. *The Journal of Development Studies 39*(6), 23–66.

Fitrani, F., B. Hofman, and K. Kaiser (2005). Unity in diversity? The creation of new local governments in a decentralising Indonesia. *Bulletin of Indonesian Economic Studies 41*(1), 57–79.

Ford, J. F. G. and G. Fitz (2000). Inter-governmental fiscal relations and state building: The case of Indonesia. *Washington, DC: Mimeograph.*

Foundation, A. (2008). Indonesia local economic governance index, 2007.

Gelb, A. H. (1988). *Oil windfalls: Blessing or curse?* Oxford University Press.

Gilmore, E., N. P. Gleditsch, P. Lujala, and J. K. Rood (2005). Conflict diamonds: A new dataset. *Conflict Management and Peace Science 22*(3), 257–272.

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.

Glaeser, E. L., S. P. Kerr, and W. R. Kerr (2012). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics* (0).

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, 2). 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, February). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics 114*(1), 83–116.

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.

Hirschman, A. (1958). The strategy of economic development.

Hofman, B. et al. (2006). *Evaluating fiscal equalization in Indonesia*, Volume 3911. World Bank Publications.

Hooker, M. A. and M. M. Knetter (2001). Measuring the economic effects of military base closures. *Economic Inquiry 39*(4), 583–598.

Horn, M. K. (2004). Giant oil fields 1868-2003, data on a CD-ROM. *Giant oil and gas fields of the decade 1999*, 2000–2006.

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, September). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

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.

Jarvis, M., G.-M. Lange, K. Hamilton, D. Desai, B. Fraumeni, B. Edens, S. Ferreira, H. Li, L. Chakraborti, W. Kingsmill, et al. (2011). *The changing wealth of nations: Measuring sustainable development in the new millennium.*

Jeanty, P. W. (2012). Nearstat: Stata module to calculate distance-based variables and export distance matrix to text file. *Statistical Software Components*.

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.

Kirchberger, M. (2010). Distributional effects of natural disasters: Who are the winners and losers? Technical report, working paper.

Kotsadam, A. and A. Tolonen (2013). Mineral mining and female employment. *Oxcarre Research Papers 114*.

Kraay, A. (2012). Government spending multipliers in developing countries: Evidence from lending by official creditors.

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).

Lei, Y.-H. and G. Michaels (2011). Do giant oilfield discoveries fuel internal armed conflicts? *CEPR Discussion paper series* (6934).

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

Libecap, G. D. (1989). The political economy of crude oil cartelization in the United States, 1933–1972. *The Journal of Economic History 49*(04), 833–855.

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.

Libman, A. (2013). Natural resources and sub-national economic performance: Does sub-national democracy matter? *Energy Economics 37*, 82–99.

Lippert, A. (2014). Spill-overs of a resource boom: Evidence from Zambian copper mines. Technical report, Oxford Centre for the Analysis of Resource Rich Economies, University of Oxford.

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.

Loayza, N., A. Mier y Teran, and J. Rigolini (2013). Poverty, inequality, and the local natural resource curse. Technical report, Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit.

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

Lowe, M. (2014). Rail revival in africa? the impact of privatization. *New Perspectives*.

Lujala, P., J. K. Rood, and N. Thieme (2007). Fighting over oil: Introducing a new dataset. *Conflict Management and Peace Science 24*(3), 239–256.

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

McCulloch, N. and E. Malesky (2011). Does better local governance improve district growth performance in Indonesia? *IDS Working Papers 2011*(369), 1–48.

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.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics 90*(4), 683–701.

Michaels, G. (2011). The long term consequences of resource-based specialization. *The Economic Journal 121*(551), 31–57.

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.

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.

Monteiro, J. and C. Ferraz (2010). Does oil make leaders unaccountable? Evidence from Brazil's offshore oil boom. *unpublished, PUC-Rio*.

Moretti, E. (2010). Local multipliers. *The American Economic Review*, 373–377.

Moretti, E. (2011). Local labor markets. *Handbook of labor economics 4*, 1237–1313.

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.

Notowidigdo, M. J. (2010). Essays on the economics of local labor markets.

Oates, W. E. (1972). Fiscal federalism.

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

Papyrakis, E. and R. Gerlagh (2004). The resource curse hypothesis and its transmission channels. *Journal of Comparative Economics 32*(1), 181–193.

Papyrakis, E. and O. Raveh (2014). An empirical analysis of a regional Dutch disease: The case of Canada. *Environmental and Resource Economics 58*(2), 179–198.

Perez-Sebastian, F. and O. Raveh (2012). The natural resource curse, fiscal decentralization, and agglomeration economies. Technical report, Mimeo.

Perry, G. and M. Olivera (2009). Natural resources, institutions and economic performance. *Documento de trabajo, noviembre 15*.

Pesaran, M. H. (1990). An econometric analysis of exploration and extraction of oil in the U.K. Continental Shelf. *The Economic Journal 100*(401), pp. 367–390.

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

Raleigh, C., A. Linke, H. Hegre, and J. Karlsen (2010). Introducing ACLED: An armed conflict location and event dataset special data feature. *Journal of peace research 47*(5), 651–660.

Raveh, O. (2013). Dutch disease, factor mobility, and the Alberta Effect: The case of federations. *Canadian Journal of Economics/Revue canadienne d'économique 46*(4), 1317–1350.

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

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

Sachs, J. D. and A. M. Warner (1995). Natural resource abundance and economic growth. Technical report, National Bureau of Economic Research.

Smith, A. (1937). *The Wealth Of Nations [1776].* na.

Stigler, G. (1957). The tenable range of functions of local government.

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, January). Foreign direct investment and the risk of expropriation. *Review of Economic Studies 61*(1), 81–108.

Tiebout, C. M. (1956). A pure theory of local expenditures. *The Journal of Political Economy*, 416–424.

Toews, G. (2013). Inflated expectations and natural resource booms: Evidence from Kazakhstan. *OxCarre Research Paper 109*.

Toews, G., A. F, and J. Rud (2014). What happens when a mine closes? The case of coal mines in UK. *mimeo*.

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.

United Nations (1958). United Nations Convention on the Territorial Sea and the Contiguous Zone. *United Nations Treaty Series*.

United Nations (1982). United Nations Convention on the Law of the Sea. *United Nations Treaty Series 1833*, 3.

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 (2009). Volatility and the natural resource curse. *Oxford economic papers*, gpp027.

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

Van Wijnbergen, S. (1984). The Dutch Disease: A disease after all? *The Economic Journal*, 41–55.

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.

Vermeulen, W. (2013). Resource income and the effect on domestic neighbours.

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.

Zhang, T. and H.-f. Zou (1998). Fiscal decentralization, public spending, and economic growth in China. *Journal of public economics 67*(2), 221–240.