STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2019–2020

The Local Effect of Oil on Women's Employment and Empowerment - Evidence from Africa

Nadja Friedl (41346) and Sailee Sakhardande (41430)

Abstract

Does oil extraction impact local communities in developing countries? The socio-economic ramifications of oil extraction on the local population are not clear a priori. Recent evidence pointing towards a gender-differential impact of natural resources motivates an analysis focused on women. This paper examines the local effect of oil extraction on women's employment and empowerment in Africa. Using a novel data set, we exploit spatial and temporal variation in oilfield opening to estimate the local effect of oil extraction. We match precise oilfield data from Rystad Energy with geo-coded DHS survey data for six African oil-producing countries. Employing difference-in-difference estimation, we compare women close to and far from oilfields, before and after extraction starts. Our main results are three-fold: First, we find that the local industrial development stemming from oil extraction increases women's probability of service sector employment by 18%. This effect is robust to excluding migrants. We hypothesise that the economic benefits of oil extraction proliferate through local multiplier effects (spillovers). Second, there is no change in women's empowerment, as measured in terms of decision-making power and self-stated barriers to healthcare. Third, our results indeed capture a highly localised effect, as the positive impact on service sector employment vanishes with distance.

Keywords: Local Impact of Natural Resources, Oil, Employment, Female Empowerment, Africa
JEL: J16, N57, R11, O13

Supervisor: Elena Paltseva
Date submitted: 09 December 2019
Date examined: 16 December 2019
Discussants: Lotta Adeborg and Sebastian Ekberg
Examiner: Magnus Johannesson

Acknowledgements

We would like to express our sincere gratitude to our supervisor, Assistant Professor Elena Paltseva, for her time, valuable guidance and insightful discussions. We would like to acknowledge that this thesis was funded by *The Research Council of Norway*, Grant 275387, for which we are very thankful. Finally, we are grateful for valuable advice and suggestions provided by Assistant Professor Abhijeet Singh, Roman Bobilev, Felix Schafmeister, Martina Dosser and Johannes Matt.

Contents

List of Tables

List of Figures

1 Introduction

Does oil extraction impact local communities in developing countries? While resource rents are generally collected and distributed at the national level, there are likely to be distinct effects of oil extraction on local communities. People living in resource-producing areas may experience changes in economic, social and environmental outcomes. A priori, the socio-economic ramifications of oil extraction on the local population are not clear. While economic benefits in terms of increased employment opportunities may arise, local ecosystems (e.g. water, air and soil) are likely to be adversely affected, which may in turn impact health and employment.

Recent evidence points towards a gender-specific impact of extractive industries at both the macro and micro level (Benshaul-Tolonen and Baum, 2019). The overall impact of resource extraction on women is described as a mixed blessing. Women may benefit from increased employment, though not necessarily in the same way as men. Resource extraction may even exacerbate gender inequality in employment, as men are more likely to gain direct employment in extractive industries. In contrast, employment gains for women are expected to mainly arise in non-resource sectors. In the context of large-scale mining in Africa, women are found to benefit mainly from service sector employment, while males gain manual employment (Kotsadam and Tolonen, 2016). Additionally, women may also be disproportionately put at risk by the social, economic and environmental consequences of resource extraction (World Bank, 2013). Thus, the gender-differential impact of extractive industries motivates a study focused on women.

The main research question of this paper is to explore the local effect of oil extraction on women's employment in Africa. The key channel through which we expect oil extraction to impact local employment is that of linkages and multipliers. Direct employment in the capital-intensive and highly globalised oil industry is generally low for women (Fajana, 2005). However, indirect employment effects may be generated through local multiplier effects (spillovers) to non-resource sectors of the economy. Additionally, we test whether there is a local impact on women's empowerment, which for the purpose of this analysis is narrowly defined as decision-making power within the household and agency over one's own health. As it has been argued that "petroleum perpetuates patriarchy" (Ross, 2008, p.120), women's empowerment is an important outcome to explore in the context of oil. Further, women's employment and empowerment are interrelated. For example, benefits from greater labour market participation may result in increased bargaining power, autonomy and delayed marriage and childbearing for women (Heath and Mobarak, 2012; Jensen, 2012). At the same time, higher bargaining power in the household may also result in greater female labour force participation (Basu, 2006).

To the best of our knowledge, our paper is the first to explore the local effect of oil extraction on women's employment and empowerment in Africa. Using novel data on oil extraction from Rystad Energy, we compile and analyse an innovative dataset that includes the most recent and non-publicly available information on oilfields'[1] precise location and start-up date. We match the oilfield data with geo-coded survey data on women's employment and empowerment outcomes from the Demographic and Health Surveys (DHS). Exploiting spatial and temporal variation in the opening of oilfields, we estimate the local effect using a difference-in-difference design. We divide our sample into a treatment group – defined as women living within a 15 km radius of an

[1] Unless stated otherwise, oilfields in this paper refer to conventional onshore oilfields.

oil deposit[2] – and a control group, capturing women living between 15 and 100 km from an oil deposit. Our sample includes all African oil-producing countries for which geo-coded survey data is available; it encompasses about 100,000 women from Angola, Chad, Egypt, Gabon, Madagascar and Nigeria.

Our main results are three-fold. First, we find that the local industrial development stemming from oil extraction increases women's probability of service sector employment by 3.9 percentage points. That is, women living in the vicinity of active oilfields are on average 18% more likely to work in services. This result is robust to excluding migrant women. Second, there is no change in women's empowerment, measured in terms of decision-making power and self-stated barriers to healthcare. Third, our results indeed capture a highly localised effect, as the positive impact on service sector employment vanishes when increasing the treatment radius from 15 to 25 km.

Our paper mainly relates to the growing literature on the local and regional effects of natural resources (for an overview, see Cust and Poelhekke, 2015). Several studies have examined the local impact of mineral resources, such as gold or copper, in the African context (e.g. Aragón and Rud, 2016; Kotsadam and Tolonen, 2016; Lippert, 2014). The global importance of oil as a natural resource raises the question whether the findings of these local impact studies also hold in the context of oil extraction. Generalisation of these studies to oil extraction may not hold due to systematic differences between oil and mineral extraction, for example in terms of capital intensity, infrastructure requirements and land usage. Thus, we contribute to the empirical literature on the local effect of natural resources by examining the local effect of oil extraction on women's employment and empowerment.

The remainder of this paper is structured as follows: Section 2 reviews the previous literature and Section 3 provides background on oil extraction. We then describe our data and summary statistics in Section 4. Section 5 describes the empirical strategy and discusses our study's internal validity. We then present our results in Section 6 and provide robustness checks in Section 7. Section 8 discusses the results, comparing them to similar studies on other natural resources and addresses the external validity of the study. Section 9 concludes.

2 Literature Review

There are three strands of literature that this paper relates to: first, the literature on the local and regional impact of natural resources; second, the literature on the gender-specific impact of natural resources; and third, the literature specific to employment effects of oil. Below we provide a short overview on these three strands.

First and foremost, this paper relates to the economic literature on the local impact of natural resources. This is a relatively new strand of literature - earlier generations of scholars have tended to explore the impact of natural resources on a macroeconomic level. Notably, Auty's (1994) resource curse hypothesis posits that resource dependence is detrimental to long-run economic growth. This hypothesis has been tested in many cross-country analyses, but the empirical evidence remains mixed. For example, while growth in the resource sector is found to decrease manufacturing exports and expand the service sector in some studies (e.g. Harding and Venables, 2010; Sachs and Warner,

[2] Oil deposits refer to oilfields that will turn operational during our period of study.

1995), others find no significant trade-off between the manufacturing sector and the resource-rich sector (e.g. Hutchinson, 1994; Spatafora and Warner, 1999). For a comprehensive review of the empirical literature on the resource curse, see, for example, Frankel (2010) and van der Ploeg (2011).

However, many of the early cross-country studies suffer from weak identification, as differences in countries' resource wealth and dependence are not purely exogenous, but intertwined with factors such as political stability, institutions or economic development. To address the endogeneity and identification concerns, new methodologies have emerged in the field of natural resource economics. Common methodologies, such as instrumental variable approaches or natural experiments (e.g. Vicente, 2010) have been applied, but could not fully resolve identification problems. Van der Ploeg and Poelhekke (2017) survey the existing quantitative evidence on the impact of natural resources, and identify the study of the *local* impact of natural resources as an important methodology that improves identification.

2.1 The local and regional impact of natural resources

Many natural resources are geographically concentrated, which motivates the study of local and regional effects. In the case of oil, prominent examples include the Niger Delta region in Nigeria or Saudi Arabia's Eastern Province. Comparing within-country regions (e.g. provinces, counties, districts) differing in resource wealth reduces the unobserved heterogeneity in terms of national characteristics, such as institutional quality or political stability. Thus, one branch of the local and regional impact literature focuses on a single country, exploiting within-country differences in resources, which improves econometric identification.

Another key advantage of local impact studies is that they can provide detailed insight on region-specific outcomes. These include environmental and health effects or changes in local employment, wealth and income. Surveying the literature on the local impact of natural resources, Cust and Poelhekke (2015) highlight the importance of within-country effects, as they may be important transmission mechanisms for outcomes observed at the country-level. Regions differ in many aspects that are key channels in determining resources' costs and benefits, such as infrastructure, prices, local political institutions, rent seeking and corruption, which motivates micro-level explorations of such factors.

A driving force behind the emergence of local impact studies has been improved data availability. Firstly, large-scale socioeconomic surveys with geographic identifiers, such as the DHS surveys used in this paper, enable to identify individuals or households' exact geographic location. Most commonly, the geographic location is obtained through latitude and longitude GPS coordinates. Secondly, detailed data on natural resources provide variables such as geographic location, discovery year, production status and volumes. Combining this individual- or household-level and natural resource data, many authors have been able to exploit spatial or temporal variation within countries (Cust and Poelhekke, 2015). This is generally seen as having improved causal inference in the field, as many of the endogeneity concerns present in earlier studies - such as differences in institutions, culture or stability - can be mitigated.

One strand of literature examines the impact of individual projects; the most prominent example is the study of the local economic impact of Yanacocha, Peru's largest gold mine (Aragón and Rud, 2013). Utilising a difference-in-difference design, the authors compare households close to the mining area with households further away. Exploiting a change in procurement policies by the mine owner, they find positive effects on the local population's standard of living. A ten percent increase in the gold mine's activity is associated with a 1.7 percent increase in the real income of nearby households. Backward linkages are hypothesised to be the key reason why local populations benefit. However, it is crucial to note that these findings are short-run effects and cannot be interpreted as a long-term impact on income. Next, as the world's second largest gold mine, the finding is context-specific and not representative for gold mining in general, limiting the results' external validity.

Similarly, Lippert (2014) finds a positive impact of natural resources on locals' living standards. Zambians in copper mining constituencies are found to experience increases in household expenditure and consumer durable ownership, have better housing conditions and lower unemployment. Despite the capital-intensive nature of copper mining, local populations are found to benefit from resource extraction if the mining sector is sufficiently integrated with the regional economy. Both urban and rural households benefit, but the latter to a smaller extent. This study on the local impact of copper mining in a developing country highlights that a multitude of outcomes can be affected, though different sub-populations may not benefit equally. This highlights the importance of testing for heterogeneous effects for different sub-populations, such as rural and urban communities.

Comparing mining and non-mining districts in Peru, Loayza and Rigolini (2016) find evidence for higher average consumption per capita and lower poverty rates in mining districts. While this provides further evidence for positive local effects of resource extraction, mining districts are also found to have larger income inequality than non-mining districts. This indicates a trade-off. On the one hand, the share of the poor and extreme poor population is 2.6 percentage points lower in mining districts and per capita consumption is 9% higher. On the other, mining has a negative distributional effect. Inequality, measured by the Gini coefficient, is found to be 0.6 percentage points higher in mining districts. This dual effect is partly attributed to mining attracting highly educated migrants. This emphasises the importance of taking migration into resource-rich regions into account.

While the generally positive effects of natural resources on local populations found in the studies summarised above provide evidence against a "local resource curse", there is one outcome for which a resource curse seemingly exists at both the national and sub-national level: conflict. Various local impact studies suggest that natural resources, including oil, tend to increase conflict on a sub-national level (e.g. Berman et al., 2017; Dube and Vargas, 2013; Lei and Michaels, 2014).

Negative local impacts of natural resources are also found for health and environmental outcomes. The environmental pollution arising from mining may detrimentally affect local soil and water conditions. This is supported by evidence from Ghana, where Aragón and Rud (2015) find that gold mining significantly reduces agricultural productivity. Studying the effect of mining in 44 developing countries, von der Goltz and Barnwal (2019) find evidence suggesting a health-wealth trade-off, and highlight negative health impacts on pregnant women. Similarly, oil spills are found to have detrimental effects on women's reproductive health in Peru (Sebastián, Armstrong and Stephens, 2002) and on neonatal mortality in Nigeria (Bruederle and Hodler, 2019).

2.2 Gender-specific effects of natural resources

Second, this paper relates to the literature on the gender-specific effects of natural resources. While the majority of the economic literature does not explicitly account for gender-specific effects of natural resources, a recent review of the empirical literature shows that there is strong evidence for gender-differential impacts of extractive industries at both the macro and micro level (Benshaul-Tolonen and Baum, 2019). Women may benefit from arising economic opportunities, but are often also disproportionately put at risk by the social, economic and environmental consequences of resource extraction (World Bank, 2013). Key outcomes affected include female labour force participation, marriage markets, health and security – the overall impact of extractive industries on women is described as a mixed blessing, as the impact varies across sectors and contexts (Benshaul-Tolonen and Baum, 2019).

Two papers on natural resources' gender-specific impacts that are closely related to our paper concern the local effect of mining on women in Africa (Kotsadam and Tolonen, 2016; Tolonen, 2019). These papers are the most relevant to our study as they estimate a highly localised impact and consider similar micro-level employment and empowerment outcomes in the African context.

Kotsadam and Tolonen (2016) examine the local employment impact of large-scale mining, their study covers mines for the major minerals[3] extracted in 29 African countries. Exploiting spatial and temporal variation, they find that mine opening has two significant effects on women living within a 20 km radius of a mine: they either leave the workforce or move from self-employed agriculture to services. Mine opening is found to decrease overall female employment by 8%. At the same time, women are two percentage points more likely to work in services, which corresponds to an increase in the likelihood of service employment of over 50%. These results indicate that mine openings induce a structural shift in female employment. Industrial mining seems to enable some women to switch from the primary to tertiary sector in Africa, though overall, the results suggest that living close to mineral extraction sites is a mixed blessing. As this paper explores the local employment effects on women across Africa, this makes the setting very comparable to our study. It provides evidence for gender-specific local employment effects of natural resources in developing countries. This motivates us to explore the local employment effects of oil - another key resource in Africa, which is not included in the mineral dataset used by Kotsadam and Tolonen - on women.

In a related study on gold mining in sub-Saharan Africa, Tolonen (2019) finds evidence for improvements in women's gender norms. Gender norms in this context are understood as "attitudes, constraints, and bargaining power in household decisions" (2019, p.2), which are mainly measured by women's justification of domestic violence, their access to healthcare and decision-making power. Similar to the methodology applied in Kotsadam and Tolonen (2016), identification relies on exploiting spatial and temporal variation in gold mine opening in eight African countries. Employing a difference-in-difference design, Tolonen compares differences in gender norms for treated (located within a 15 km radius of mines) and untreated (located within 15-100 km from mines) communities, before and after gold mines became active. The results indicate a significant and positive change in gender norms: women are 19% less likely to justify domestic violence and self-stated barriers to healthcare decrease by 23%. This local effects study is the first to explore the causal effect of gold mining on gender norms, thereby highlighting the methodology's potential to examine a new range

[3]The ten most important minerals in this study are gold, silver, platinum, aluminium, copper, lead, nickel, tin, zinc and palladium (Kotsadam and Tolonen, 2016).

of outcome variables and providing further insight into resources' gender-specific local impact. This study is closely related to our paper, both in terms of methodology and outcomes studied. The key distinguishing feature is the natural resource of interest, which is gold in Tolonen's paper and oil in our study. Our paper thus also tests whether Tolonen's finding for gold mines can be generalised to oil, another key resource in Africa that differs in capital intensity, land usage and average production time. Our baseline specification is similar to Tolonen; however, we modify it to account for the specific features of oil extraction and additionally test for persistence of treatment effects to identify potential long-term changes in outcomes.

Having discussed these relevant findings on the gender-specific local impact of mining on women, we will now review the existing evidence related to oil. However, it is crucial to note that the existing literature on the local impact of oil on our outcomes of interest is not extensive, especially in developing country contexts. We thus also include relevant findings from national-level studies and evidence from developed countries that provide insight on the gender-specific effect of oil on employment and empowerment, our two outcomes of interest.

Gender-specific effects of oil on employment

The structural transformation arising from oil production may affect men and women differently, which has led to studies on gender-specific employment effects of oil. Overall, the effect of an oil resource shock on female employment is ambiguous at best. The literature suggests two countervailing mechanisms through which oil can impact female labour force participation. On the one hand, an oil boom increases the prevailing wage rate, which increases women's incentive to join the workforce. On the other hand, the countervailing force of partners' higher wages and government transfers may raise women's reservation wage, which reduces their incentive to join the labour force (Ross, 2008). The extent to which one effect dominates the other is contingent on the prevailing gender-based segregation in society. One would expect an oil boom to reduce female labour force participation if women are not allowed in the non-tradable sector because of the crowding out of the tradable sector and the increase in the reservation wage. Employing a first-difference model with country fixed effects, Ross finds that the extraction of oil and gas results in a reduction in female labour force participation, and, by extension, constrains their political influence. This leads Ross to conclude that "petroleum perpetuates patriarchy" (2008, p.120). While Ross' study was the first to explicitly link the natural resource curse and gender, it has been criticised for not adequately accounting for cross-country differences in culture and politics (Benshaul-Tolonen and Baum, 2019).

By exploiting within-country variation in large oilfield discoveries across different US counties, Maurer and Potlogea (2017) address some of the concerns inherent in Ross' (2008) study, as their methodology can better account for differences in institutions, culture and other unobservable confounding factors at the national level. The authors consider the discovery of oil as a male-biased labour shock that causes an increase in men's earnings. According to their model, an increase in male wages tends to crowd out the tradable sector, which leads to the loss of female employment. This effect may be offset by an increase in the demand for local services, which is predominantly catered to by women. The extent to which one effect outweighs the other is an empirical question. The authors employ a difference-in-difference design and compare the evolution of oil-rich counties to non-oil rich counties over the same period. They do not find a negative impact of the oil boom on female employment, rather a slight increase in employment for unmarried women, which stands in contrast to the cross-country results of Ross (2008).

Gender-specific effects of oil on empowerment

Few studies have addressed the effect of oil on women's empowerment and gender norms, and the existing ones fall mainly in the realm of political science. For example, Liou and Musgrave (2016) argue that autocratic leaders in oil-rich economies purposefully design policies to limit gender equality, which is one way to ensure government survival. Another related study is the study of gender-specific violence in the US. Following the 2008 oil boom in the Bakken region of Montana and North Dakota, domestic and dating violence have been found to increase (Jayasundara et al., 2018).

2.3 The effect of oil on employment

Third and finally, our paper relates to studies examining the effect of oil on employment. This effect has been studied on both the national and local level, and we will focus specifically on the channels through which effects may occur.

The oil industry may generate direct and indirect employment effects. Direct employment includes both high-skilled occupations, mainly in technical and managerial positions, and lower-skilled positions, e.g. security personnel, construction and maintenance workers and administrative support. Oil extraction's high capital intensity explains why the total number of direct oil-related employment is typically small (Adusah-Karikari, 2015). Additionally, specialist positions in developing countries' oil industry are often occupied by high-skilled internationals. Thus, no large direct employment effects on the local population are expected – this is especially true for women, who are generally underrepresented in the oil industry in Africa (Fajana, 2005). Local employment is thus not expected to be primarily generated through direct employment, but through indirect employment requiring lower skills (Obeng-Odoom, 2014). However, it is a priori ambiguous whether oil extraction has a significant impact on local employment in other sectors.

The enclave hypothesis

Producing oilfields may not impact local employment at all. This would lend support to the enclave theory, first posited by Hirschman (1958). According to this hypothesis, the linkages from natural resource industries to local economies are nonexistent or very small at best. In this strand of literature, natural resource extraction and production are assumed to form an enclave that does not induce spillover effects to local areas in the vicinity. There is little empirical research on the enclave hypothesis, but the oil industry in Africa has often been characterised as "secured enclaves, often with little or no economic benefit to the wider society" (Ferguson, 2005, p.378). Offshore oil production, in particular, is described as an enclave economy, for instance in the cases of Ghana (Ackah-Baidoo, 2013) or Angola (Ferguson, 2005). Some empirical support for the enclave hypothesis has been found in the context of gold mining in Peru, which would have constituted an enclave in absence of the local procurement policy that created strong linkages to local communities (Aragón and Rud, 2013).

Local Multipliers

In contrast to the enclave hypothesis, there is a strand of the natural resource literature emphasising the presence of linkages and spillovers. Spillovers here refer to the effect of one industry on another – in our context, they would arise if the oil industry affects other sectors of the economy.

The size of these spillovers may vary across different sectors and over the short- and long-term. According to this literature, natural resource extraction will increase local employment if the spillovers and linkages are sufficiently strong.

Empirically, spillovers are typically measured by job multipliers, the number of jobs in other sectors that are created by one additional job in the directly impacted industry. Local multipliers, a framework formulated by Moretti (2010), arise when a new job increases the demand for local goods and services, thereby creating additional jobs. One example of local multipliers to the non-resource sector arising from oil extraction is increased demand for materials required for the operation and maintenance of an oilfield. The income earned by employees in the oil and manufacturing industry may then be "spent on restaurant meals, home improvements, other locally provided goods and services" (Marchand and Weber, 2018, p.481). This would increase demand in manual, agricultural and service sectors, and thus generate additional local employment. Whether the employment generated in the indirectly-impacted industries is actually taken up by locals depends on the size of the resource industry relative to the local labour market. If the latter is sufficiently large, locals will benefit from increased employment, whereas a small local labour market relative to the natural resource sector may induce migration into the resource areas (Marchand and Weber, 2018).

In the natural resource literature, the local multiplier approach was first applied to the coal industry, where Black et al. (2005) found that 0.174 additional jobs were generated by each coal mining job in the US. Most empirical evidence for local multipliers from the oil industry comes from North America. Positive local multiplier effects have been established for the oil and gas industry in the US, recent estimates range from 0.3 (Weinstein, 2014) to 2.17 (Fetzer, 2014) additional jobs. Empirical evidence for a long-term multiplier effect of 0.3 jobs was found for oil extraction in Texas (Lee, 2015). Similar effects have been found for energy extraction in Canada, which mainly affected the construction, retail and service sector (Marchand, 2012). The effect that each resource job creates between one and two additional jobs in the local economy is considered as one of the better established findings in the natural resource literature (Marchand and Weber, 2018). However, to estimate the precise local multiplier, the number of directly generated employment in the impacted industry has to be known, which is not always the case. Thus, scholars have been using various measures, which include resource dependence, endowment and extraction. For instance, an additional oil rig was found to create 37 immediate and 224 long-run jobs in the US (Agerton et al., 2015).

Overall, the argument of sufficiently strong linkages and local multipliers stands in contrast to the hypothesis of the resource industry forming a closed enclave. In our empirical analysis, we will examine whether women's employment is impacted in the specific context of oil extraction in Africa, and which sectors are affected.

3 Background

We next provide an overview of the life cycle of an oilfield and discuss the distinguishing features of oil extraction, which have important implications for the outcomes we study.

3.1 The life cycle of an oilfield

The life cycle of an oilfield can be divided into five distinct phases: exploration, appraisal, development, production and closure. Borthwick et al. (1997) and Darko (2014) provide an overview of the different phases. The process starts with an exploration survey to identify major sedimentary basins. The entire exploration phase may take between one and five years and often, the quest for oil resources ends with no new discoveries. A temporary and often highly-skilled and international workforce is associated with the exploration activity. When exploratory drilling is successful, the appraisal phase begins, during which more wells are drilled to quantify the size of the field. This usually lasts between four and ten years.

Before oil production can begin, development of the field is undertaken. The development phase can be divided into the planning phase and the construction phase. The construction phase usually lasts between four and six years (Dietsche et al., 2013). Overall, the development phase may last for about ten years. However, it is crucial to note that all these time frames are approximations - there is considerable variation in the duration of the exploration and development phases across countries and projects. Often, multiple production wells are drilled to reduce land requirements and the overall infrastructure cost. The size of an oilfield is directly proportional to the well sites constructed and the total area occupied typically spans several hectares of land.

Finally, the production phase, which usually lasts between 20 and 50 years, may commence. Initially, oil wells are freely flowing and the rate of flow depends on a variety of factors, such as the properties of the reservoir rock, the underground pressure and the viscosity of the oil. These factors vary during the lifetime of the well. Extracted oil is then transported using pipelines, which are connected to the refineries or to ports where the oil is shipped using tankers (Kojima et al., 2010; Meili et al., 2018). The labour requirements during the construction and production phases include various skill levels. While some work is highly technical and managerial, requiring specialised training and education, there are also unskilled manual labour requirements.

3.2 Specific features of oil extraction

As discussed in the literature review, several studies have examined the local impact of mineral resources, such as gold or copper (e.g. Kotsadam and Tolonen, 2016; Lippert, 2014). However, generalisation of these studies to oil extraction may not hold if there are systematic differences between oil and mineral extraction. These differences could be apparent in terms of extraction methods, capital and labour intensity or infrastructure requirements. Further, such differences may translate into different local effects, even for the same outcomes studied.

Cameron and Stanley (2017) outline key differences between oil and mineral extraction. First, capital intensity is higher for oil extraction than for mining. Second, oilfields generally occupy a smaller area than mines. Mining operations are entirely land-based and cover a larger area, which makes them more likely to impact local communities. Third, external inputs and support infrastructure are present to a greater degree in the case of mining. Oil projects require less infrastructure development in terms of roads and railways in comparison to high-volume commodities such as coal, iron-ore and copper. Instead, oil production requires the development of a substantial network of pipelines, which is the most common way of oil transportation. Thus, the oil industry can be characterised as more capital-intensive, occupying smaller areas of land and less reliant on local infrastructure than mining. These key differences between oil and other extractive resources may result in different implications for the local development of the region, which motivates local impact studies focusing specifically on oil.

Another specific feature of oil extraction in Africa are local content requirement (LCR) policies, aimed at creating backward and forward linkages to the local economy. A key objective is to generate direct employment by requiring oil companies to locally source a given percentage of their total workforce. With the exception of Madagascar, all countries in our sample have LCR policies in place. In Angola for instance, LCR policies require that oil companies locally source 100% of unskilled manual workers, 80% of the middle-level workers and 70% of the higher-level workers in the oil industry (Oyewole, 2018). However, women make up a very small percentage of the total labour directly employed in the oil industry (Fajana, 2005). While LCR policies can be expected to induce greater participation of local workers in the oil extraction process, the capital-intensive nature of oil extraction may truncate the magnitude of such direct employment effects, as well as indirect employment effects generated through local multipliers.

4 Data

4.1 Data sources

We compile a unique dataset for the purpose of our study by combining data from three sources. The oilfield data is obtained from Rystad Energy, a Norwegian energy consultancy; individual-level data on employment, empowerment and control variables comes from the Demographic and Health Surveys (DHS); and additional geographical data is taken from the Global Administrative Area Database (GADM).

DHS Survey Data

The DHS Program provides a comprehensive dataset of standardised household surveys. The main areas covered are population, health and nutrition; the specific outcomes we use for this study are women's occupation, decision-making power and self-stated barriers to healthcare. The DHS data also contains extensive information on women's characteristics, such as their age, level of education, household characteristics and area of residence, which we employ as controls. The standardised surveys have been widely used in Development and Health Economics, and the DHS Program is generally described as the best available source for pan-African survey data. We use repeated cross-sectional data comprising standard DHS surveys for six African countries: Angola, Chad, Egypt, Gabon, Madagascar and Nigeria. Our data covers a period of twenty-six years, ranging from 1990 to 2016. We utilise all available standard DHS surveys with geo-coded data for oil-producing countries in Africa. This enables us to use seven survey rounds for Egypt (1992-93, 1995-96, 2000, 2003, 2005, 2008, 2014), four survey rounds for Nigeria (1990, 2003, 2008, 2013), two survey rounds for Madagascar (1997, 2008-09) and one survey round for Gabon (2012), Chad (2014-15) and Angola (2015-16) each. The choice of countries is restricted by the availability of locational data for the surveyed DHS clusters. The lack of geo-coded survey data precludes the inclusion of other oil-producing African countries, such as Niger or the Democratic Republic of Congo. We obtain an overall sample size of about 100,000 women; the average sample size per DHS cluster is 14. Our sample only includes usual residents, visitors are excluded.

Oilfield Data

Our oilfield data is obtained from Rystad Energy, an independent energy research and business intelligence company located in Oslo, Norway. This data is not publicly available and provides detailed and up-to-date information on oilfields. Notably, it includes the GPS coordinates of the centre-points of oilfields, which enables us to determine oilfields' precise geographic location. We further obtain the start-up year of each oilfield, i.e. the year when production started. Our sample contains all 237 oilfields that started production during the period from 1995 to 2018. This choice is determined by the availability of DHS data, which ranges from 1990 to 2016. As no geo-coded DHS survey data is available before 1990, we could not convincingly analyse pre-treatment trends for oilfields opening before 1995. Therefore, our analysis does not cover oilfields starting prior to 1995, and we exclude all DHS clusters that are within the treatment radius of such oilfields.

The Rystad data enables us to identify oilfields' precise location and start-up year, which is a key advantage distinguishing our study from the previous literature. Comprehensive data on oilfields is not publicly available, which may explain why many scholars have used Horn's (2004) dataset on giant oil and gas fields worldwide. However, this data set has two key limitations. First, it only encompasses giant fields with a pre-extraction size of 500 million barrels of oil or equivalent, which limits the analysis, as smaller and medium-sized oilfields are not included. Second, it does not include fields discovered after 2004, which restricts the analysis to less recent oil discoveries. An alternative commercial source used in the literature is the Wood Mackenzie PathFinder database, which includes similar information to that provided by Rystad Energy.

Further Geographical Data

We further use the Global Administrative Area Database (GADM), which provides geospatial data. This enables us to divide countries into subdivisions, which is necessary in order to identify DHS clusters' and oilfields' location in regions (GADM level 1) and districts (GADM level 2). This geographical information is needed to control for district fixed effects and regional time-trends.

4.2 Data compilation

We use the Geographic Information System (GIS) application QGIS to match the DHS clusters to oil deposits that become operational oilfields during the period from 1995 to 2018. We create circular buffers of different radii[4] around each oil deposit in order to identify the treatment and the control group. Clusters within smaller radii will be our treated group, as they identify women living in the vicinity of oil deposits. Clusters located between the smaller and the larger radii comprise the control group. In our baseline analysis, women living within 15 km of an oil deposit are considered as treated, and women living between 15 and 100 km of an oil deposit comprise the control group. The treatment group comprises 7.9% and the control group 92.1% of the total sample.

The summary statistics for our outcome variables and sample characteristics are displayed in Table 1. Women are on average 32 years old and have completed 7 years of education. 56% of women live in rural areas; and almost two-thirds of sampled women have never moved, i.e. always lived in the city, town or village they currently reside in. The majority of women are married (81%), and women have on average three children. About one-third of women state that they are

[4]The choice of treatment and control radii will be further explained in Section 5.2.

currently working. The service sector is the largest sector providing employment to women (21%), followed by agriculture (8%) and manual work (2%).

Table 1: Summary Statistics

	Mean Value					Min	Max
Sample	Entire Sample	Control Group	Treatment Group	Control Group	Treatment Group		
Time Period	both	pre	pre	post	post		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Characteristics							
Age	32.192	32.574	31.296***	32.055	31.142	15	49
Education	6.71	6.132	5.079***	7.301	6.191	0	24
Never-mover	0.635	0.614	0.729***	0.641	0.741	0	1
Female household head	0.126	0.121	0.133**	0.13	0.121	0	1
Rural	0.564	0.534	0.788***	0.559	0.691	0	1
Married	0.812	0.824	0.837**	0.803	0.799	0	1
Number of children	2.794	2.906	3.045***	2.686	2.886	0	15
Outcome Variables							
Working	0.313	0.315	0.294***	0.314	0.3	0	1
Services	0.214	0.203	0.173***	0.226	0.202	0	1
Agriculture	0.077	0.09	0.098	0.064	0.085	0	1
Manual	0.022	0.022	0.023	0.022	0.014	0	1
Decision-making power	0.685	0.663	0.577***	0.7	0.674	0	1
Barriers to healthcare	0.242	0.262	0.243**	0.233	0.26	0	1
(%) of sample	100%	39.1%	4.2%	53%	3.7%		

Treated = 1 indicates living within 15 km from an oil deposit
Treated = 0 indicates living within 15-100 km from an oil deposit
Active = 1 indicates an operational oil deposit (i.e. active oilfield)
Active = 0 indicates a non-operational oil deposit
Pre-treatment, control group: *Treated* = 0 and *Active* = 0
Pre-treatment, treatment group: *Treated* = 1 and *Active* = 0
Post-treatment, control group: *Treated* = 0 and *Active* = 1
Post-treatment, treatment group: *Treated* = 1 and *Active* = 1
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ for a t-test between control group, pre (2) and treatment group, pre (3)

5 Methodology

Exploiting the spatial and temporal variation in the opening of oilfields across six countries, we test whether the industrial development induced by oil extraction leads to changes in women's employment and empowerment. In this section, we explain our methodology and empirical specification by firstly describing our outcome variables, secondly elaborating on the empirical strategy, and finally discussing the assumptions needed for estimation.

5.1 Outcomes

We examine two broad categories of outcomes: women's employment and empowerment. While the choice of outcome variables for the former is relatively straight-forward, the latter is more ambiguous, due to the increased difficulty of measuring and quantifying empowerment. We will therefore explain our choice and measurement of outcome variables below.

5.1.1 Employment outcomes

We examine four employment outcomes. The first outcome measure is the broadest; it measures whether women are working. The DHS surveys record whether respondents are currently working, which 31.3% of women in our sample answer affirmatively. The fact that over two-thirds of our sample are not working emphasises the importance of examining female employment. *Working* represents an indicator variable taking on the value 1 if a woman is currently working, and 0 otherwise.

Our second, third and fourth outcome measures capture employment in different sectors. This allows us to test whether oil extraction generates indirect employment opportunities for women in different sectors of the economy. The DHS Program classifies occupations based on the ILO International Standard Classification of Occupations (ISCO). This information on women's occupation enables us to divide the working respondents into different sectors. Following standard economic theory, we divide the economy into three sectors: primary, secondary and tertiary. Utilising the occupational categories provided by the DHS, the primary sector includes agricultural self-employment and agricultural employment; the secondary sector is comprised of skilled and unskilled manufacturing; and the tertiary sector includes professional, clerical, household, sales and service employment. These three sectors, to which we will hereafter refer to as agriculture, manual and services, are not equally important in providing employment to women in our sample countries. About one-fifth of our sampled women work in services, followed by agriculture (7.7%) and manual work (2.2%). The service sector is thus the one providing employment to the majority of working women; over two thirds of all working women are employed in services. One quarter of working women are engaged in agriculture, and 7% of working women in manual work. A small section of our sample could not be classified to any sector, as their occupation was recorded as "other" in the DHS. The variables *Agriculture, Manual* and *Services* are binary variables taking on the value 1 if a woman is working in the respective sector, and 0 otherwise.

5.1.2 Empowerment outcomes

There is no single definition for empowerment, but in its broadest sense, it is "the expansion of freedom of choice and action" (World Bank, 2002). Female empowerment, in particular, describes the "process by which women gain power and control over their own lives and acquire the ability to make strategic choices" (EIGE, 2019). According to the UN, there are five components to women's empowerment: "women's sense of self-worth; their right to have and to determine choices; their right to have access to opportunities and resources; their right to have the power to control their own lives, both within and outside the home; and their ability to influence the direction of social change to create a more just social and economic order, nationally and internationally" (UNDP, 2017). These dimensions highlight the challenge of adequately measuring women's empowerment – one cannot easily quantify self-worth or the power to control one's life.

The data collected by the DHS Program can, however, shed light on some of the dimensions of female empowerment. Notably, the questions on women's decision-making power in the household and the obstacles faced in accessing healthcare are closely related to the UN's definition of women's empowerment – they describe women's ability to make choices and to exercise control over their own lives.

Decision-making power in the household

Our first measure of empowerment explicitly relates to women's decision-making power, and thus the power dynamics within the household. About 70% of the women in our sample do not work, implying that they are largely involved in domestic activities and considerably confined to their homes. Consequently, their decision-making power within the household becomes crucial to understanding how empowered they are. The DHS data contains a number of variables which capture whether women have a "final say" in various areas.

We capture women's decision-making power by an index combining three variables. The three questions used cover women's decision-making power regarding their own health, making large household purchases and visiting family and relatives. These questions thus relate to decision-making power in the domain of health, money and family dynamics. We capture each of the three questions by a dummy variable indicating whether a woman has decision-making power or not, where 1 indicates that women have decision-making power and 0 indicates that their husband/partner or another person alone has the final say. We refrain from testing each of the variables individually, as the risk of observing a significant result by chance increases with the number of hypotheses tested. Instead, we take the average of the three variables to create an overall index of decision-making, which reduces multiple inference problems. Crucially, all questions have similar numbers of observations and a sufficiently large overlap, which enables us to combine them into an index. We refrain from including questions of arguably lesser importance to female empowerment, such as who decides on "the food to be cooked each day". While still indicating some decision-making power, this cannot be considered as equally important as deciding on one's own health, large purchases or mobility.

Ideally, we would have liked to also include information on women's decision-making capacity regarding contraceptive use and sexual relations in order to obtain a more comprehensive measure. The DHS surveys are suited to cover these aspects; relevant questions include those on who "decides on contraceptive use" and whether a woman "can refuse sex" (DHS Program, 2018). However, as these questions are asked in different survey modules, we only have a limited overlap in observations and varying sample sizes, which prohibits including them in our index measure.

Self-stated barriers in accessing healthcare

Our second measure of women's empowerment relates to accessing healthcare. Since independently seeking medical help for oneself is strongly related to decisions concerning one's own health and body, this is an important dimension of female empowerment. The DHS Program collects information on factors that constitute a "big problem" to women in their access to healthcare. Three questions are directly related to women's empowerment, they cover whether "having to ask for permission", "not wanting to go alone" or "getting the money needed" constitute an obstacle in women's access to healthcare. To avoid problems arising from multiple hypothesis testing, we combine them into an index by taking the average. This is possible as all questions are asked in the same DHS survey module, which ensures a large overlap in observations. For each of the three questions, any factor that does not constitute a big problem in accessing healthcare is coded as 0, any factor that is a big problem takes on value 1. Our index on barriers to healthcare ranges from 0 to 1; larger values correspond to more obstacles in accessing healthcare, and thus limited empowerment.

Crucially, we do not consider factors such as "distance to health clinic", as this is a factor predominantly related to infrastructure, not to women's empowerment. A woman stating that the distance is a large problem in accessing healthcare may live in a remote rural region and thus be hindered from easily accessing healthcare, but the distance by itself does not provide any information on her ability to make independent choices in accessing healthcare. In contrast, if she has to ask for permission or money or is uncomfortable going alone, this is directly related to her ability to decide for herself and to exercise power over her own life.

Overall, it is crucial to note that our two outcome measures – decision-making in the household and self-stated barriers to healthcare – cannot provide a comprehensive picture on women's empowerment. Important aspects, such as political participation and representation, are not covered by DHS data and hence cannot be examined in this analysis.

5.1.3 The interrelated nature of women's employment and empowerment

The two outcome categories we examine – women's employment and empowerment – are closely related. Female employment, and the income it generates, can improve women's empowerment. This is illustrated by the World Bank's current gender strategy, which has increasing women's labour force participation and income-earning opportunities as one of four key objectives (World Bank, 2015). The channels of female labour force participation impacting women's empowerment are manifold. They include delayed childbearing and marriage – empirical evidence for this has, for example, been found for female garment workers in Bangladesh (Heath and Mobarak, 2012) or in case of increased office employment opportunities in an experimental setting in rural India (Jensen, 2012). Additionally, female labour force participation is found to increase women's bargaining power (e.g. Duflo, 2003; Majlesi, 2016). Of course, increased employment is not the only channel driving positive changes in women's empowerment. Other factors, such as access to media, have also been found to significantly improve female autonomy, reduce fertility and improve decision making power (Cheung, 2012; Jensen and Oster, 2009; La Ferrara et al., 2012).

Women's employment does not only impact their empowerment – causality can also go in the other direction. For example, analysing mother-child pairs from the National Longitudinal Survey of Youth, Farré et al. (2012) find that the intercultural transmission of gender role attitudes contributes to heterogeneity in women's labour supply. The two-way causality is also addressed by Basu's (2006) model of household behaviour, which posits that female labour supply determines the household's balance of power, but the household's balance of power also impacts women's labour supply.

Women's employment and empowerment are correlated; and they both affect economic development. The relationship between women's labour force participation and the level of economic development has traditionally been described as U-shaped (e.g. Boserup, 1970; Mammen and Paxson, 2000). According to these models, low levels of development correspond to high female employment, mainly driven by agriculture and self-employment. As development increases, women leave the labour force – this pattern only reverses at higher stages of development, when more tertiary sector employment arises. This broad pattern can be observed in our six sample countries: Egypt is the most developed and has the lowest female labour force participation rate of 23%, compared to 51% in Nigeria or 65% in Chad (World Bank Data, 2019). However, as developing countries have experienced simultaneous and interrelated increases in female education levels and female labour force participation over the past three decades, Heath and Jayachandran (2017) posit that the U-curve has shifted upwards, which can explain increased female labour supply despite most low-income countries being on the downward sloping part of the U-shape. The dual relationship between economic development and women's empowerment is emphasised by Duflo (2012). While economic development is generally assumed to decrease inequalities between men and women, persistent discrimination against women can also hinder development (Sen, 1990).

Overall, the simultaneity makes it difficult to establish a causal effect of economic development on women's employment and empowerment. In this study, we use oilfield openings as a plausibly exogenous shock to local economic development, and test whether it brings about changes in women's employment and empowerment. A caveat of our analysis is that we cannot disentangle the direction of causality between employment and empowerment.

5.2 Empirical strategy

Our identification strategy is based on a difference-in-difference framework that relies on exploiting the temporal and spatial variation in oilfield opening. We use the opening of an oilfield as a source of exogenous shock to the local economic area. The identifying assumption states that the location and timing of an oilfield opening are exogenous to changes in local population and labour market characteristics.

There are three characteristics of the oil industry that lend support to this assumption. First, oil extraction can only take place where there are oil reserves. Therefore, the primary requirement for the opening of an oilfield is the existence of oil below the surface of the earth, which is a geological phenomenon independent of local population characteristics. Second, the oil industry is capital intensive. Labour is not the main input of production, which makes it less likely that the local labour supply influences the location and timing of oil extraction. Third, the oil industry is globalised and dominated by transnational corporations, which employ many expatriate workers. This also makes the oil extraction business less reliant on local labour market and population characteristics. Thus, the necessary condition of oil reserves and the industry's capital intensive and highly globalised nature support the identifying assumption. Potential threats to identification, such as the role of institutions, infrastructure and local inputs, will be addressed in Section 5.4.

We are interested in the local effect of oil extraction on women's employment and empowerment outcomes. Thus, we use the opening year of an oilfield as the start of our treatment period. Using natural resource deposits as a plausibly exogenous measure for resource extraction is one of the common methodologies in the natural resource literature; and our identification strategy is similar to Kotsadam and Tolonen (2016) and Tolonen (2019). We do not use levels of oil production as the treatment variable, as the measurement of oil production may vary across countries and there are likely to be measurement errors.

We define an indicator variable – *Treated* – that is switched on when a DHS cluster is within the treatment radius of an oil deposit, which is time-invariant. Our analysis only includes oil deposits that start production at some point between 1995 and 2018, and continue production throughout the period of study. That means, all deposits included in our analysis will turn operational during the period of our study. The variable *Treated* enables us to classify our observations into the treatment and control group. We define another indicator variable – *Active* – which indicates whether an oil deposit was operational during the DHS survey year. The main treatment effect is given by the estimated coefficient on the interaction (*Active* × *Treated*) between the time-invariant oil deposit indicator (*Treated*) and the time-varying indicator variable for the operational status of an oil deposit (*Active*). We employ the following specification:

$$Y_{icdt} = \beta_0 + \beta_1 Treated_c + \beta_2 Active_{ct} * Treated_c + X_i + FE_{kt} + FE_d + \epsilon_{icdt} \tag{1}$$

where i indicates an individual observation, c indicates the DHS cluster, d denotes the district (GADM level 2), k is representative of the country and t indicates the year. It is important to cluster standard errors when clusters of units rather than individual units are assigned to the treatment (McKenzie, 2017). Since all the individuals in a cluster are either treated or untreated in our analysis, we cluster our standard errors at the DHS cluster level. The specification includes year-country fixed effects (FE_{kt}) that allow for a difference-in-difference interpretation. It also includes district fixed effects (FE_d) which control for any district-specific time-invariant characteristics that influence the outcome variables. Further, we also include a vector of individual-level control variables (X_i), which includes women's education, age, an indicator variable for living in a rural area and an indicator variable for a female household head.

While the opening of an oilfield is the exogenous shock that we consider, it is important to factor in the construction phase of an oilfield. It is possible that the outcome variables we study start changing before the oilfield turns operational, e.g. through spillover effects during the construction phase. While we do not have information on the precise duration of construction for each oilfield in our sample, we calculate the average duration of construction (*expected startup year* – *approval year*) for all African onshore oilfields currently under development, which is 5.2 years (Rystad data). This is in line with general estimates of the duration of the construction phase of oilfields, which usually takes between four and six years (Dietsche et al., 2013).

Therefore, we augment our baseline specification (1) to account for the construction phase of oilfields. As our treated population may be impacted during the construction phase, we include the interaction term *Treated* × *Construction*. As discussed above, the construction usually lasts up to six years prior to oilfield opening.

$$Y_{icdt} = \beta_0 + \beta_1 Treated_c + \beta_2 Treated_c * Con_{ct} + \beta_3 Active_{ct} * Treated_c \ + X_i + FE_{kt} \ + FE_d + \epsilon_{icdt} \tag{2}$$

In our third specification, we divide the construction phase into two separate periods to allow for differential impacts in the early and late period of construction, as the intensity and nature of construction work may differ between these two periods. The early construction period covers years four to six prior to oil extraction, and the late construction period covers the three years immediately prior to extraction. Both construction periods have an approximately equal number of observations. It is not possible to consider each year of construction separately since we have repeated cross-sectional data with considerable variation in the number of observations per year. Again, both construction periods are interacted with the indicator *Treated* to account for a potential treatment effect during the construction periods. The augmented specification is as follows:

$$Y_{icdt} = \beta_0 + \beta_1 Treated_c + \beta_2 Treated_c * Early\ Con_{ct} \ + \beta_3 Treated_c * Late\ Con_{ct} + \beta_4 Active_{ct} * Treated_c + \ X_i + FE_{kt} + FE_d + \epsilon_{icdt} \tag{3}$$

The coefficient on *Active* × *Treated* gives us the mean effect computed over the entire post-treatment period. This specification assumes that oil extraction induces a level-shift. To test this assumption, we modify specification (3) by replacing *Active* × *Treated* with three distinct post-treatment time dummies, which are each interacted with our *Treated* variable.

$$\begin{aligned} Y_{icdt} = {} & \beta_0 + \beta_1 Treated_c + \beta_2 Treated_c * Early\ Con_{ct} \ + \beta_3 Treated_c * Late\ Con_{ct} + \\ & \beta_4 Active(0-3y)_{ct} * Treated_c \ + \beta_5 Active(4-8y)_{ct} * Treated_c + \\ & \beta_6 Active(post\ 8y)_{ct} * Treated_c + \ X_i + FE_{kt} + FE_d + \epsilon_{icdt} \end{aligned} \tag{4}$$

The three post-treatment time dummies denote three different periods of oil extraction. Similar to the construction phase, we divide the entire post-treatment period into three distinct time periods by allocating an approximately equal number of observations to each group. These three post-treatment categories broadly represent the immediate period after start-up (zero to three years), the medium-term (four to eight years) and the long-term (nine to nineteen years). We do not include each post-treatment year separately - which would have been the most detailed division of the *Active* × *Treated* interaction - for two reasons. First, the number of observations is not the same in each year due to the repeated cross-sectional nature of our data; including an interaction for each post-treatment year would lead to imprecisely estimated coefficients for years with few observations. Second, the estimation of 19 additional parameters would not be justified by economic reasoning. For instance, we are not aiming to capture the differential effect of being in the sixth or seventh year post start-up.

Specification (4) enables us to test the assumption of a post-treatment level shift in the outcomes. Failure to reject the equality of the three post-treatment coefficients ($\beta_4, \beta_5, \beta_6$) would support the assumption of a level shift in outcomes, as estimated in specifications (1) to (3).

Choice of Treatment and Control Radii

The choice of an appropriate treatment radius is vital for the correct estimation of the treatment effect. Defining a treatment radius inevitably involves a trade-off. If the radius is too large, individuals considered as treated may not actually be affected by the oil extraction; and if the radius is too small, part of the control group may be affected. We define the baseline treatment radius to be 15 km and the control radius to be 100 km. Women living within 15 km of an oil deposit constitute the treatment group; and women living within 15 and 100 km constitute the control group. The choice of the baseline treatment and control radii is based on the local impact literature and on measures of mobility in Africa.

First, it is guided by the existing literature on the local effect of natural resources. Tolonen (2019) uses a 15 km radius to estimate the impact of gold mining on gender norms. In another study estimating the impact of being in the vicinity of gold mines on infant mortality, Tolonen (2018) uses a radius of 10 km, but confirms that the economic footprint extends to a wider radius. Similar treatment radii are utilised in other contexts. For example, Aragón and Rud (2015) employ a distance of 20 km when estimating the local impact of gold mining on agricultural productivity in Ghana, and Kotsadam and Tolonen (2016) employ a radius of 20 km in a study focusing on the local effect on mining on employment across Africa.

Second, we aim to ensure that women are sufficiently mobile to take up potentially arising employment opportunities within the treatment radius. Commuting distances provide insight into the mobility of individuals, and it is reasonable to assume that the local employment effects of oil extraction will be largely confined to the area within which commuting is possible. Common modes of commuting in African countries include walking, cars, minibuses or motorcycles (Chen et al., 2017). Independent of household structure and marital status, women are found to commute less than males (Neto et al., 2015). This general pattern also holds in our sample, as illustrated by the shorter average commuting time for women (23 minutes) than for men (35 minutes) in Egypt (Ehab, 2018). Common commuting distances in African countries are between 5 km and 15 km (Aidoo et al., 2013; Kung et al., 2013). For instance, commuting distance to work in urban areas in Nigeria were up to 13 km in 1996 and up to 19 km in 2006 (Ogunbodede, 2008). These distances support the assumption that our baseline treatment radius of 15 km is not too large for women to take up potentially arising employment opportunities. Additionally, as the coordinates of DHS clusters are displaced up to 5 km to ensure respondents' anonymity[5] , we refrain from using a smaller treatment radius than 15 km.

When choosing the size of our control group, we aim to capture reasonably similar individuals and still ensure a large enough sample size. The choice of a control group between 15 km and 100 km is also guided by the literature (e.g. Tolonen, 2019). We decide not to include all non-oil producing regions in a country, as these would exhibit larger heterogeneity in local characteristics, such as economic development, political stability, vulnerability to natural disasters or demographics.

5.3 Assumptions

The difference-in-difference framework identifies a mean causal effect conditional on fulfilling a standard set of assumptions, which we will discuss in this section.

[5]Urban clusters are displaced up to 2 km. Rural clusters are displaced up to 5 km; and 1% of rural clusters are displaced up to 10 km.

A crucial assumption for any difference-in-difference analysis is the common trends assumption. The common trends assumption requires that the counterfactual behaviour of the treatment and control group is the same (Angrist and Pischke, 2009). In the absence of a treatment effect, the treatment and control group should follow a similar trend before and after the introduction of treatment, which in our case is the oilfield becoming operational. The assumption can be written as follows (Lechner, 2010):

$$
\begin{aligned}
&E(Y_1^0|X=x, D=1) - E(Y_0^0|X=x, D=1) = \\
&E(Y_1^0|X=x, D=0) - E(Y_0^0|X=x, D=0) = \\
&\qquad E(Y_1^0|X=x) - E(Y_0^0|X=x); \forall x \in X
\end{aligned}
$$

The term $E(Y_1^0|X=x, D=1)$ represents the expected value for the treated group in the post-treatment period, whereas $E(Y_1^0|X=x, D=0)$ denotes the expected value for the treated group pre-treatment. Similarly, for the control group, $E(Y_0^0|X=x, D=1)$ and $E(Y_0^0|X=x, D=0)$ indicate the post-treatment and pre-treatment expected values, respectively. In the absence of a treatment effect, we would expect the difference between the treatment and control group in the pre-treatment period to be equal to the difference between the control and treatment group in the post-treatment period, as represented in the equation above.

The difference-in-difference framework allows for a difference in levels of outcomes, as long as both treatment and control group follow similar trends pre-treatment. We note that there are differences in the levels of outcomes between our treated and control groups pre-treatment. These differences in levels do not pose a threat to a difference-in-difference estimation. Compared to women in the control group pre-treatment (see Table 1, column 2), treated women in the pre-treatment period (see Table 1, column 3) are less likely to work and less likely to be employed in services. They have less decision-making power, but face fewer barriers in accessing healthcare. The overall picture thus seems to be that prior to the onset of oil extraction, the treatment group is less progressive than the control group.

It is crucial to note that treatment occurs at multiple points in time. Oil deposits in our data become operational between 1995 and 2018; treatment is thus not identified by a single calendar year. Thus, instead of showing trends in outcomes over time, we employ a modified regression specification (5) to assess common trends.

$$
\begin{aligned}
Y_{icdt} = {} & \beta_0 + \beta_1 Treated_c + \beta_2 Treated_c * (pre\ 13y)_{ct} + \beta_3 Treated_c * (4-6y\ pre)_{ct} + \\
& \beta_4 Treated_c * (1-3y\ pre)_{ct} + \beta_5 Treated_c * Active(0-3y)_{ct} + \\
& \beta_6 Treated_c * Active(4-8y)_{ct} + \beta_7 Treated_c * Active(post\ 8y)_{ct} + \\
& X_i + FE_{kt} + FE_d + \epsilon_{icdt}
\end{aligned} \tag{5}
$$

We regress our outcome variables of interest on the grouped years before and after oilfield opening, each interacted with *Treated*. As before, the regression also includes the *Treated* dummy, year-country fixed effects, district fixed effects and covariates. As elaborated in our empirical strategy, the heterogeneity in the number of observations precludes the inclusion of each year relative to oilfield opening individually. We now divide the pre-treatment period into four time periods, ensuring approximately equal number of observations. The division of the construction period and post-treatment period remains the same as in our main regressions. The construction phase is the earliest period where we may expect a treatment effect, based on the life cycle of oilfields as described in Section 3.1. As construction lasts approximately four to six years in Africa, we omit the pre-treatment period immediately before the construction phase.

Each estimated coefficient thus denotes whether the difference between treatment and control in that particular period is significantly different from the difference between treatment and control in the omitted period. Therefore, if the estimated coefficient is significantly different from zero before treatment, this indicates a violation of the common trends assumption. In graphs 1 to 6, the x-axis depicts the relative time before and after oilfield opening, year zero corresponds to an oilfield's start-up year.

For service sector employment, the estimated coefficients are shown in Figure 1. The coefficients in the pre-treatment period (except for the late construction phase, i.e. the three years immediately prior to oilfield opening) are not significantly different from zero. This indicates that there is no treatment effect prior to the late construction phase, which lends support to the assumption of common trends. The estimated coefficient on the late construction phase is significantly different from zero, which indicates a treatment effect. Similarly, we observe significant coefficients post-oilfield opening, indicating the presence of a treatment effect which we will verify in our full regression analysis in Section 6.

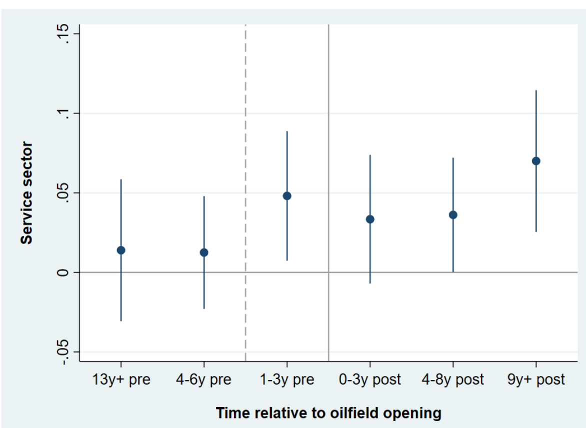

Figure 1: Service Sector Employment

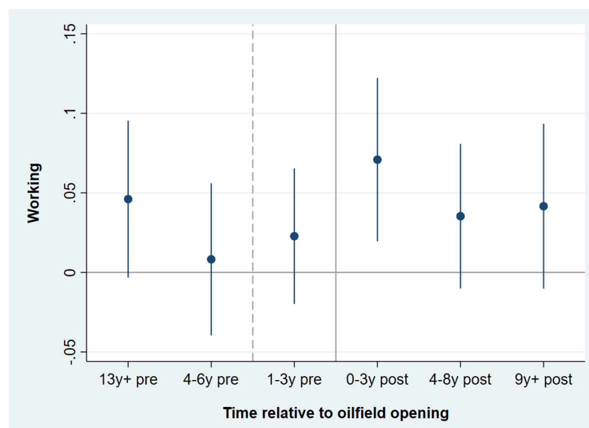

Figure 2: Total Employment

For total employment (Figure 2), none of the coefficients in the pre-treatment period are significantly different from zero. This lends support to the common trends assumption. We do, however, observe some treatment effect in the period immediately after oilfield opening.

For the agricultural sector (Figure 3), the 95% confidence interval of all estimated coefficients intersects zero, indicating no significant divergence of common trends throughout the period of study. Similarly, none of the coefficients for the manual sector (Figure 4) in the pre-treatment periods are significantly different from zero.

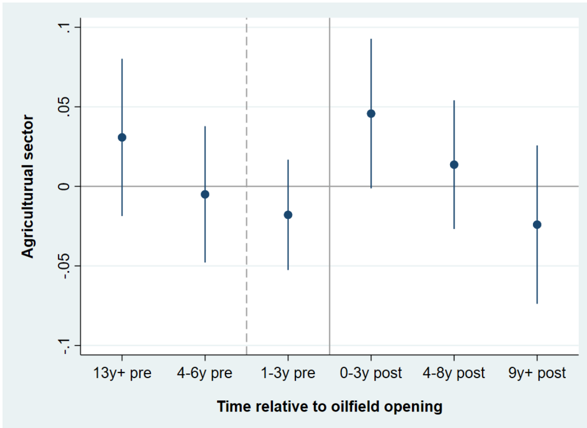

Figure 3: Agricultural Sector Employment

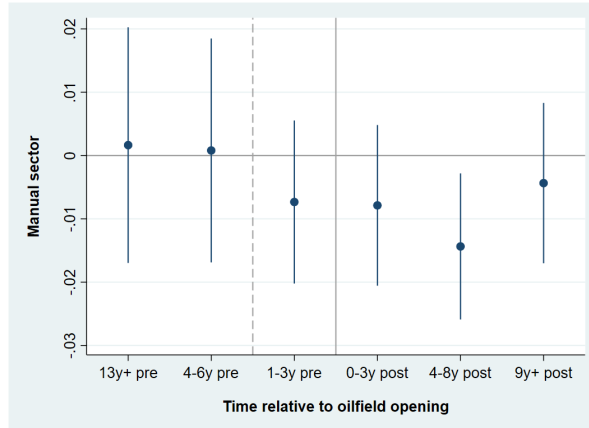

Figure 4: Manual Sector Employment

The empowerment outcomes that we study are not available in the earlier DHS surveys, limiting the pre-treatment periods we can examine. However, all the coefficients, both before and after an oilfield opening, are not significantly different from zero (Figures 5 and 6). While this supports the common trends assumption pre-treatment, it may also suggests that no treatment effect exists, which we will further analyse in Section 6.

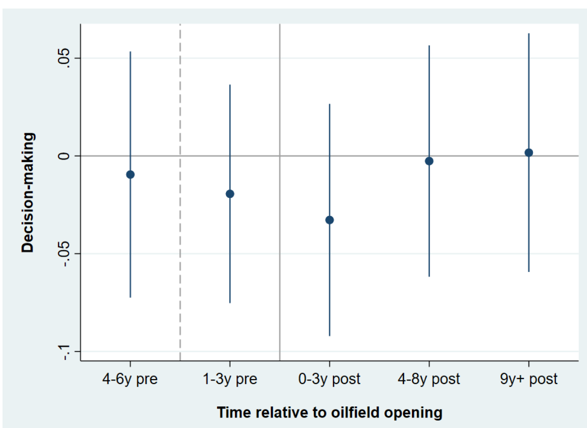

Figure 5: Decision-making in the Household

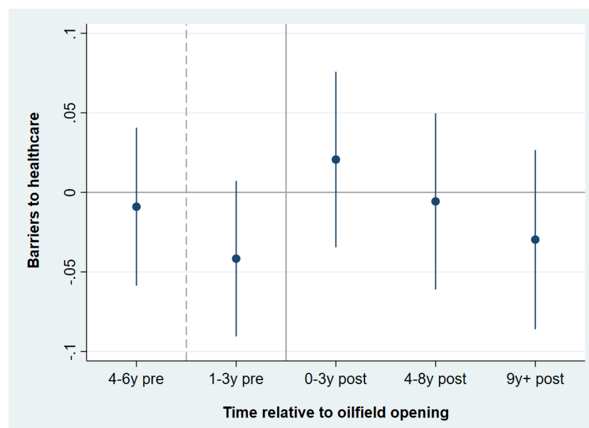

Figure 6: Self-stated Barriers to Healthcare

Overall, Figures 1 to 6 confirm that trends do not diverge significantly prior to the late construction period, which supports the notion of common trends over a sufficiently long period before treatment. This suggests that the early stages of the oilfield life cycle, such as the appraisal phase or the planning period during development, do not pose a serious threat to the common trends assumption.

Another assumption specific to the strict difference-in-difference framework is that the treatment should have no impact on the pre-treatment population in the pre-treatment period. However, since an oilfield opening is preceded by a construction phase, we partly relax this assumption in our analysis by controlling for the years immediately prior to the opening of an oilfield, as elaborated in our empirical strategy.

Another important assumption for causal inference is the Stable Unit Treatment Value Assumption (SUTVA). This assumption requires that only one of the potential outcomes is observable for every member of the population (Lechner, 2010). This precludes any relevant interaction between the treatment and control group. Our choice of the treatment radius, although based on the literature and on common commuting distances in the region, cannot rule out any spillover treatment effects beyond the 15 km radius. Hence, it is possible that the SUTVA assumption is violated in our analysis. However, we do not consider this to be a big cause of concern. We expect the local impact of an oilfield to decay with distance and hence expect any treatment effects beyond 15 km, if existent, to be weak at best. One of our robustness checks confirms this hypothesis: estimating the treatment effect based on a 25 km treatment radius, we find that our most precisely estimated baseline result – the positive impact on service sector employment – disappears completely.

Next, the exogeneity assumption (EXOG) requires that the covariates in our analysis are not influenced by the treatment. Characteristics such as age, education, having a female household head[6], and being in a rural or urban setting are highly unlikely to be influenced by the opening of an oilfield. These are standard controls in studies estimating the local impact of natural resources (e.g. Aragón and Rud, 2013; Kotsadam and Tolonen, 2016; Tolonen, 2019), which supports the exogeneity of our control variables.

Additionally, one might want to control for the number of children, marital status or partners' occupation in the regressions for employment outcomes. However, these controls would violate the exogeneity assumption, as they may be "bad controls", which are variables that may themselves be outcomes (Angrist and Pischke, 2009). It is plausible that oil extraction affects fertility, marital status and partners' employment, which is why we decide not to control for these covariates in our regressions. This is justified by findings of the previous literature, where Maurer and Potlogea (2017) find an increased marriage rate of young women in oil-rich counties in the US. Similarly, natural resources have been found to decrease the fertility of young women (Tolonen, 2019).

Classical Assumptions and the Linear Probability Model

As our employment outcomes are indicator variables taking on 0 or 1, we employ a Linear Probability Model in the employment regressions. There are a number of assumptions pertaining to the classical linear regression model that must continue to hold in case of the Linear Probability Model. The zero conditional mean assumption has to be fulfilled in order for the estimates to be unbiased. The assumption requires the idiosyncratic error term (u) to have an expected value of zero conditional on any value of the independent variables (x_1, x_2, ...) (Wooldridge, 2013). This can be represented as:

$$E(u|x_1, x_2, ..., x_k) = 0$$

Omitting an important variable that is correlated with any of the independent variables causes the zero conditional mean assumption to be violated. For our analysis, this assumption would be violated if any confounding time-varying factors, such as evolving regional institutions, influence

[6]This may be considered as an empowerment outcome. However, the household head in our context generally refers to the eldest household member. Having a female household head therefore does not mean that the respondent is the household head, it thus does not imply anything about the respondents' own empowerment.

the decision to open an oilfield. The year-country fixed effects cannot entirely capture such time-varying factors at the regional level. Therefore, as a robustness check, we include regional time trends to capture such time-varying confounding factors at the regional level.

Further, the strict exogeneity assumption requires the idiosyncratic error term in each year to be uncorrelated with any independent variable in the same year as well as in the preceding and succeeding years. However, the Linear Probability Model violates the assumption of homoskedasticity, which is a limitation of the model. This assumption is crucial for justification of the t- and the F-statistics, although it does not cause the coefficients to be biased. However, we address this issue of heteroskedasticity by clustering our standard errors at the DHS cluster level.

5.4 Internal validity

The internal validity of our study, i.e. our ability to correctly identify the causal effect of onshore oil extraction on the outcomes we study, crucially depends on the validity of the assumptions made. Below, we discuss potential threats to identification.

Exogeneity of Treatment

Our identifying assumption states that timing and location of oilfield opening are exogenous to changes in local characteristics, such as trends in population characteristics and labour market trends. The necessary condition of oil reserves and the oil industry's capital intensive and globalised nature support this identifying assumption. However, there are potential threats to identification, as other factors may drive the decision to open an oilfield. According to Eggert (2001), the placement of the oil deposits may be random, but the discovery of such deposits is not. The discovery and hence the location of an oilfield may be influenced by access to and cost of inputs, water supply, transportation and institutions. Thus, it would pose a threat to identification if inputs, existing infrastructure or institutional quality drive the timing and location of oilfield openings. We will address each of these threats to identification in turn.

First, due to its high capital intensity, the oil industry uses few domestic inputs (Söderling, 2006). Therefore, access to domestic inputs cannot be considered as a key factor driving the location and timing of oilfield opening. For example, a large transnational oil company is highly unlikely to decide on the opening of an oilfield due to large amounts of cement production in one specific location. Thus, we do not consider domestic inputs to pose a serious threat to the identifying assumption.

A second threat to identification is posed by variations in infrastructure, including water supply and transportation, which are important for oil production (Eggert, 2001). It does not pose a problem to our identification strategy if the opening of an oilfield spurs development of infrastructure in the region. However, it would be a cause for concern if existing infrastructure determines the location and timing of oil extraction. Oil is primarily transported through pipelines, which are the least expensive way of transportation given the economies of scale (Kojima et al., 2010; Meili et al., 2018). Thus, the low reliance on local road and railways mitigates the concern that existing infrastructure drives the timing and location of oilfield openings. However, it might be argued that new oilfields may be opened closer to ports or existing pipeline networks. We mitigate this concern by employing district fixed effects, which control for time-invariant features of a district, such as its strategic location, proximity to ports or existing pipeline infrastructure.

Further, oil production requires electricity and water. As in the case of transportation, it is not a problem for our analysis if the development of an oilfield brings about better water supply or electricity in the region, as long as it is not the other way round. Transnational oil companies typically gather inputs such as power and water supply to ensure reliable infrastructure for their operations (Cameron and Stanley, 2017) – it is not the case that the availability of power and water supply determines the location and timing of opening an oilfield. Thus, the overall dependence of oil on existing infrastructure is low. While we cannot completely rule out the concern that local infrastructure poses a threat to the identifying assumption, the nature of oil extraction and the inclusion of district fixed effects mitigate this concern. Importantly, the threat to identification is lower than in comparable studies on the local effect of mineral resources, which are much more reliant on railways and roadways compared to oil (Weng et al., 2013).

Third, factors such as institutional quality at the national level, royalties and taxation may be considered as driving the decision to open an oilfield. As these factors are the same within a country, we control for them by including year-country fixed effects, which also account for national changes in institutions and taxation policy during our period of study. The institutional quality and regulatory framework that may drive transnational oil corporations' decision to open an oilfield is determined at the national or regional level. Therefore, we also account for region-specific time trends in our robustness checks. Time-invariant differences in institutional quality across districts are accounted for by employing district fixed effects. Changes in institutional quality at the district level do not pose a serious threat to identification, as investment decisions of large transnational oil corporations are guided by the regulatory and legal framework at the national or regional level, not by differences in districts' institutional quality. Thus, while national and regional institutional quality and regulatory frameworks may influence the decision of oilfield opening, our regression specification accounts for these factors by including fixed effects and time trends.

Confounding Factors

A threat to internal validity would be the existence of a factor other than oil extraction that only affects our treated population, and hence drives a potential treatment effect. However, this is very unlikely to be the case due to the way in which treatment is defined in our study. The sample is divided into treatment and control based on the distance to oil deposits that turn operational during our study period; and the start-up of oil extraction defines the post-treatment group. A confounding factor posing a threat to internal validity would have to impact only the treatment group and not the control group. Additionally, it would have to mimic the timing of our treatment. It is highly unlikely that there is such a factor that mimics both the location and the timing of oil extraction, and hence only affects what we define as our treated population.

The difference-in-difference design eliminates the impact of confounding factors that affect the treatment and control group in the same way. However, it would pose a threat to internal validity if there are confounding factors that affect treatment and control groups differentially. We address this concern by limiting the geographical area of our control group to women living within 100 km of oil deposits. If, instead, we had used all observations in a country, control and treatment groups might be differentially impacted by political or economic crises or natural disasters. In the smaller geographic area we are considering, we do not expect factors such as economic crisis, coups, floods, droughts or storms to affect treatment and control areas in a systematically different way.

6 Results

The results from our empirical analysis are presented below. We will first report the results on women's employment and sectoral effects, and then present the results for women's empowerment. In each of the tables, we present results for our four specifications. Recall that the basic specification (1) is augmented by accounting for the construction phase in (2), (3) and (4). The construction phase is split up in early and late construction in (3) and (4), the latter of which additionally divides the post-treatment period into three distinct time periods. The main coefficient of interest in our difference-in-difference estimation is that on the interaction between *Active* and *Treated*, as it captures the effect on our post-treated treatment group.

6.1 Results for employment outcomes

Women's Employment

First, we examine the effect of local industrial development arising from oil extraction on women's employment. The results are reported in Table 2.

Without accounting for the construction phase (Table 2, column 1), we find that women living in the vicinity of active oilfields are 3.1 percentage points more likely to work. However, the estimated treatment effect in specification (1) may be biased, as it does not account for any employment effects on the treated women during the construction phase of oilfields. As elaborated in the empirical strategy, we thus augment our regression with interactions between the treated group and the construction phase (specifications (2) to (4)). After accounting for potential employment effects during construction, our coefficient of interest slightly decreases in magnitude. However, as the coefficient on *Active* $\times$ *Treated* is only marginally significant at 10%, we cannot conclude that oil extraction significantly changes women's probability to work. If existent, any local employment effect seems to be positive – however, we refrain from drawing any strong conclusions due to the low precision of our estimated coefficient of interest in specifications (2) and (3). Overall, the results from Table 2 indicate that women's likelihood to work does not seem to be significantly changed by the presence of oil extraction within 15 km.

Table 2: Total Employment

	(1)	(2)	(3)	(4)
	Working	Working	Working	Working
Treated (15 km)	-0.0360***	-0.0342**	-0.0342**	-0.0361**
	(0.0128)	(0.0156)	(0.0156)	(0.0157)
Active × Treated	0.0310**	0.0291*	0.0291*	
	(0.0145)	(0.0175)	(0.0175)	
Construction × Treated		-0.00318		
		(0.0169)		
Early Construction × Treated			-0.0117	-0.0129
			(0.0222)	(0.0223)
Late Construction × Treated			0.00405	0.00328
			(0.0191)	(0.0191)
Active × Treated (0-3y) (β_4)				0.0513**
				(0.0239)
Active × Treated (4-8y) (β_5)				0.0140
				(0.0202)
Active × Treated (post 8y) (β_6)				0.0211
				(0.0237)
N	97660	97660	97660	97660
R^2	0.258	0.258	0.258	0.258
mean	0.313	0.313	0.313	0.313
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6 = 0$)	-	-	-	1.55
p-value	-	-	-	0.198
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Sectoral Effects

While women's overall probability to work is not significantly changed, the local industrial development arising from oil extraction may still affect female employment in specific sectors. Therefore, we now turn to a sectoral analysis in order to test for any local employment effects in specific industries.

If the enclave hypothesis fails and oil extraction generates sufficiently strong spillovers to local communities, we would expect increases in women's service sector employment. This is confirmed by Table 3, which depicts the results for the service sector.

Table 3: Service Sector Employment

	(1) Services	(2) Services	(3) Services	(4) Services
Treated (15 km)	-0.00701	-0.0214*	-0.0213*	-0.0224*
	(0.00964)	(0.0118)	(0.0118)	(0.0119)
Active × Treated	0.0235**	0.0388***	0.0387***	
	(0.0109)	(0.0136)	(0.0136)	
Construction × Treated		0.0252*		
		(0.0136)		
Early Construction × Treated			0.00619	0.00612
			(0.0145)	(0.0145)
Late Construction × Treated			0.0412**	0.0422**
			(0.0173)	(0.0174)
Active × Treated (0-3y) (β_4)				0.0275
				(0.0176)
Active × Treated (4-8y) (β_5)				0.0297**
				(0.0143)
Active × Treated (post 8y) (β_6)				0.0638***
				(0.0196)
N	97660	97660	97660	97660
R^2	0.255	0.255	0.255	0.255
mean	0.214	0.214	0.214	0.214
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6 = 0$)	-	-	-	3.64
p-value	-	-	-	0.012
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6$)	-	-	-	2.1
p-value	-	-	-	0.122
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Specification (1) suggests that women living in the vicinity of oil extraction are 2.4 percentage points more likely to work in services. This effect is significant at the 5% level. However, this effect would be biased if women's service sector employment already changes during the construction phase. Once we account for potential effects during construction, the coefficient of interest increases in magnitude. This increase in magnitude indicates that the treatment effect in specification (1) may indeed be underestimated. Women's likelihood to work in services already increases during construction, as evident in the significant coefficient on the late construction phase in specification (3).

We find that the local industrial development stemming from oil extraction increases women's probability to work in the service sector by 3.9 percentage points. This effect is significant at the 1% level. Based on the sample average, this treatment effect corresponds to an 18% increase in women's likelihood to be employed in services. This employment effect on a non-resource sector supports the hypothesis of oil extraction inducing local multiplier effects, rather than being an enclave economy.

To test the assumption of oil extraction inducing a level shift, we estimate specification (4). The coefficients on the three periods of oil extraction ($\beta_4, \beta_5, \beta_6$) are jointly significant at 5%. Performing an F-test on the equality of these three coefficients, we fail to reject the null hypothesis of equality at the 10% level. Therefore, we cannot conclude that oil extraction has a differential impact in the short term relative to the medium or long term of our examined period. This lends support to the assumption that oil extraction induces a level shift in service sector employment, increasing women's probability to work in services by 18%.

For the agricultural sector, the coefficient of interest is not significant in any of the specifications (Table 4). Living in the vicinity of active oilfields does not change women's likelihood to work in agriculture. An additional analysis excluding agricultural self-employment confirms this result (see Appendix, Table 14); there are no significant changes in women's likelihood to be employed in agriculture. We observe a marginally significant negative impact on the likelihood to work in agriculture in the late construction phase (Table 4, specification (3)). This may be explained by some women switching from agricultural to service sector employment in the years immediately prior to oilfield opening, i.e. taking up the opportunity of less physical work.

Similar to the agricultural sector, our empirical analysis yields no strong evidence for local employment effects on the manual sector (Table 5). In specifications (1) to (3), the coefficient of interest is negative. However, it is only marginally significant at 10% and the lack of precision precludes any strong conclusions. If at all, women living in the vicinity of active oilfields seem to be less likely to work in the manual sector. One potential explanation may be that the increase in service sector employment opportunities induces women to switch from manual to service professions.

In summary, we find that the local industrial development stemming from oil extraction increases women's likelihood to work in services by 3.9 percentage points (18%). However, we cannot conclude that women are more likely to work overall. One potential explanation why the increase in service sector employment does not translate into an overall increase in women's probability to work is that some women may switch from the primary and secondary to the tertiary sector.

Table 4: Agricultural Sector Employment

	(1) Agriculture	(2) Agriculture	(3) Agriculture	(4) Agriculture
Treated (15 km)	-0.0310**	-0.0173	-0.0173	-0.0176
	(0.0122)	(0.0145)	(0.0145)	(0.0145)
Active × Treated	0.0150	0.000344	0.000356	
	(0.0128)	(0.0159)	(0.0160)	
Construction × Treated		-0.0241*		
		(0.0140)		
Early Construction × Treated			-0.0183	-0.0191
			(0.0188)	(0.0187)
Late Construction × Treated			-0.0290*	-0.0309**
			(0.0149)	(0.0148)
Active × Treated (0-3y)				0.0326
				(0.0211)
Active × Treated (4-8y)				-0.000600
				(0.0168)
Active × Treated (post 8y)				-0.0377*
				(0.0210)
N	97660	97660	97660	97660
R^2	0.206	0.206	0.206	0.206
mean	0.077	0.077	0.077	0.077
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 5: Manual Sector Employment

	(1) Manual	(2) Manual	(3) Manual	(4) Manual
Treated (15 km)	0.00216	0.00460	0.00459	0.00394
	(0.00394)	(0.00508)	(0.00507)	(0.00506)
Active × Treated	-0.00736*	-0.00996*	-0.00996*	
	(0.00383)	(0.00556)	(0.00555)	
Construction × Treated		-0.00429		
		(0.00682)		
Early Construction × Treated			0.000289	0.0000454
			(0.00948)	(0.00948)
Late Construction × Treated			-0.00815	-0.00804
			(0.00652)	(0.00653)
Active × Treated (0-3y) (β_4)				-0.00857
				(0.00652)
Active × Treated (4-8y) (β_5)				-0.0151**
				(0.00602)
Active × Treated (post 8y) (β_6)				-0.00509
				(0.00670)
N	97660	97660	97660	97660
R^2	0.025	0.025	0.025	0.025
mean	0.022	0.022	0.022	0.022
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6 = 0$)	-	-	-	2.54
p-value	-	-	-	0.054
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6$)	-	-	-	1.91
p-value	-	-	-	0.148
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Heterogeneous effects on employment

The sectoral composition of the economy differs in rural and urban settings. Naturally, agricultural employment is more common in rural areas. In our sample, 39% of working women in rural settings are engaged in agriculture, compared to only 5.8% of urban working women. In contrast, service sector employment is much higher in urban areas, where 86% of working women are employed in the service sector, compared to 54% of rural working women. The share of manual employment is 6.5% and 7.6% in rural and urban areas, respectively. The strength of local multiplier effects may vary between rural and urban areas. For example, oil workers may prefer consuming services in urban settings, which would suggest stronger spillover effects on women's employment in urban areas. Thus, we test whether the local employment effects from oil extraction differ between rural and urban areas.

We employ specification (6) to test for a differential impact on rural and urban women.

$$\begin{aligned} Y_{icdt} = \beta_0 &+ \beta_1 Treated_c * Rural_c + \beta_2 Treated_c * Urban_c + \beta_3 Treated_c * Early\ Con_{ct} * Rural_c + \\ &\beta_4 Treated_c * Early\ Con_{ct} * Urban_c + \beta_5 Treated_c * Late\ Con_{ct} * Rural_c + \\ &\beta_6 Treated_c * Late\ Con_{ct} * Urban_c + \beta_7 Active_{ct} * Treated_c * Rural_c + \\ &\beta_8 Active_{ct} * Treated_c * Urban_c + \beta_9 Rural_c + X_i + FE_{kt} + FE_d + \epsilon_{icdt} \end{aligned} \tag{6}$$

Table 6 depicts the coefficients of interest (β_7 and β_8), which are the interactions of the treatment group post-treatment with the rural and urban sample, respectively. As expected from our preceding analysis, the two coefficients of interest are not jointly significant for the agricultural and manual sectors. For working and services, the coefficients β_7 and β_8 are jointly significant at the 5% level.

Table 6: Total Employment - Differential Impact

	(1) Working	(2) Services	(3) Agriculture	(4) Manual
Active × Treated × Rural (β_7)	0.0147	0.0322**	-0.00771	-0.00974
	(0.0189)	(0.0137)	(0.0176)	(0.00637)
Active × Treated × Urban (β_8)	0.0948***	0.0574	0.0463	-0.00871
	(0.0361)	(0.0391)	(0.0316)	(0.00953)
N	97660	97660	97660	97660
R^2	0.258	0.255	0.206	0.025
F-test ($H_0 : \beta_7 = \beta_8 = 0$)	3.57	3.71	1.19	1.65
p-value	0.028	0.025	0.305	0.193
F-test ($H_0 : \beta_7 = \beta_8$)	4.25	0.38	-	-
p-value	0.039	0.538	-	-
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

We then perform an F-test to test for the equality of these coefficients. For services, we fail to reject the null hypothesis of equal coefficients. Thus, we cannot conclude that rural and urban women are differentially affected by oil extraction in their vicinity, supporting our earlier result of an 18% increase in the likelihood to work in services for the entire sample. In contrast, we reject the null hypothesis of equal coefficients at the 5% level for working. Thus, oil extraction seems to differentially affect rural and urban women's probability to work. All else equal, urban women living near active oilfields are eight percentage points (0.0948 – 0.0147 = 0.0801) more likely to work than rural women living near active oilfields. This is an interesting result refining our previous analysis, as it provides evidence for a heterogeneous local impact on rural and urban women's probability to work.

Having identified differential effects on total employment in rural and urban settings, we now estimate specification (3) considering the rural and urban sample separately. The results in Table 7 provide further insight on the heterogeneous effect of oil extraction on women's probability to work.

Table 7: Total Employment - Rural and Urban Sample

	(1) Working All	(2) Working Rural	(3) Working Urban	(4) Working All	(5) Working Rural	(6) Working Urban
Treated (15 km)	-0.0342**	-0.0292*	-0.0633	-0.0361**	-0.0308*	-0.0749
	(0.0156)	(0.0170)	(0.0449)	(0.0157)	(0.0169)	(0.0510)
Early Construction × Treated	-0.0117	-0.0262	0.0167	-0.0129	-0.0270	0.0171
	(0.0222)	(0.0251)	(0.0428)	(0.0223)	(0.0252)	(0.0429)
Late Construction × Treated	0.00405	-0.0130	0.0816**	0.00328	-0.0134	0.0817**
	(0.0191)	(0.0220)	(0.0334)	(0.0191)	(0.0219)	(0.0337)
Active × Treated	0.0291*	0.0257	0.115***			
	(0.0175)	(0.0196)	(0.0357)			
Active × Treated (0-3y) (β_4)				0.0513**	0.0479*	0.137**
				(0.0239)	(0.0257)	(0.0550)
Active × Treated (4-8y) (β_5)				0.0140	0.00659	0.0904*
				(0.0202)	(0.0229)	(0.0464)
Active × Treated (post 8y) (β_6)				0.0211	0.0240	0.115***
				(0.0237)	(0.0276)	(0.0421)
N	97660	55123	42537	97660	55123	42537
R^2	0.258	0.278	0.262	0.258	0.278	0.262
mean	0.313	0.311	0.315	0.313	0.311	0.315
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6$=0)	-	-	-	1.55	1.32	3.63
p-value	-	-	-	0.198	0.267	0.013
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6$)	-	-	-	-	-	0.26
p-value	-	-	-	-	-	0.768
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Restricting the sample to the urban population, we find that women living near active oilfields are 11.5 percentage points more likely to work (column 3). This effect is significant at the 1% level. Based on the sample mean, urban women's likelihood to work increases by 36.5%. Column 6 lends support to assuming a level shift – the three coefficients ($\beta_4, \beta_5, \beta_6$) are jointly significant at 5%, but we fail to reject the null hypothesis of equality of coefficients even at 10%.

Thus, while we cannot conclude that there is a local employment effect of oil extraction for the whole sample, urban women living near active oilfields are found to be more likely to work. Service sector employment is found to increase for the entire sample. Thus, rural women are also more likely to work in services, but not significantly more likely to work overall. One possible explanation is that services make up a smaller share of total employment in rural than in urban areas. Thus, an increase in service sector employment is more likely to translate into an increase in total employment in urban areas. Another potential explanation may be that some rural women switch from agriculture, which makes up almost 40% of rural employment, to less physically demanding service sector jobs.

6.2 Results for empowerment outcomes

As discussed in Section 5.1.2, we measure changes in women's empowerment by constructing two indices: the first index captures women's decision-making power in the household, the second index captures self-stated barriers in accessing healthcare. Results are presented in Table 8 and Table 9.

Table 8 presents the results for the self-stated barriers to healthcare. We find that women living in the vicinity of active oilfields do not experience any changes in barriers to healthcare. As in the analysis for empowerment outcomes, we divide the post-treatment period into three distinct time periods based on equality of observations, and interact each of them with *Treated*. As empowerment outcomes may take time to evolve, the division of the post-treatment period into three different time periods allows testing for a lagged effect on the outcomes we study. We do not observe significant changes in barriers to healthcare in any of the specifications from (1) to (4). This confirms that there is no treatment effect, not even with a time lag.

Table 9 represents the results for decision-making power in the household. Recall that our index combines three variables capturing decision-making regarding one's own health, large household purchases and visiting family and relatives. We do not find that women living in the vicinity of active oilfields experience any changes in their decision-making power. The coefficient of interest in all specifications is statistically insignificant. Specification (4) further disproves the notion of any delayed treatment effect.

In summary, we do not find any significant change in women's decision-making power in the household or in their self-stated barriers to healthcare. Thus, we cannot conclude that the local development arising from oil extraction changes women's empowerment, which will be discussed further in Section 8.

Table 8: Self-stated Barriers to Healthcare

	(1) Barriers to Healthcare	(2) Barriers to Healthcare	(3) Barriers to Healthcare	(4) Barriers to Healthcare
Treated (15 km)	-0.0177	0.00104	-0.000255	0.000961
	(0.0172)	(0.0270)	(0.0270)	(0.0272)
Active × Treated	0.0172	-0.00211	-0.000926	
	(0.0160)	(0.0258)	(0.0258)	
Construction × Treated		-0.0272		
		(0.0230)		
Early Construction × Treated			-0.00739	-0.00902
			(0.0255)	(0.0253)
Late Construction × Treated			-0.0406	-0.0417*
			(0.0251)	(0.0249)
Active × Treated (0-3y)				0.0206
				(0.0281)
Active × Treated (4-8y)				-0.00566
				(0.0283)
Active × Treated (post 8y)				-0.0297
				(0.0287)
N	70735	70735	70735	70735
R^2	0.187	0.187	0.188	0.188
mean	0.242	0.242	0.242	0.242
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 9: Decision-making in the Household

	(1) Decision making	(2) Decision making	(3) Decision making	(4) Decision making
Treated (15 km)	-0.0134	-0.00253	-0.00269	-0.00225
	(0.0176)	(0.0279)	(0.0279)	(0.0280)
Active × Treated	-0.00191	-0.0131	-0.0130	
	(0.0176)	(0.0279)	(0.0279)	
Construction × Treated		-0.0163		
		(0.0269)		
Early Construction × Treated			-0.0112	-0.00950
			(0.0321)	(0.0321)
Late Construction × Treated			-0.0203	-0.0194
			(0.0286)	(0.0285)
Active × Treated (0-3y)				-0.0327
				(0.0303)
Active × Treated (4-8y)				-0.00262
				(0.0302)
Active × Treated (post 8y)				0.00170
				(0.0311)
N	62198	62198	62198	62198
R^2	0.155	0.155	0.155	0.155
mean	0.685	0.685	0.685	0.685
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

7 Robustness

Our baseline estimation involves the choice of a treatment radius of 15 km and includes the entire sample, irrespective of migration status. In this section, we employ three robustness checks. First, we test for a potential treatment effect in case of a different treatment radius in order to examine whether we have indeed estimated a highly localised treatment effect. Second, we test whether our results are robust to excluding migrants. Third, we allow for a regional time trend to relax the assumption that time-varying unobservables at the regional level, such as institutions, do not significantly impact our results.

7.1 Varying the treatment radius

As a first robustness check, we change the radius that classifies the DHS clusters into treatment and control groups by replicating the analysis for a treatment radius of 25 km. While the choice of our baseline treatment radius is guided by the literature, it cannot be ruled out that some treatment effect may persist beyond the 15 km radius. We therefore test the employment outcomes for a treatment radius of 25 km in order to ascertain whether the treatment effects found in our baseline analysis persist beyond 15 km. The treatment group is now defined as living within 25 km from an oil deposit, whereas the control group is defined as being within 25 to 100 km. Intuitively, we would expect this to attenuate the treatment effect that we obtained for the 15 km radius, because a part of the former control group now belongs to the treated group. We would therefore expect a non-existent or weaker treatment effect for our employment outcomes.

Results for overall employment and service sector employment from changing the treatment radius to 25 km are presented in Table 10 and Table 11, respectively. Compared to the baseline results for working, the coefficient of interest declines in magnitude in each specification. When accounting for the construction phase, the estimated treated effect declines from 2.9 percentage points in the baseline to 2.7 percentage points, and remains only marginally significant at 10%. Due to the low precision of the estimated treatment, we again cannot conclude that living next to active oilfields increases women's probability to work, which is consistent with our expectations.

We next turn to the sectoral analysis. For agriculture and manual work, the coefficient of interest is not statistically significant at 5% (see Appendix, Table 16 and Table 17). The absence of a treatment effect is consistent with the results from the 15 km treatment radius.

For service sector employment, the difference in the estimated treatment effect between the 15 km and 25 km radius is starkly apparent. The coefficient of interest strongly declines in magnitude and is not statistically significant in any of the specifications in Table 11. This confirms that the treatment effect on services that we observe in our baseline specification is, in fact, highly localised and does not extend to a larger area. This supports our hypothesis that the treatment effect wanes as the distance from the oilfield increases. This is further confirmed by testing for a treatment effect employing a much larger treatment radius of 50 km, where we also find no significant treatment effect on service sector employment (see Appendix, Table 18).

Table 10: Total Employment (25 km radius)

	(1)	(2)	(3)	(4)
	Working	Working	Working	Working
Treated (25 km)	-0.00845	-0.00709	-0.00714	-0.00881
	(0.0135)	(0.0152)	(0.0152)	(0.0152)
Active × Treated	0.0288***	0.0272*	0.0271*	
	(0.0109)	(0.0142)	(0.0142)	
Construction × Treated		-0.00235		
		(0.0138)		
Early Construction × Treated			-0.00652	-0.00570
			(0.0201)	(0.0200)
Late Construction × Treated			0.000121	0.000493
			(0.0150)	(0.0150)
Active × Treated (0-3y)				0.0303*
				(0.0170)
Active × Treated (4-8y)				0.0182
				(0.0154)
Active × Treated (post 8y)				0.0446**
				(0.0180)
N	95422	95422	95422	95422
R^2	0.250	0.250	0.250	0.250
mean	0.306	0.306	0.306	0.306
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 11: Service Sector Employment (25 km radius)

	(1) Services	(2) Services	(3) Services	(4) Services
Treated (25 km)	-0.00105	0.000346	0.000179	-0.00194
	(0.00887)	(0.00981)	(0.00981)	(0.00977)
Active × Treated	0.00924	0.00764	0.00713	
	(0.00805)	(0.00989)	(0.00987)	
Construction × Treated		-0.00242		
		(0.00948)		
Early Construction × Treated			-0.0166	-0.0170
			(0.0128)	(0.0128)
Late Construction × Treated			0.00598	0.00529
			(0.0107)	(0.0108)
Active × Treated (0-3y)				0.0175
				(0.0121)
Active × Treated (4-8y)				-0.00812
				(0.0110)
Active × Treated (post 8y)				0.0229*
				(0.0133)
N	95422	95422	95422	95422
R^2	0.252	0.252	0.252	0.252
mean	0.210	0.210	0.210	0.210
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

7.2 Excluding migrants

Our second robustness check excludes any migrant women. We conduct our baseline analysis for a subset of the local population, the so-called “never-movers”. These are women who have always lived in the city, village or town of their current residence. An increase in economic activity due to the opening of an oilfield may induce people to move into the vicinity of oilfields. It would not pose a threat to our analysis if women move within the treatment radius, for example due to marriage. Such women would always be a part of the treated group. However, if the expectation of better economic prospects induces women to move into the treatment radius, it can be argued that such women are fundamentally different from the originally local population (e.g. in terms of education or experience), and may drive the results we obtain.

To understand if oil extraction benefits the originally local population, we would ideally exclude all women who migrated from a control to a treatment area. For such an analysis, we would need information on the previous location of all women who migrated in order to segregate between women that moved within the treatment area and women who moved into the treatment area. However, the DHS data does not include information on women’s exact previous location of residence and hence we cannot conduct the desired analysis.

There is, however, another variable in the DHS dataset that denotes the years lived in the current place of residence (city, village or town). While this does not enable us to clearly identify the migrant women moving from control into treatment, the variable allows us to identify a specific sub-group of originally local women, which we refer to as “never-movers”. These women have lived all their lives in the current place of residence. In order to test whether our baseline results hold for this subset of the originally local population, we now exclude all migrants by restricting the sample to the never-movers. Our sample size now shrinks from 97,660 to 43,007.

We test whether the positive treatment effect we found on service sector employment also holds for the never-movers. The results are displayed in Table 12. We find a strong positive treatment effect on service sector employment for never-movers. After accounting for the construction phase, the likelihood of engaging in services increases by 4.2 percentage points (27%). Thus, the magnitude of the treatment effect is higher for never-movers than for the whole sample (18%), which confirms that the gains in service sector employment are not driven by migrants. Conducting an F-test on the three separate post-treatment coefficients ($\beta_4, \beta_5, \beta_6$ in specification (4)), we find that they are jointly significant at 5%. We further test for equality of these three coefficients and fail to reject the null hypothesis that the coefficients are equal to each other at 10%. This lends support to the assumption of a level shift in service sector employment. In other words, the increase in service sector employment also seems to be persistent for women who have never moved.

Consistent with the baseline results including the full sample, the never-movers do not experience a significant change in overall employment, agriculture and manual employment (results in Appendix: Table 19, Table 20 and Table 21). Thus, the robustness check of restricting the sample to never-movers confirms that our baseline results are not driven by migrants.

Table 12: Service Sector Employment (never-movers)

	(1) Services	(2) Services	(3) Services	(4) Services
Treated (15 km)	-0.00318	-0.0249**	-0.0255**	-0.0268**
	(0.0101)	(0.0121)	(0.0121)	(0.0122)
Active × Treated	0.0166	0.0410***	0.0418***	
	(0.0123)	(0.0142)	(0.0142)	
Construction × Treated		0.0389***		
		(0.0134)		
Early Construction × Treated			0.0289*	0.0279*
			(0.0158)	(0.0158)
Late Construction × Treated			0.0494***	0.0486***
			(0.0172)	(0.0171)
Active × Treated (0-3y) (β_4)				0.0507***
				(0.0180)
Active × Treated (4-8y) (β_5)				0.0309*
				(0.0170)
Active × Treated (post 8y) (β_6)				0.0307
				(0.0212)
N	43007	43007	43007	43007
R^2	0.241	0.241	0.241	0.241
mean	0.152	0.152	0.152	0.152
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6 = 0$)	-	-	-	2.99
p-value	-	-	-	0.029
F-statistic ($H_0 : \beta_4 = \beta_5 = \beta_6$)	-	-	-	0.52
p-value	-	-	-	0.593
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

7.3 Region-specific time trends

Accounting for country-specific changes in each year in our main analysis, we find that women's likelihood to be employed in the service sector increases by 3.9 percentage points (18%). However, it may be argued that changes in factors at the regional level, such as slowly evolving institutions rather than oil extraction, may drive local economic development and hence our results. Such time-varying unobservables are not captured by the district fixed effects that we include. In order to verify that the gains in service sector employment are not driven by time-varying unobservables at the regional level, we augment our specification with region-specific time trends. A region is defined by the GADM level 1 data. Our sample includes clusters in 57 different regions. By allowing for region-specific time trends, we account for more variation in trends than in our baseline specification.

Results from the inclusion of region-specific time trends for service sector employment are reported in Table 13. The coefficient of interest in specification (3) drops slightly in magnitude, from 0.039 (Table 3) to 0.031 (Table 13), but retains statistical significance at 5%. Thus, the positive treatment effect of oil extraction on women's service sector employment is robust to accounting for region-specific time trends.

Table 13: Service Sector Employment (including regional time trend)

	(1) Services	(2) Services	(3) Services	(4) Services
Treated (15 km)	-0.00356	-0.0193	-0.0183	-0.0186
	(0.00971)	(0.0128)	(0.0128)	(0.0128)
Active × Treated	0.0155	0.0322**	0.0312**	
	(0.0113)	(0.0147)	(0.0147)	
Construction × Treated		0.0258*		
		(0.0148)		
Early Construction × Treated			0.0100	0.0115
			(0.0168)	(0.0169)
Late Construction × Treated			0.0375**	0.0392**
			(0.0179)	(0.0181)
Active × Treated (0 - 3y)				0.0199
				(0.0182)
Active × Treated (4 - 8y)				0.0367**
				(0.0159)
Active × Treated (post 8y)				0.0418*
				(0.0224)
N	97660	97660	97660	97660
R^2	0.258	0.258	0.258	0.258
mean	0.214	0.214	0.214	0.214
Country-Year FE	No	No	No	No
Regional Time Trend	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

8 Discussion

8.1 Contribution and comparison to previous findings

Oil extraction may be a curse or a blessing for local communities. Exploiting spatial and temporal variation in the opening of oilfields, we explore the local effect of oil extraction on women's employment and empowerment in Africa. The contribution of our empirical analysis is three-fold.

First, analysing a novel dataset that combines detailed oilfield data from Rystad with geo-coded DHS survey data, we add to the empirical literature on the local effect of oil. Second, by focusing exclusively on women and examining their employment and empowerment, we provide insight into gender-specific effects of oil extraction, a topic that has long been neglected when studying the impact of natural resources. Third, the developing country context of our study addresses a gap in the literature, as most local impact studies on oil have been conducted in developed countries, where the impact of local economic development arising from oil production in nearby communities is likely to differ.

We find that women living near active oilfields are 18% more likely to be employed in the service sector. This provides evidence against the common portrayal of the oil industry as an enclave without any linkages to local communities. The increase in employment in a non-resource sector supports the existence of sufficiently strong local multiplier effects from oil extraction. Oil extraction seems to increase the demand for local services, thereby generating indirect employment opportunities for women. Examples of service sector professions recorded in the DHS data include street vendors, cashiers, shop assistants, receptionists, restaurant staff, bar staff, administrative and office support, cleaners, governesses, child care or teaching staff. Our estimated treatment effect thus suggests that the demand for female employment in such service professions increases in the vicinity of active oilfields. This is in line with the structural change literature, which posits that economic development induces a shift from primary to tertiary sector employment.

However, we note that the magnitude of the estimated increase in service sector employment is lower than in comparable studies on the local impact of other natural resources. Large-scale mining in Africa has been found to increase female service sector employment by 50% (Kotsadam and Tolonen, 2016), which suggests that the multiplier effects on services are weaker in oil communities than in mining communities. The lower magnitude of the local employment impact of oil on services is in line with the differences between oil and mineral extraction. As elaborated in Section 3.2, oil extraction is even more capital intensive than mining, resulting in less direct employment, which in turn may generate less indirect employment in non-resource sectors. Additionally, while Kotsadam and Tolonen only consider large-scale mines, our study is not limited to giant oilfields, which may also explain the relatively lower magnitude of the treatment effect on services.

In contrast, the negative employment effect on agriculture found for large-scale mining in Africa (Kotsadam and Tolonen, 2016) is not evident in oil communities. Both mining and oil have been found to have negative environmental consequences, which decrease agricultural productivity (e.g. Aragón and Rud, 2015; Kadafa, 2012). However, we do not observe a significant reduction employment in the agricultural sector. One potential explanation is that mining operations are entirely land-based, covering a larger area than oil extraction sites (Cameron and Stanley, 2017) and hence are more likely to directly impact areas under cultivation. Another explanation is that air pollution and the resulting acid rain arising from mining may be even more detrimental to agricultural productivity in comparison to oil spills, which are regarded as the most serious environmental consequence of oil extraction in developing countries. However, a localised analysis estimating the effect of oil spills on agricultural employment would be needed to shed further light on this question. We would expect agricultural employment to decrease in areas that are most affected by oil spills.

We find that overall female employment is not significantly affected by the local industrial development stemming from oil extraction. This indicates that the local multiplier effects vary by sector, as only the service sector experiences increases in employment. However, restricting the sample to urban women, we find that they are significantly more likely to work. Intuitively, this can be explained by services making up a larger share of total female employment in urban areas. In contrast to our finding of no significant change in the entire sample's probability to work, Kotsadam and Tolonen (2016) find that overall female employment decreases by 8% in the context of large-scale mining in Africa. In their study, this effect is driven by the decrease in agriculture. As we do not observe a significant decrease in the agricultural sector that would offset increases in services, we do not find a reduction in overall employment. Thus, oil extraction does not drive women in nearby communities out of the labour market, as one might have expected from earlier studies establishing a negative effect of oil on female employment on the national level (Ross, 2008).

However, we find no impact on women's empowerment. A priori, the effect of oil extraction on empowerment is unclear. Positive changes would be expected through channels such as increased female employment, and hence income, or through better access and larger exposure to media. On the other hand, if local conflict and violence increase with oil extraction, one may also expect that women become more constrained in their security and ability to make choices. The fact that we find no change in women's empowerment may suggest that these countervailing forces cancel each other out, leaving the empowerment outcomes we study unchanged.

There are few comparable studies focusing on women's empowerment, but Tolonen (2019) finds a positive local impact on women's empowerment and service sector employment (31% increase) in the context of gold mining in sub-Saharan Africa. Tolonen does not ascertain causality between employment and empowerment outcomes. However, if service sector employment is an important channel that affects women's empowerment, then the comparatively modest gains in service employment in our analysis (18% increase) may explain the absence of improvement in empowerment outcomes. Alternatively, women's empowerment may be affected through the channel of increased media exposure, for which Tolonen (2019) finds evidence in the context of gold mining. Even though we do not find any significant effect on empowerment, we test whether the channel of media exposure changes for treated women. Again, we find no significant treatment effect (see Appendix, Table 15). The absence of a change in media exposure, which has been established as an important channel impacting women's empowerment (Chong and La Ferrara, 2009; Jensen and Oster, 2009), may thus also explain why we do not observe any significant changes in women's decision-making power and self-stated barriers to healthcare.

Multiple Hypothesis Testing

The two empowerment outcomes we test are indices, which we construct in order to reduce multiple inference problems. The risk of observing significant results by chance increases with the number of hypotheses tested - therefore, we refrain from testing the six relevant empowerment questions separately. An alternative method to address problems arising from multiple hypothesis testing is to use Bonferroni-corrected p-values. This correction redefines the cut-off level of statistical significance by dividing the desired significance level by the number of hypotheses tested. If we had decided to test for each of the utilised decision-making and barriers to healthcare variables separately, we would have divided our desired significance level of 5% by the six hypotheses tested, resulting in an adjusted p-value of 0.008.

However, if testing each question separately had been our preferred approach, we could have also included relevant empowerment questions from other DHS survey modules, such as women's decision-making power on contraception, and questions with fewer observations, such as decision-making power on husbands' earnings. Due to the limited overlap in observations, we refrained from including these questions in our empowerment indices. Considering each question separately would have resulted in testing eight decision-making outcomes and six barriers to healthcare outcomes, resulting in a total of 14 hypotheses tested in the empowerment section. This large number of hypotheses would have necessitated the use of the Bonferroni-corrected p-value of 0.0036 (0.05/14=0.0036). However, we do not consider it warranted to examine 14 separate outcomes in the realm of women's empowerment. Our definition of empowerment is already narrow, and the separate examination of 14 sub-questions would not allow any meaningful conclusion on women's empowerment. We do not aim to establish the local effect of oil extraction on decision-making in one specific aspect of life, e.g. family visits and therefore do not test each outcome separately.

8.2 External validity

Our results cannot be extended to developed countries. First, the channels impacting women's employment and their decision to work differ significantly between developed and developing countries. Second, prevalent societal gender norms are also different between the African countries in our study and developed countries. For example, "having to ask for permission" to access healthcare would not constitute a relevant outcome in women's empowerment in most developed countries. To explore the effect of oil extraction on women's empowerment in developed countries, different outcomes would have to be studied due to differences in prevalent gender norms and obstacles to empowerment. Thus, the underlying differences in female employment and gender norms, as well as the channels impacting these outcomes, constrain the generalisability of our results.

Further, our analysis only considers onshore oilfields. Therefore, our results are not applicable in the context of offshore oilfields. Notably, offshore oilfields employ fewer people and are more isolated from local communities by the virtue of being away from land. Thus, the ability of offshore oilfields to generate linkages and multiplier effects to nearby local economies is naturally limited and hence different from onshore oilfields. Finally, our results may not entirely be extended to other natural resources, especially if the resources systematically vary in terms of local reliance on infrastructure, capital intensity and their ability to generate linkages to the local economy.

We expect our study's findings to have external validity for other oil-rich countries in Africa, for example Niger or Congo, which are engaged in conventional onshore oil production. Further, some African countries, such as Kenya, have recently discovered oil resources. Our findings may thus also be relevant for African countries starting onshore oil production in the future.

9 Conclusion

Our paper examines the local impact of oil extraction on women's employment and empowerment. Specifically, we examine whether living in the vicinity of active oilfields significantly impacts women's occupational status and female empowerment. We use the opening of an oilfield as a plausibly exogenous shock to the local economic area. The capital-intensive and globalised nature of the oil industry, relatively low reliance on local infrastructure including transportation and little dependence on local inputs support the validity of the identifying assumption. Further, district fixed effects control for any time-invariant local characteristics, such as the level of local economic development or distance to the nearest city or port, which may affect the decision to invest in a particular district. Additionally, region-specific time trends employed in our robustness analysis control for any time-varying factors, such as institutions, which may change slowly at the regional level and hence affect transnational oil corporations' decision to invest in a particular region.

Our main results are three-fold: First, we find that the local industrial development stemming from oil extraction increases women's probability of service sector employment by 18%. This effect is robust to excluding migrants and to accounting for region-specific time trends. This suggests that oil extraction cannot be considered as an enclave economy, as there are local multiplier effects on women's service sector employment. However, this local employment effect is relatively modest in magnitude; the increase in women's service sector employment is found to be smaller than in comparable studies examining the local impact of large-scale mining in Africa. Second, there is no change in women's empowerment. The local development from oil extraction neither affects women's decision-making power within the household nor their self-stated barriers in accessing healthcare. While our study does not aim to establish causality between employment and empowerment, it might be the case that the modest gains in terms of service sector employment alone are insufficient to translate into enhanced empowerment. Third, our results indeed capture a highly localised effect, as the positive impact on service sector employment vanishes when increasing the treatment radius to 25 km

Finally, although this paper identifies local spillover effects of oil extraction on female service sector employment, the aggregate welfare effects remain to be determined. A full welfare analysis taking into account additional outcomes such as income, health and conflict would be necessary to provide a more comprehensive picture on the overall local impact of oil extraction on women in Africa.

References

Ackah-Baidoo, A. (2013). Fishing in troubled waters: Oil production, seaweed and community-level grievances in the Western Region of Ghana. *Community Development Journal, 48*(3), 406-420.

Adusah-Karikari, A. (2015). Black gold in Ghana: Changing livelihoods for women in communities affected by oil production. *The Extractive Industries and Society, 2*(1), 24-32.

Agerton, M., Hartley, P., Medlock, K., & Temzelides, T. (2015). Employment impacts of upstream oil and gas investment in the United States. IMF Working Paper WP/15/28.

Aidoo, E. N., Amoh-Gyimah, R., & Ackaah, W. (2013). The effect of road and environmental characteristics on pedestrian hit-and-run accidents in Ghana. *Accident Analysis & Prevention, 53*, 23-27.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.

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

Aragón, F. M., & Rud, J. P. (2015). Polluting industries and agricultural productivity: Evidence from mining in Ghana. *The Economic Journal, 126*(597), 1980-2011.

Auty, R. M. (1994). Industrial policy reform in six large newly industrializing countries: The resource curse thesis. *World Development, 22*(1), 11-26.

Basu, K. (2006). Gender and say: A model of household behaviour with endogenously determined balance of power. *The Economic Journal, 116*(511), 558-580.

Benshaul-Tolonen, A., & Baum, S. (2019). Structural transformation, extractive industries and gender equality. Available at *SSRN* 3464290.

Berman, N., Couttenier, M., Rohner, D., & Thoenig, M. (2017). This mine is mine! How minerals fuel conflicts in Africa. *American Economic Review, 107*(6), 1564-1610.

Black, D., McKinnish, T., & Sanders, S. (2005). The economic impact of the coal boom and bust. *The Economic Journal, 115*(502), 444–471.

Borthwick, I., Balkau, F., Read, T., & Monopolis, J. (1997). *Environmental management in oil and gas exploration and production* (Report No. 37, 4-7). UNEP Technical Publication. Retrieved from https://wedocs.unep.org

Boserup, E. (1970). *The role of women in economic development.* New York: St. Martin's.

Bruederle, A., & Hodler, R. (2019). Effect of oil spills on infant mortality in Nigeria. *Proceedings of the National Academy of Sciences, 116*(12), 5467-5471.

Cameron, P. D., & Stanley, M. C. (2017). *Oil, gas, and mining: A sourcebook for understanding the extractive industries.* The World Bank. Retrieved from https://openknowledge.worldbank.org/handle/10986/26130

Chen, Y., Jaupart, P., Moreno-Monroy, A., & Picarelli, N. (2017). *Unequal commutes: Job accessibility and employment in Accra* (Report No. C-33401-GHA-1). International Growth Centre.

Cheung, M. (2012). *Edutainment radio, women's status and primary school participation: Evidence from Cambodia* (No. 2012:5). Stockholm University, Department of Economics.

Chong, A., & Ferrara, E. L. (2009). Television and divorce: Evidence from Brazilian novelas. *Journal of the European Economic Association, 7*(2-3), 458-468.

Cust, J., & Poelhekke, S. (2015). The local economic impacts of natural resource extraction. *Annual Review of Resource Economics, 7*(1), 251-268.

Darko, E. (2014). *Short guide summarising the oil and gas industry lifecycle for a non-technical audience.* London: Overseas Development Institute.

DHS Program. (2018). *Description of the demographic and health surveys individual recode data file.* Retrieved from https://dhsprogram.com/publications/publication-DHSG4-DHS-Questionnaires-and-Manuals.cfm

Dietsche, E., Dodd, S., Haglund, D., Henstridge, M., Jakobsen, M., Sindou, E., & Slaven, C. (2013). *Extractive industries, development and the role of donors.* London: Oxford Policy Management. Retrieved from https://assets.publishing.service.gov.uk/media/57a08a21e5274a31e0000440/Extractive-industries-development-and-the-role-of-donors.pdf

Dube, O., & Vargas, J. F. (2013). Commodity price shocks and civil conflict: Evidence from Colombia. *The Review of Economic Studies, 80*(4), 1384-1421.

Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa. *The World Bank Economic Review, 17*(1), 1–25.

Duflo, E. (2012). Women's empowerment and economic development. *Journal of Economic Literature, 50*(4), 1051–1079.

Eggert, R. G. (2001). *Mining and economic sustainability: National economies and local communities.* Colorado School of Mines. Retrieved from https://pubs.iied.org/pdfs/G00952.pdf

Ehab, M. (2018). *The commuting gender gap and females' labor supply and earnings in the Egyptian labor market* (Working Paper No.1211). Economic Research Forum. Retrieved from http://erf.org.eg/wp-content/uploads/2018/06/1211˙Final.pdf

EIGE. (2019). *Empowerment of women.* European Institute for Gender Equality. Retrieved from https://eige.europa.eu/thesaurus/terms/1102

Fajana, S. (2005). *Industrial relations in the oil industry in Nigeria* (No. 993857973402676). International Labour Organization. Retrieved from https://www.ilo.org/sector/Resources/publications/WCMS˙161189/lang–en/index.htm

Farré, L., Klein, R., & Vella, F. (2012). Does increasing parents' schooling raise the schooling of the next generation? Evidence based on conditional second moments. *Oxford Bulletin of Economics and Statistics, 74*(5), 676-690.

Ferguson, J. (2005). Seeing like an oil company: space, security, and global capital in neoliberal Africa. *American Anthropologist, 107*(3), 377-382.

Fetzer, T. (2014). *Fracking growth* (CEP Discussion Paper No. 1278). Centre for Economic Performance, London School of Economics and Political Science.

Frankel, J. A. (2010). *The natural resource curse: A survey* (NBER Working Paper w15836). National Bureau of Economic Research.

Harding, T., & Venables, A. J. (2010). *Foreign exchange windfalls, imports and exports. Unpublished*. University of Oxford.

Heath, R., & Jayachandran, S. (2016). *The causes and consequences of increased female education and labor force participation in developing countries* (NBER Working Paper w22766). National Bureau of Economic Research.

Heath, R., & Mobarak, M. (2012). *Supply and demand constraints on educational investment: Evidence from garment sector jobs and the female stipend program in Bangladesh* (Working Paper). New Haven, CT: Yale School of Management.

Hirschman, A. O. (1958). *The strategy of economic development*. New Haven: Yale University Press.

Horn, M. K. (2004). *Giant fields 1868–2004* (CD-ROM). Version 1.2 (revision to 2003 version), AAPG Datapages.

Hutchison, M. M. (1994). Manufacturing sector resiliency to energy booms: Empirical evidence from Norway, the Netherlands, and the United Kingdom. *Oxford Economic Papers*, 311-329.

Jayasundara, D., Legerski, E., & Danis, F. S. (2018). The impact of oil development on sexual assault in the Bakken. *Family & Intimate Partner Violence Quarterly, 10*(4), 63-72.

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *The Quarterly Journal of Economics, 127*(2), 753-792.

Jensen, R., & Oster, E. (2009). The power of TV: Cable television and women's status in India. *The Quarterly Journal of Economics, 124*(3), 1057-1094.

Kadafa, A. A. (2012). Environmental impacts of oil exploration and exploitation in the Niger Delta of Nigeria. *Global Journal of Science Frontier Research Environment & Earth Sciences, 12*(3), 19-28.

Kojima, M., Matthews, W., & Sexsmith, F. (2010). *Petroleum markets in sub-Saharan Africa: Analysis and assessment of 12 countries* (Extractive industries and development series No. 15). Washington, DC: World Bank. Retrieved from http://documents.worldbank.org/curated/en/907171467990361424/Petroleum-markets-in-Sub-Saharan-Africa-analysis-and-assessment-of-12-countries

Kotsadam, A., & Tolonen, A. (2016). African mining, gender, and local employment. *World Development*, 83, 325-339.

Kung, K. S., Greco, K., Sobolevsky, S., & Ratti, C. (2014). Exploring universal patterns in human home-work commuting from mobile phone data. *PloS one, 9*(6), e96180.

La Ferrara, E., Chong, A., & Duryea, S. (2012). Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics, 4*(4), 1-31.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics, 4*(3), 165-224.

Lee, J. (2015). The regional economic impact of oil and gas extraction in Texas. *Energy Policy, 87*, 60–71.

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

Liou, Y. M., & Musgrave, P. (2016). Oil, autocratic survival, and the gendered resource curse: When inefficient policy is politically expedient. *International Studies Quarterly, 60*(3), 440-456.

Lippert, A. (2014). *Spill-overs of a resource boom: Evidence from Zambian copper mines* (No. 131). Oxford Centre for the Analysis of Resource Rich Economies, University of Oxford.

Loayza, N., & Rigolini, J. (2016). The local impact of mining on poverty and inequality: Evidence from the commodity boom in Peru. *World Development, 84*, 219-234.

Mammen, K., & Paxson, C. (2000). Women's work and economic development. *Journal of Economic Perspectives, 14*(4), 141-164.

Marchand, J. (2012). Local labor market impacts of energy boom-bust-boom in Western Canada. *Journal of Urban Economics, 71*(1), 165–174.

Marchand, J., & Weber, J. (2018). Local labor markets and natural resources: A synthesis of the literature. *Journal of Economic Surveys, 32*(2), 469-490.

Maurer, S., & Potlogea, A. (2017). *Male-biased demand shocks and women's labor force participation: Evidence from large oil field discoveries* (Working Paper). Universität Konstanz. Retrieved from
http://www.uni-konstanz.de/FuF/wiwi/workingpaperseries/WP˙08˙Maurer˙Potlogea˙2017.pdf

McKenzie, D. (2017). *When should you cluster standard errors? New wisdom from the econometrics oracle.* World Bank Blogs. Retrieved from
https://blogs.worldbank.org/impactevaluations/when-should-you-cluster-standard-errors-new-wisdom-econometrics-oracle

Meili, C., Jungbluth, N., & Annaheim, J. (2018). *Life cycle inventories of crude oil extraction.* ESU-services Ltd. commissioned by BFE, BAFU, Erdöl-Vereinigung, Schaffhausen, Switzerland. Retrieved from www.esu-services.ch/data/public-lci-reports

Moretti, E. (2010). Local multipliers. *American Economic Review, 100*(2), 373-77.

Neto, R. S., Duarte, G., & Páez, A. (2015). Gender and commuting time in São Paulo metropolitan region. *Urban Studies, 52*(2), 298-313.

Obeng-Odoom, F. (2014). Oil, sex, and temporary migration: The case of Vienna City, Sekondi-Takoradi, Ghana. *The Extractive Industries and Society, 1*(1), 69-74.

Ogunbodede, E. F. (2008). Urban road transportation in Nigeria from 1960 to 2006: Problems, prospects and challenges. *Ethiopian Journal of Environmental Studies and Management, 1*(1), 7-18.

Oyewole, B. (2018). *Overview of local content regulatory frameworks in selected ECCAS countries*. United Nations Conference on Trade and Development (UNCTAD), Geneva. Retrieved from https://unctad.org/en/PublicationsLibrary/ditccominf2018d4˙en.pdf

Ross, M. L. (2008). Oil, islam, and women. *American Political Science Review, 102*(1), 107-123.

Sachs, J. D., & Warner, A. M. (1995). *Natural resource abundance and economic growth* (NBER Working Paper w5398). National Bureau of Economic Research.

Sebastián, M. S., Armstrong, B., & Stephens, C. (2002). Outcomes of pregnancy among women living in the proximity of oil fields in the Amazon basin of Ecuador. *International Journal of Occupational and Environmental Health, 8*(4), 312-319.

Sen, A. (1990). More than 100 million women are missing. *The New York Review of Books, 37*(20), 61-66.

Söderling, L. (2006). After the oil: Challenges ahead in Gabon. *Journal of African Economies, 15*(1), 117-148.

Spatafora, N., & Warner, A. (1999). *Macroeconomic effects of terms-of-trade shocks: the case of oil-exporting countries*. The World Bank. Retrieved from https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-1410

Tolonen, A. (2018). Local industrial shocks and infant mortality. *The Economic Journal, 129*(620), 1561-1592.

Tolonen, A. (2019). Endogenous gender norms: Evidence from Africa's gold mining industry. Available at *SSRN* 3284519.

UNDP. (2017). *Annex 1 to the UNDP gender equality strategy 2014 - 2017: The future we want: Rights and empowerment*. United Nations Development Programme. Retrieved from https://www.undp.org/content/dam/undp/library/gender/Annex%201%20Terminology.pdf

Van der Ploeg, F. (2011). Natural resources: curse or blessing?. *Journal of Economic Literature, 49*(2), 366-420.

Van der Ploeg, F., & Poelhekke, S. (2017). The impact of natural resources: Survey of recent quantitative evidence. *The Journal of Development Studies, 53*(2), 205-216.

Vicente, P. C. (2010). Does oil corrupt? Evidence from a natural experiment in West Africa. *Journal of Development Economics, 92*(1), 28-38.

Von der Goltz, J., & Barnwal, P. (2019). Mines: The local wealth and health effects of mineral mining in developing countries. *Journal of Development Economics, 139*, 1-16.

Weinstein, A. (2014). Local labor market restructuring in the shale boom. *Journal of Regional Analysis and Policy, 44*(1). 71–92.

Weng, L., Boedhihartono, A. K., Dirks, P. H., Dixon, J., Lubis, M. I., & Sayer, J. A. (2013). Mineral industries, growth corridors and agricultural development in Africa. *Global Food Security, 2*(3), 195-202.

Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach.* 5th edition. Mason, OH: South-Western.

World Bank. (2002). Chapter 2: What is empowerment? In Narayan, D. (ed). *Empowerment and poverty reduction: A sourcebook*, pp. 10–23. PREM (Word Bank), Washington, DC.

World Bank. (2013). *Gender in extractive industries.* World Bank Group. Retrieved from http://www.worldbank.org/en/topic/extractiveindustries/

World Bank. (2015). *Gender equality, poverty reduction and inclusive growth: Gender strategy 2016 - 2023.* World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/820851467992505410/pdf/102114-REVISED-PUBLIC-WBG-Gender-Strategy.pdf

World Bank Data. (2019). Labor force participation rate, female. World Bank Group. Retrieved from https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS

A Appendix

Table 14: Agricultural Sector (excluding self-employment)

	(1) Agriculture	(2) Agriculture	(3) Agriculture	(4) Agriculture
Treated	-0.0238***	-0.0205**	-0.0205**	-0.0211**
	(0.00768)	(0.0101)	(0.0101)	(0.0103)
Active × Treated	0.0102	0.00663	0.00662	
	(0.00690)	(0.00942)	(0.00941)	
Construction × Treated		-0.00587		
		(0.00885)		
Early Construction × Treated			-0.0115	-0.0119
			(0.0105)	(0.0105)
Late Construction × Treated			-0.00116	-0.00170
			(0.00998)	(0.00998)
Active × Treated (0-3y)				0.0182
				(0.0120)
Active × Treated (4-8y)				0.00243
				(0.0109)
Active × Treated (post 8y)				-0.00216
				(0.0112)
N	97660	97660	97660	97660
R^2	0.153	0.153	0.153	0.153
mean	0.038	0.038	0.038	0.038
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 15: Frequency of Media Access

	(1)	(2)	(3)	(4)
	Radio	Radio	TV	TV
Treated	-0.0492	-0.0513	-0.103	-0.103
	(0.0711)	(0.0717)	(0.116)	(0.117)
Active × Treated	0.0374		0.0723	
	(0.0718)		(0.107)	
Early Construction × Treated	0.0988	0.100	0.102	0.107
	(0.0851)	(0.0851)	(0.0974)	(0.0970)
Late Construction × Treated	0.0316	0.0330	0.0773	0.0803
	(0.0899)	(0.0899)	(0.0974)	(0.0974)
Active × Treated (0-3y)		0.00963		0.0106
		(0.0824)		(0.117)
Active × Treated (4-8y)		0.0401		0.0964
		(0.0798)		(0.107)
Active × Treated (post 8y)		0.0790		0.137
		(0.0881)		(0.117)
N	79784	79784	79730	79730
R^2	0.343	0.343	0.591	0.591
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 16: Agricultural Sector Employment (25 km radius)

	(1)	(2)	(3)	(4)
	Agriculture	Agriculture	Agriculture	Agriculture
Treated (25 km)	-0.00186	-0.00660	-0.00640	-0.00536
	(0.0129)	(0.0138)	(0.0137)	(0.0137)
Active × Treated	0.0158*	0.0212*	0.0218*	
	(0.00914)	(0.0116)	(0.0116)	
Construction × Treated		0.00819		
		(0.0113)		
Early Construction × Treated			0.0256	0.0263
			(0.0173)	(0.0173)
Late Construction × Treated			-0.00210	-0.00133
			(0.0118)	(0.0118)
Active × Treated (0-3y)				0.0144
				(0.0148)
Active × Treated (4-8y)				0.0308**
				(0.0121)
Active × Treated (post 8y)				0.0164
				(0.0149)
N	95422	95422	95422	95422
R^2	0.196	0.196	0.196	0.196
mean	0.074	0.074	0.074	0.074
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 17: Manual Sector Employment (25 km radius)

	(1) Manual	(2) Manual	(3) Manual	(4) Manual
Treated (25 km)	-0.00549*	-0.000830	-0.000917	-0.00152
	(0.00297)	(0.00325)	(0.00326)	(0.00325)
Active × Treated	0.00378	-0.00154	-0.00181	
	(0.00313)	(0.00361)	(0.00360)	
Construction × Treated		-0.00805**		
		(0.00355)		
Early Construction × Treated			-0.0154***	-0.0149***
			(0.00394)	(0.00392)
Late Construction × Treated			-0.00368	-0.00338
			(0.00445)	(0.00446)
Active × Treated (0-3y)				-0.00154
				(0.00365)
Active × Treated (4-8y)				-0.00439
				(0.00434)
Active × Treated (post 8y)				0.00534
				(0.00474)
N	95422	95422	95422	95422
R^2	0.024	0.024	0.024	0.024
mean	0.021	0.021	0.021	0.021
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 18: Service Sector Employment (50 km radius)

	(1)	(2)	(3)	(4)
	Services	Services	Services	Services
Treated (50 km)	0.000422	0.00211	0.00213	-0.000316
	(0.00947)	(0.00976)	(0.00977)	(0.00970)
Active × Treated	-0.000808	-0.00290	-0.00272	
	(0.00613)	(0.00766)	(0.00766)	
Construction × Treated		-0.00327		
		(0.00653)		
Early Construction × Treated			0.00526	0.00693
			(0.00740)	(0.00757)
Late Construction × Treated			-0.0156*	-0.0143*
			(0.00839)	(0.00850)
Active × Treated (0-3y)				0.00249
				(0.00886)
Active × Treated (4-8y)				-0.00872
				(0.00881)
Active × Treated (post 8y)				0.0174
				(0.0130)
N	81983	81983	81983	81983
R^2	0.253	0.253	0.253	0.253
mean	0.202	0.202	0.202	0.202
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 19: Total Employment (never-movers)

	(1) Working	(2) Working	(3) Working	(4) Working
Treated (15 km)	-0.0157	-0.0213	-0.0216	-0.0247
	(0.0163)	(0.0190)	(0.0190)	(0.0190)
Active × Treated	0.0194	0.0257	0.0260	
	(0.0203)	(0.0233)	(0.0234)	
Construction × Treated		0.0100		
		(0.0209)		
Early Construction × Treated			0.00588	0.00296
			(0.0262)	(0.0263)
Late Construction × Treated			0.0144	0.0118
			(0.0244)	(0.0243)
Active × Treated (0-3y)				0.0502*
				(0.0287)
Active × Treated (4-8y)				0.0144
				(0.0313)
Active × Treated (post 8y)				-0.0354
				(0.0314)
N	43007	43007	43007	43007
R^2	0.182	0.182	0.182	0.182
mean	0.259	0.259	0.259	0.259
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 20: Agricultural Sector Employment (never-movers)

	(1) Agriculture	(2) Agriculture	(3) Agriculture	(4) Agriculture
Treated (15 km)	-0.0165	-0.00264	-0.00236	-0.00417
	(0.0136)	(0.0163)	(0.0164)	(0.0165)
Active × Treated	0.00563	-0.00989	-0.0102	
	(0.0153)	(0.0180)	(0.0181)	
Construction × Treated		-0.0248		
		(0.0163)		
Early Construction × Treated			-0.0203	-0.0221
			(0.0219)	(0.0219)
Late Construction × Treated			-0.0295*	-0.0316*
			(0.0162)	(0.0162)
Active × Treated (0-3y)				0.00478
				(0.0212)
Active × Treated (4-8y)				-0.00235
				(0.0284)
Active × Treated (post 8y)				-0.0749***
				(0.0226)
N	43007	43007	43007	43007
R^2	0.214	0.214	0.214	0.214
mean	0.089	0.089	0.089	0.089
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Table 21: Manual Sector Employment (never-movers)

	(1) Manual	(2) Manual	(3) Manual	(4) Manual
Treated (15 km)	0.00404	0.00632	0.00640	0.00627
	(0.00473)	(0.00623)	(0.00621)	(0.00618)
Active × Treated	-0.00264	-0.00519	-0.00530	
	(0.00516)	(0.00751)	(0.00749)	
Construction × Treated		-0.00408		
		(0.00796)		
Early Construction × Treated			-0.00274	-0.00276
			(0.00972)	(0.00979)
Late Construction × Treated			-0.00549	-0.00530
			(0.00830)	(0.00836)
Active × Treated (0-3y)				-0.00489
				(0.00809)
Active × Treated (4-8y)				-0.0141
				(0.00901)
Active × Treated (post 8y)				0.00881
				(0.0124)
N	43007	43007	43007	43007
R^2	0.022	0.022	0.022	0.023
mean	0.018	0.018	0.018	0.018
Country-Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Standard errors in parentheses

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