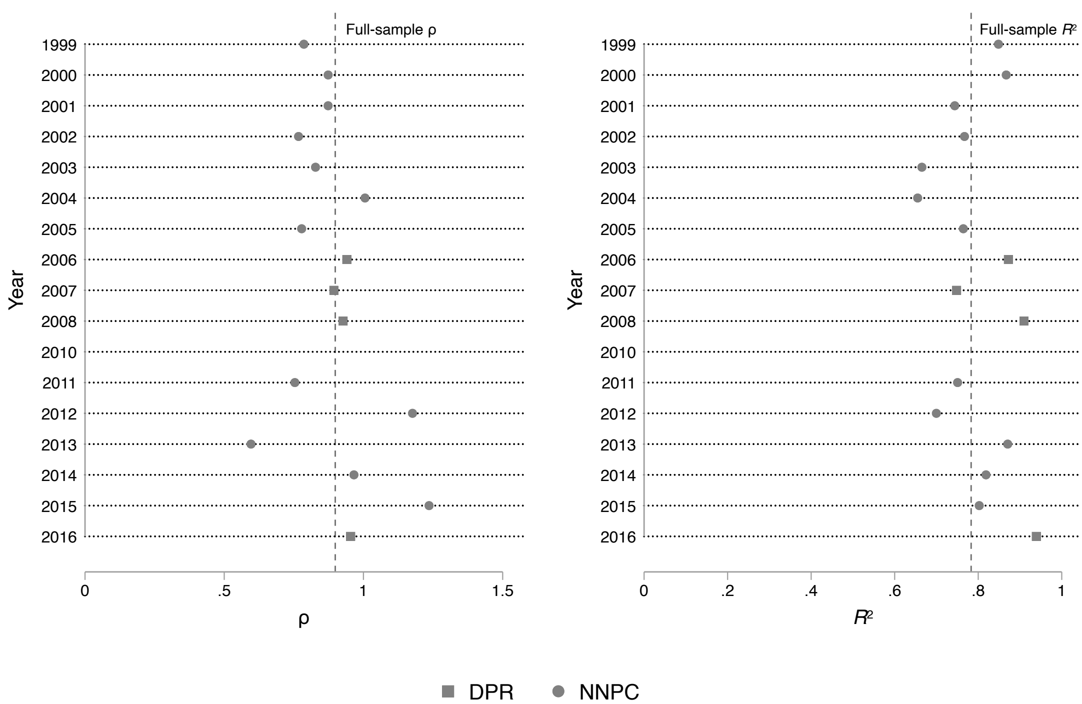
ONLINE APPENDIX
— For Online Publication Only —

DATA APPENDIX

A1. Oil production and infrastructure data

Information on 314 active Nigerian oilfields forms the core of the data. These field-level data come from Annual Statistical Bulletin of the NNPC, augmented with confidential data from the Department of Petroleum Resources (DPR)[25] for years in which NNPC data is unavailable. Between these two sources, I observe oil production for each oilfield from 1998-2016.[26] Because of uneven coverage, some fields are missing in certain years after the field first appears in the data. I assign output in these field-years to missing, while coding output as zero only when it is explicitly indicated as such in a DPR or NNPC source. A "shut-in" field is defined as a field that is nonproducing in a given time period.

Figure A1. Year-to-year correlations in oil output



Note: Figure shows coefficient estimates (left panel) and R^2 (right panel) from separate AR(1) regressions of oil output for each consecutive year pair in the data, indicated on the vertical axis by the second year of the pair. Vertical dashed lines indicates the coefficient or R^2 from an AR(1) regression on the pooled full sample. Marker symbols indicate data source by year. Sample is an unbalanced panel of 314 oilfields from 1998-2016. Oil production data are missing for 2009, so estimates for 2008-2009 and 2009-2010 are excluded.

[25]The DPR is Nigeria's primary petroleum sector regulatory body.
[26]Unfortunately, disaggregated data are unavailable for 2009.

There are significant reporting format and content differences between the DPR and NNPC data. DPR data covers a larger number of fields and companies, while NNPC reports are provisional and may aggregate across neighboring fields for smaller operators, or even exclude them entirely. Unfortunately, DPR data are only available for four years of the sample: 2006-2008 and 2016, none of which overlap with years in which NNPC data is available. To validate the comparability of the two series', I estimate AR(1) regressions for each pair of consecutive years in the sample. The resulting R^2 and autocorrelation coefficient ρ for these regressions are plotted in Figure A1. Year-to-year correlation is generally high and similar across both data sources, and remains high in year-pairs when the data source changes.

DRP also provides time-invariant field covariates: the number of wells (field size), date of completion of the first well (field age), and the depth of the deepest well. I use infrastructure maps to obtain centroid locations for the fields in the DPR-NNPC data, which are then used to link fields to information on oil theft, militancy, and various control variables. The fields are mapped in Figure 2, with the color indicating the year in which the field was indigenized.

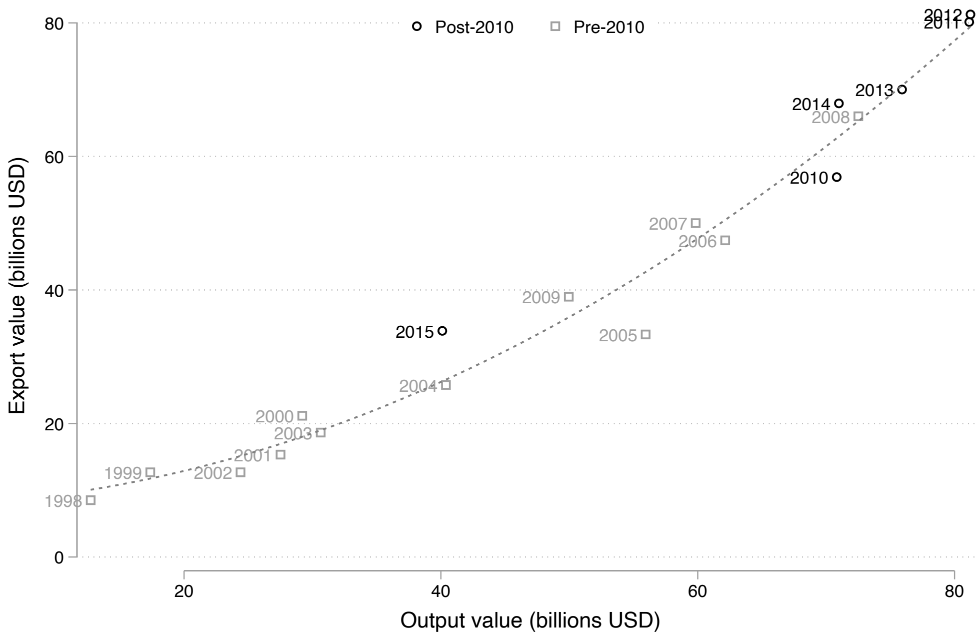
Measurement error in output may be correlated with localization. For example, firms may underreport output in order to evade royalty taxation. If multinational firms tend to have a greater incentive to misreport, then any local output advantage that materializes after divestment may simply be driven by reporting biases. I corroborate the reliability of Nigerian administrative data on output using independent data on export values from UN COMTRADE. This is a reasonable comparison, since the vast majority of Nigerian oil (85-90%) is exported. I aggregate all administrative NNPC data from 1998-2015 to the year level and value this production at annual world crude oil prices. I then correlate this series against reported COMTRADE crude oil export values in Figure A2. The two series are highly correlated over time. Furthermore, if selective reporting drives the results, the correlation between the two series should strengthen over time (in particular, after 2010) as the local market share grows and under-reporting falls. There is no evidence that post-2010 observations are systematically more correlated; observations in both periods are tightly clustered around the regression line.

A2. *Corporate transactions data*

Data on corporate transactions comes from DrillingInfo (DI), a paid-subscription database on the oil and gas sector. From DI I obtain a list of 155 corporate transactions in the Nigerian oil and gas sector from 2006-2020. I digitize PDF files for each transaction which contain the announcement and closing dates, name of buyers, sellers, and assets, deal value, deal status at the time of reporting (closed, terminated, or in progress), and the type of transaction. I then drop new exploration awards, which cover license awards from the Nigerian government to private firms, since these licenses contain unexplored fields that do not enter into the data. I also drop corporate M&A transactions, which typically do not refer to specific assets but rather reflect changes in the ownership structure of entire firms. 117 transactions remain after these sample restrictions. For each asset transaction, I retain the nationality of buyers, sellers, the transaction opening and closing dates, and whether it was successful.

Many transactions contain information on both fields and block, since the former is typically, though not always, contained in the latter. If field-level information is available, I use that, since some fields within a block may be divested while others are not; otherwise I take the block-level information. Of the 117 DI transactions, 74 contain specific fields, covering 104 unique fields. The remaining 43 transactions mention only blocks, covering 44 blocks. In total, 43 out of 104 field-level transactions and 15 out of 44 blocks are matched to the field panel data. Since only 27% of the 117 transactions cover assets that are actively producing at

Figure A2. Output measurement validation: export value



Note: Figure shows time-series correlation between aggregate oil production value as measured by Nigerian administrative sources and Nigerian oil export values from UN Comtrade data, both measured in billions of USD. Pre- and post-2010 observations are indicated in figure.

the time of the transaction, these match rates are not unreasonable.

A3. Oil block concessions data

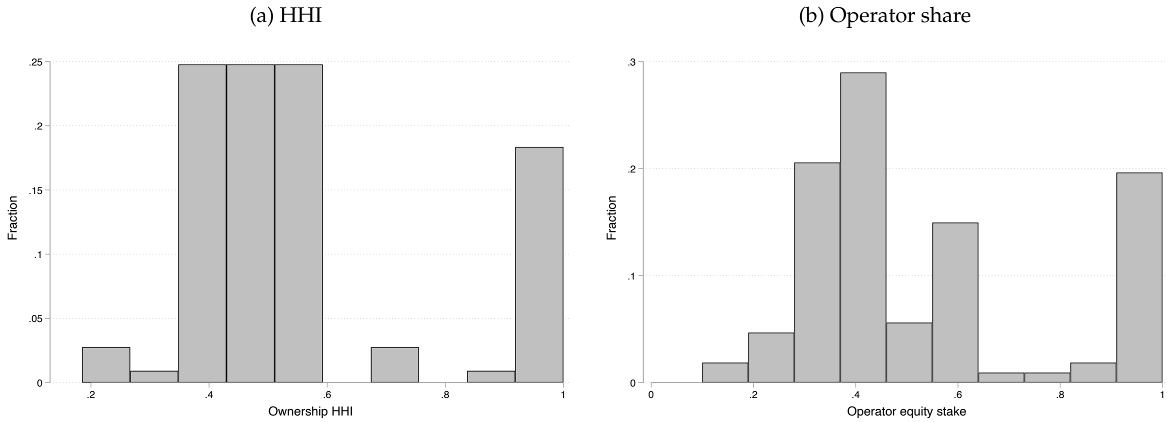
Concessions – large blocks of territory, typically containing several oilfields – are the primary unit of ownership in the Nigerian oil market. The exceptions to this rule are "marginal" fields, independently-owned fields awarded to local operators that do not belong to larger concession blocks. Concessions are typically jointly owned by several partners, often including an equity stake for the Nigerian National Petroleum Corporation (NNPC).

Data on 73 concessions and 11 marginal fields for the years 2006-2016 comes from the DPR and the Nigerian Extractive Industries Transparency Initiative (NEITI). These data contain the concession size, location, operator, license type, equity breakdown, and depth (for offshore properties). These are the relevant parameters for calculating royalty rates. These concessions match 304 of the 314 fields in the main field-level data. Figure A3 shows the block-level distribution of ownership shares in concessions as of 2016. Panel A shows the Herfindahl index of ownership, while Panel B gives the share owned by the primary operating firm. There is substantial variation in ownership structure across assets in the Niger Delta.

A4. Treatment definition

Both main sources of data contain information on the identity of firms in control of oil fields. Still, both datasets contain substantial gaps and drawbacks; therefore, in my preferred speci-

Figure A3. Distribution of ownership shares, 2016



Note: Figure shows histograms of ownership concentration, measured as the Herfindahl index (Panel A), and the stake owned by the operating company (Panel B). Sample is a cross-section of 106 active oil blocks (licenses) in 2016.

fications, I combine information from both sources to measure the localization "treatment" at the field-year level. The DPR-NNPC dataset includes information on the firm operating each field in each year. However, as mentioned, this data has substantial gaps – many fields are missing information for years after they first enter the data. In addition, operatorship information is likely to lag divestments given administrative data collection challenges. Therefore, it is difficult from this data alone to determine the exact year in which a given treatment occurs. Furthermore, using the operatorship measure exclusively overlooks cases in which local firms are non-operating shareholders, which may also be important.

In contrast, the DI data provides detailed information on a substantially wider set of transactions, covering all cases in which any ownership stake in a given field is transferred from a multinational to a local firm. It also contains precise information on the date of a transaction. However, it does not have information before 2006, so we cannot identify which fields are always-treated (that is, divested as of 2006) in our difference-in-differences setup using only the DI data. Furthermore, while it provides a wealth of transaction-level detail, it does not include comprehensive information on operatorship per se, simply changes in ownership.[27]

However, when combined, the longer panel of the NNPC data can fill in the always-treated firms, while the DI data can provide more precise timing and information on non-operating divestments from 2006-2016. Using both data sources provides the most detailed picture of local participation, defined as any operatorship and/or ownership by an indigenous Nigerian firm. An indigenous Nigerian firm is any firm headquartered in Nigeria and not majority-owned by a firm headquartered outside of Nigeria.[28]

A detailed breakdown of the treatment definition is provided in Table A1. Firstly, I define a MNC-to-local "divestment" indicator from the DI data d_i^{DI}, which equals one for all fields that observed a transaction from 2006-2016 in which any buyer was Nigerian and any seller

[27] In the deal description, there is mention of operatorship in only 68 of the 117 transactions.

[28] In practice, this means that the local subsidiaries of oil supermajors in Nigeria – Eni, Total, Shell, Chevron, and ExxonMobil, as well as the Chinese Sinopec subsidiary Addax Petroleum – are classified as MNC; all others are indigenous. Note that the national oil company, the Nigerian Petroleum Development Company (NPDC) is considered an indigenous firm.

was multinational. Next, I create a dummy d_i^{NNPC} for all fields that are operated by local firm according to NNPC from 1998-2016. The first row of Table A1 corresponds to the 32 fields that experience a divestment from 2006-2016 according to both NNPC and DI records, so that $d_i^{DI} = 1$ and $d_i^{NNPC} = 1$. If DI and NNPC disagree on timing, these fields take as their treatment year the *earliest* year that local participation is reported by either source. These observations likely correspond to locally owned and operated fields. The next row indicates a further 24 fields that have any local participation from 2006-2016 according to DI data, but which are listed as solely multinationally operated over this period by NNPC, so $d_i^{DI} = 1$ and $d_i^{NNPC} = 0$. These are likely local firms that take stakes in multinationally operated fields without assuming operatorship. They take as their treatment year the closing date of the DI transaction, where available; otherwise the opening date.

Table A1—Field counts by treatment type

	Number of fields	Share
Both	32	34.04
DrillingInfo only	24	25.53
NNPC only (always treated)	33	35.11
NNPC only (transition)	5	5.32
Total	94	.

Table displays the number of fields that are marked as treatment by different data sources. Both refers to fields that are divested in both DI and NNPC data. DI only contains fields only divested in the DI data. NNPC only (always treated) contains fields that are treated in NNPC data before the DI data begins in 2006. NNPC only (transition) are fields that are divested between 2006-2016 in the NNPC but not DI data.

The third and fourth rows, together, contain fields where $d_i^{DI} = 0$ and $d_i^{NNPC} = 1$, so the treatment year is taken from the NNPC data as the first year that the field has a local operator. These fall into two categories. First are the 33 always-treated fields, or those that have local participation from 1998-2005 according to NNPC, but where $d_i^{DI} = 0$ by construction. These are the fields that we know from NNPC must be local prior to the start of the DI data in 2006. Lastly, the fourth row contains the final 5 treated fields, which become treated between 2006-2016 according to NNPC but not DI. Reassuringly, this is a small number, which we should expect since DI is a more expansive dataset capturing both owner and operator transactions. In total, there are 94 ever-treated fields. The final treatment indicator $local_{it}$ is equal to one for all field-years after the treatment year, following the "staggered adoption" difference-in-differences setup. Reverse divestments from locals to multinationals are exceedingly rate in the DI data, affecting less than 1% of field-years from 2006-2016.

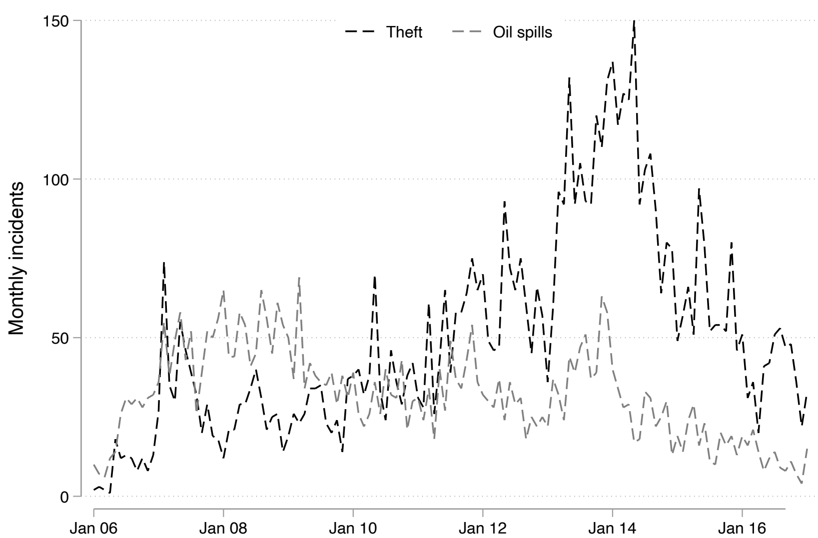
For the placebo tests, I define additional variables from the DI data, indicating a field's exposure to local-to-local (27) and all other transactions (6). Finally, I construct an indicator of delayed divestments which equals one for all years after the announcement of an MNC-to-local divestment but before it's consummation. In some cases, these are terminated/nullified transactions, while in others, this reflects a delay between the opening and closing dates. This indicator equals zero if and when the field eventually becomes "treated" according to the DI data. 36 fields are exposed to a delayed or terminated divestment in the sample period.

A5. Oil spills and theft data

Data on oil theft comes from the Nigerian Oil Spill Detection and Response Agency (NOSDRA), a division of the Federal Ministry of the Environment. NOSDRA data is taken from the Oil Spill Monitor (OSM), a comprehensive database of all 11,587 reported oil spills from 2006-2017. For each oil spill, NOSDRA investigates and files a Joint Investigative Report (JIV), verified by local communities, the oil company, and the DPR. For each spill, I observe the location and cause of the spill, as well as a text description. For those without coordinates, I georeference based on site description in the JIV, resulting in 11,145 spills with coordinates.

68.45 % of all oil spills are classified as being caused by "sabotage." I take this to be my sample of oil theft incidents, since sabotage is a reliable indicator of illegal oil tapping.[29] For each field, I define theft as the sum of all sabotage incidents that occur annually within 15 km of the centroid of the field. To measure the technical efficiency of oil production, I use all field-level spills that are not due to sabotage. In the OSM, the majority (65.3%) of these non-sabotage incidents are caused by "equipment failure" and "corrosion." They are thus a reasonable measure for losses incurred by oil companies during the normal course of business that can be controlled by the firm directly. Figure A4 charts the evolution of the black market by plotting the monthly incidents of pipeline sabotage and operational oil spills from 2006 through 2016. Oil spills due to theft rise dramatically from 2010-2014, and then fall thereafter. Oil spills due to operational failure, in contrast, decline over the whole period.

Figure A4. Sabotage and operational oil spills over time



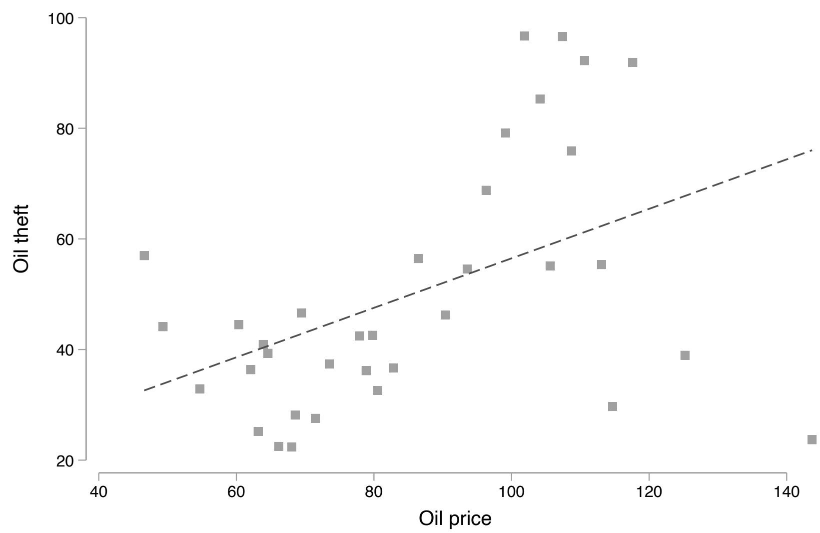
Note: Figure shows monthly totals of oil spills due to sabotage and non-sabotage (equipment failure) over time. Data come from 11,587 oil spills recorded by the NOSDRA OSM from 2006-2016. Vertical lines indicate the beginning of the federal amnesty program for ex-combatants, the end of the initial amnesty period, as well as the proposed rollback of amnesty benefits. Reprinted from Rexer and Hvinden (2022).

The most likely form of measurement error bias in the oil theft variable is that firms have incentives to mis-classify spills due to operational failure as sabotage in order to avoid legal li-

[29] Rexer and Hvinden (2022) for a discussion about measuring oil theft.

ability for compensation to host communities. In 2018, Amnesty International embarked on a large-scale effort to verify the causes of oil spills on Shell and Eni properties in the Niger Delta using JIV reports from 2011-2017. Out of 1830 oil spills analyzed, Amnesty found evidence of misclassification in only 89, less than 5% (Amnesty International, 2018). Misclassification is rare, and unlikely to drive the results. To provide further evidence that sabotage meaningfully captures organized criminal profit-seeking behavior – rather than, e.g., pipeline vandalism as community protest against oil companies – Figure A5 plots monthly sabotage events against the global oil price, controlling for the total level of violence, oil production, and seasonal effects. The positive slope demonstrates that sabotage responds strongly to price incentives, suggesting it is primarily driven by the economic logic of oil theft.

Figure A5. Oil theft and oil prices, monthly time series



Note: Figure shows the monthly time series correlation between oil theft and global oil prices. Scatterplot controls for month-of-year effects (seasonality), total conflict deaths, and total oil production. Scatterplot is binned at 35 quantiles of the price distribution.

A6. Conflict data

To measure violent conflict, I use data from the Armed Conflict Location and Event Dataset (ACLED) from 1998-2016. To measure oil-related violence, I use all conflict events that contain the following oil-industry-related strings: petroleum, petro, Agip, Shell, Eni, drilling, rig, well, pipeline, ndv, flow, NNPC, NPDC, exxon, mobil, total, addax, or gas. This captures attacks on the oil sector perpetrated by any armed groups. I then further distinguish between conflict events perpetrated by organized rebel or political militia groups, which I call "militant" attacks, and those perpetrated by unknown or unorganized groups, which I call "non-militant" attacks. For each field, I aggregate the sum of annual attacks and fatalities of different types within 15 kilometers of the field centroid. ACLED event data are derived from news media reports (see for a discussion of the methodology), and report the journalistic source for each conflict event. In some cases, I subset to only conflict events reported by local news media sources in order to test for different sources of measurement error; see Appendix E.E1 for a discussion of measurement error correlated with the divestment treatment.

A7. Gas flaring data

In addition to oil spills, gas flaring represents a major source of environmental pollution from oil production in the Niger Delta. Flaring occurs when natural gas created as a byproduct from oil production is not economically viable to capture and transport to market, and is therefore burned on site. Gas flaring pollutes air quality, vegetation, and waterways, worsens health outcomes,[30] and contributes to climate change with CO_2 emissions.

Data on gas flaring volumes comes from the Nigeria Gas Flare Tracker,[31] a joint project by NOSDRA and the NGO Stakeholder Democracy Network. I download monthly panel data on total gas flaring volume from March 2012 to May 2020, measured in thousands of cubic feet (mscf), for 210 flare sites. These location-specific volume estimates can then be converted to CO_2 emissions, since according to U.S. Energy Information administration, flared natural gas emits 54.75 kg of CO_2 per mscf.[32] I then georeference these sites manually by cross-referencing the map interface of the Gas Flare Tracker against a Google maps layer containing Nigeria's oil and gas infrastructure. I then match flares to fields using a spatial merge process. 119 flare sites fall directly within the boundaries of an identifiable field. A further 73 are matched to their nearest field within 10 kilometers. The remaining 18 flare sites either fall on the Cameroonian side of the maritime border ($n = 9$), are far from the Niger Delta ($n = 2$), or are not near any identifiable field ($n = 7$). In total, these 192 final flare sites cover 143 fields. Lastly, I merge to the production data; 180 out of 192 flare sites occur in fields actually actually contained in the DPR/NNPC output data. These matched fields account for 93.4% of the flared gas volume over the period.

A8. Law enforcement data

Data on law enforcement activity comes from the text of Nigerian news media reports. We begin by assembling a comprehensive collection of plausibly relevant news articles covering topics of oil theft, law enforcement, and crime in Nigeria by searching relevant keywords in the Dow Jones Factiva media database. We collect all articles that satisfy each of the following criteria: i) mention the word "oil", ii) mention at least one of a set of enforcement-related keywords[33] , iii) mention at least one of a set of exact oil crime-related phrases.[34] Some examples of relevant articles are shown in Figure A6. Although both local and international news media sources are included in the database, in practice the majority of oil theft-related articles come from local sources.

[30]See Ologunorisa (2009) for a review of studies on the negative impacts of Niger Delta flaring.

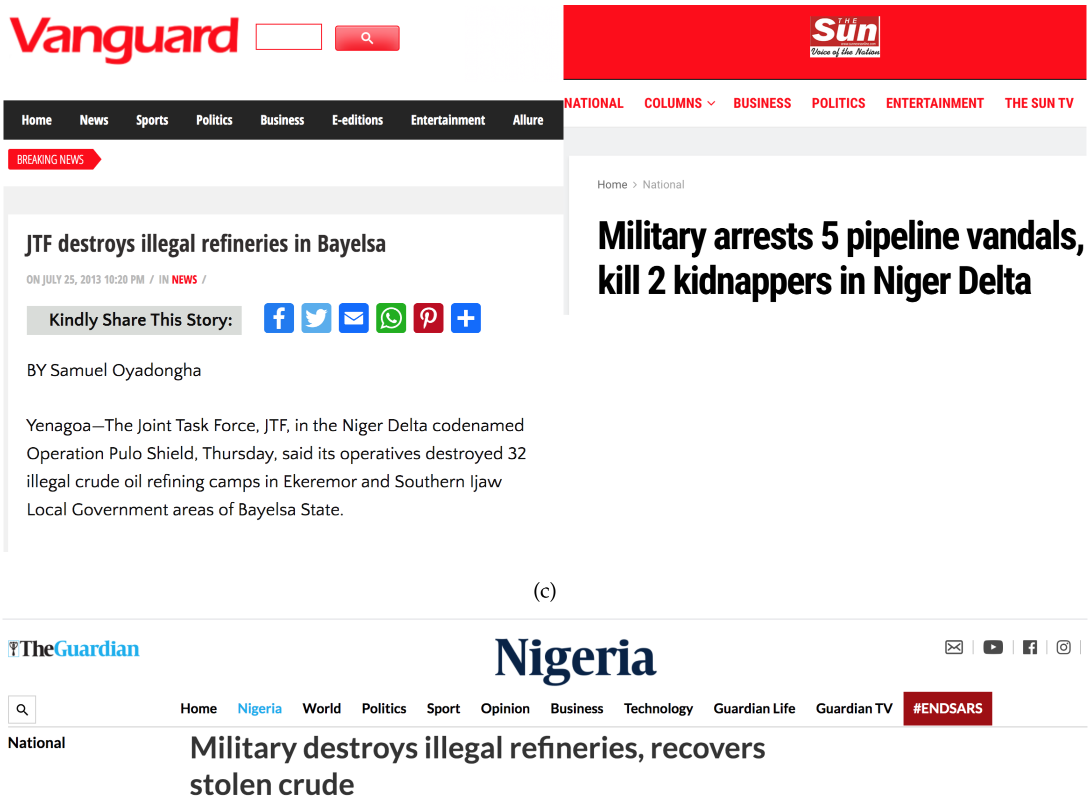
[31]`https://nosdra.gasflaretracker.ng/`

[32]`https://www.eia.gov/environment/emissions/co2_vol_mass.php`

[33]These are: "raid, raids, raided, seize, seized, seizes, seizure, seizures, destroy, destroys, destroyed, operation, capture, captures, captured, arrest, arrested, arrests, kill, killed, kills, apprehend, apprehended, apprehends, burn, burns, burned, invade, invaded, invades, search, searches, search"

[34]These are: artisanal refineries, artisanal refinery, artisanal refining, bunkerers, bunkering camp, bunkering gang, bunkering site, ex toru, illegal bunkering, illegal diesel, illegal fuel, illegal oil, illegal refineries, illegal refinery, illegal refining, illegally refined, joint task force, Nigerian military, Nigerian Navy, oil bunkerers, oil bunkering, oil smugglers, oil theft, oil thief, oil thieves, oil vandals, operation 777, operation awase, operation crocodile smile, operation delta safe, operation eagle eye, operation pulo shield, operation python dance, operation restore hope, operation river sweep, operation safety check, operation tsare teku, pipeline sabotage, pipeline vandal, pipeline vandalism, pipeline vandals, pirate, pirates, stolen crude, stolen, diesel, stolen oil, swamp buggy.

Figure A6. Sample enforcement articles from local news media sources



(a)

Vanguard

Home News Sports Politics Business E-editions Entertainment Allure

BREAKING NEWS

JTF destroys illegal refineries in Bayelsa

ON JULY 25, 2013 10:20 PM / IN NEWS /

Kindly Share This Story:

BY Samuel Oyadongha

Yenagoa—The Joint Task Force, JTF, in the Niger Delta codenamed Operation Pulo Shield, Thursday, said its operatives destroyed 32 illegal crude oil refining camps in Ekeremor and Southern Ijaw Local Government areas of Bayelsa State.

(b)

The Sun Voice of the Nation

NATIONAL COLUMNS BUSINESS POLITICS ENTERTAINMENT THE SUN TV

Home > National

Military arrests 5 pipeline vandals, kill 2 kidnappers in Niger Delta

(c)

TheGuardian Nigeria

Home Nigeria World Politics Sport Opinion Business Technology Guardian Life Guardian TV #ENDSARS

National

Military destroys illegal refineries, recovers stolen crude

Note: This figure shows screenshots from relevant articles in *The Vanguard*, *The Sun*, and *The Guardian Nigeria*, all local Nigerian newspapers.

This procedure yields 17146 total articles potentially related to oil theft enforcement.[35] We then hired Nigerian research assistants to first identify all articles that are relevant to law enforcement activity in Nigeria, yielding a total of 3932.[36] From this set of relevant articles we then manually extract all *law enforcement events*, where an event is defined as a unique interaction between law enforcement and suspected criminals that occurs in a specific location. For each event, we code the following variables: *i*) the location of the event, typically a neighborhood, village, oil asset, or local government area (municipality) *ii*) the law enforcement agency, *iii*) the illegal activity committed, selected from a pre-coded list,[37] *iv*) the items seized or destroyed in the law enforcement action, selected from a pre-coded list,[38] *v*) the to-

[35]Of course, these search terms are unlikely to be exhaustive, but they were derived from substantial reading of these articles. Also, note that this figure may be inflated because the same story is sometimes published by multiple different media outlets.

[36]We excluded articles about unrelated conflicts such as Boko Haram in Northern Nigeria, but included articles about non-oil illegal activities such as armed robbery, gang activity, and fraud

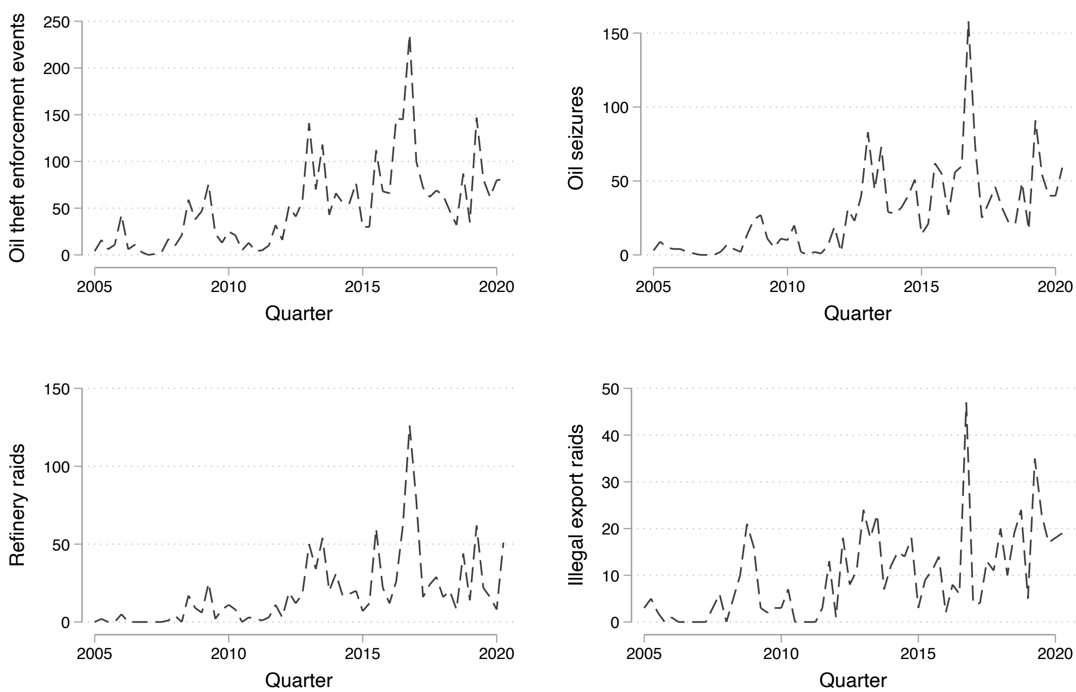
[37]These are: oil theft, piracy, illegal refining, pipeline vandalism, transportation of stolen oil, kidnapping, cultism/gang activity, militancy, and other illegal activities.

[38]These are: no item, boats, stolen oil, arms, illegal refineries, trucks, oil theft equipment, and other items.

tal number of arrests, and *vi*) the total number of fatalities. Extensive manual quality checks were conducted on weekly researcher submissions.

We consider any two articles as duplicates if they are published in the same calendar week and their headlines exceed a string similarity score threshold as defined by Levenshtein edit distance. After grouping duplicates into unique articles, we take the union of all events identified by the researchers to allow that duplicate articles may contain both repeated events as well as independent information.[39] In total, we obtain 5682 law enforcement events for which the location can be reliably geocoded, of which 3261 are related to oil theft. These events cover 3379 unique articles. 89% of all locations mentioned in relevant events were successfully geocoded. We then merge these enforcement events to villages in our sample using 5 kilometer rings, the same criteria used for oil theft. Figure A7 plots quarterly total law enforcement actions for several different criminal activity categories.

Figure A7. Anti-oil theft enforcement



Note: This figure shows total quarterly oil theft-related law enforcement actions for the following categories of enforcement: all oil theft (top left), seizures of stolen petroleum products (top right), raids on illicit refineries (bottom left), and raids on vessels engaged in illegal export of stolen oil (bottom right).

One concern about the enforcement data is selective coverage of law enforcement activity. If, for example, the local Nigerian media reports relatively more on theft when the firm is local, then an increase in enforcement following divestment may be spurious and driven en-

[39]For example, if there are two articles about the same raid, one may mention a second event, while the other does not.

tirely by reporting bias. One way to assess this investigate the relationship between oil theft events and oil theft enforcement events. First, if news articles meaningfully capture variation in underlying enforcement, then news coverage should be robustly correlated with theft over space and time. Second, if there is no differential reporting bias, this relationship should be relatively constant across assets, irrespective of their ownership status. If instead news coverage is more (or less) responsive to the underlying level of theft depending on firm ownership, this is suggestive of reporting bias. I investigate this hypothesis in Table A2. Columns (1)-(4) show a robust positive correlation between oil theft and news about oil theft enforcement, which is stable and robust to controls, year effects, and field fixed effects. However, columns (5)-(6) show that this relationship does not differ between local and multinational firms. This suggests differential reporting bias is unlikely to be a major concern.

Table A2—Enforcement and oil theft

Outcome	Oil theft enforcement					
	(1)	(2)	(3)	(4)	(5)	(6)
Oil theft events, 15 km	0.024***	0.023***	0.028***	0.026***	0.030***	0.028***
	(0.008)	(0.009)	(0.008)	(0.007)	(0.008)	(0.007)
Local firm					1.770**	1.250
					(0.873)	(0.773)
Oil theft events, 15 km × Local firm					0.001	0.008
					(0.070)	(0.059)
Field FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	No	No	Yes	No	Yes
Observations	3183	3183	3183	3183	3183	3183
R^2	0.009	0.114	0.487	0.590	0.491	0.592

Standard errors, in brackets, are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Outcome variable is oil theft enforcement, the total number of enforcement news events within 15 km of the field. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A9. *Political connections data*

To collect data on the political connections of firms in the Nigerian oil sector, I partner with a market research firm called Asoko Insight,[40], headquartered in London and Nairobi, and specializing in research on corporate entities in sub-Saharan Africa. We collected biographical information on the universe of corporate officers and shareholders in the sector. Below, I provide detailed steps that Asoko Insight uses to collect and assemble the biographical data that forms the basis of the firm-level political connections dataset.

1) Document retrieval: Asoko requests corporate filings from the Nigerian Corporate Affairs Commission (CAC), from the date of incorporation to the present. For each firm, there is at minimum an initial filing at incorporation, detailing all of the shareholders

[40] https://www.asokoinsight.com/

and directors of the firm. In nearly all cases, firms will file subsequent additional reports of current shareholders and directors when there is a change in the company ownership or personnel.

2) Search report: The documents for a given firm are then compiled by Asoko Insight into a "search report," containing basic firm-level data as well as shareholder and director history. Basic data include full name, date of incorporation in Nigeria, registered address, and issued share capital. Individual-level data includes name, address, status, and share allotment (if shareholder).

3) Transcription: Asoko then transcribes the search reports into two Excel files: one containing company-level data, and another containing person-company-filing-level data on directors and shareholders.

4) Cross-referencing: The names in the search report are then cross-referenced by Asoko against a list provided by the researcher. This list contains biographies of oil sector personnel that were able to be identified before engaging Asoko. These are taken as given and removed from the list of personnel for whom Asoko must obtain biographies. In addition, entity shareholders (e.g., holding companies) and foreign personnel are removed from the directors/shareholders list.

5) Ultimate beneficial owners (UBOs): We then take those shareholders that are themselves corporate entities and find additional information on their officers using ng-check.com – a public Nigerian corporate registry containing less detail than the CAC filings. We add all names associated with these firms to the list of personnel in 4), as "second-level UBOs." The result of steps 4 and 5 is a final list of Nigerian individuals associated with the firm for whom Asoko must obtain biographies.

6) Biographical research: The last step in the procedure is to obtain biographies for these individuals. Asoko's Nigerian field researchers use a variety of methods, including desk research of all publicly available information in the local and international business press. In addition, researchers fill in gaps by employing key informant interviews that leverage their substantial corporate network in Nigeria. The result is an individual-level dataset containing the name and biographic details for all the unique individuals identified in steps 4 and 5.

In total, we obtain a list of 706 Nigerian nationals associated with 49 distinct corporate entities, covering all the firms listed in the DI and NNPC data. We are able to find biographic data for 552 of these individuals, implying an overall match rate of 78.2%. However, this rate varies by firm; Table A3 lists all of the firms included in the data collection procedure and their match rates. We further obtain data on 90 second-level UBOs who are not listed on corporate filings. I then code these biographies into the following indicator variables:

- Any politics: if the individual has any previous political activities in Nigeria.
- Elected politician: if the individual has ever held any elected office in Nigeria.
- Technocrat: if the individual has ever worked for the DPR, the NPDC, the NNPC, the Ministry of Petroleum Resources, or any other oil-related regulatory agency in the Nigerian Federal Government.

- High-level: if the individual has ever held a cabinet-level position in the Nigerian Federal Government.
- Security forces: if the individual has ever worked for the Nigerian Federal Police or in any military branch.
- Chief: if the individual holds any non-governmental, inherited, traditional title, e.g., the Oba of Benin or the Emir of Kano.

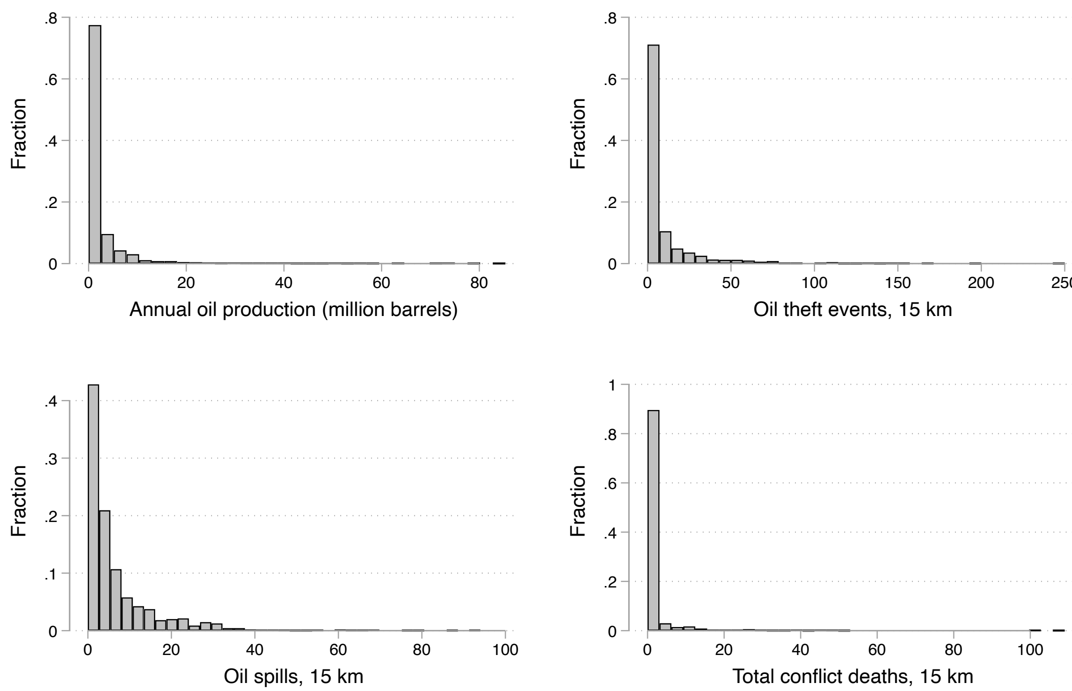
Using these dummy variables and the dates of the corporate filings as start and end dates of individual tenures, I transform the data into a firm-year panel, defining time-varying dummy variables for each connection types at the firm-level. Finally, I match this panel to the field-level data using the operators indicated in the NNPC data. Note that the political connections variables vary over-time at the field level because of both turnover in firm-level personnel and changes in field-level ownership.

A10. Sample construction

The various data sources have different time series and degrees of completeness. To harmonize the results, I take as the sample 2006-2016, for which panel data on violent conflict, piracy, oil theft, oil spills, and oil output is all available at the field-level. Within this period, oil production data is missing for some fields in each year because of incomplete coverage in the DPR-NNPC reports.Therefore, while the estimation sample for all non-production outcomes is 3,183 field-years, the sample for regressions in which production is the outcome falls to only 2476 field-years.

Figure A8 provides histograms for the main outcomes used throughout the paper – oil output, oil spills, oil theft, and conflict deaths – in our estimation sample. All of the outcome variables are non-negative and right-skewed with long tails and a mass point at zero.

Figure A8. Histogram of main outcomes



Note: Figure shows histogram of main outcomes for the primary estimation sample of field-years from 2006-2016. Oil theft events are measured as all sabotage-related oil spills within 15 km of the oilfield. Oil spills are measured as all malfunction-related oil spills within 15 km of the oilfield. Conflict deaths are all conflict deaths within 15km of the field.

Table A3—Coverage rates by firm for biographical data

Firm	Officers	Foreign	Nigerian	Coverage
AITEO EASTERN E&P CO LTD	3	0	3	1.00
ATLAS PETROLEUM INTERNATIONAL LTD	8	1	7	1.00
CAVENDISH PETROLEUM NIGERIA LTD	4	0	4	1.00
ELCREST EXPLORATION AND PRODUCTION NIGERIA LTD	13	3	10	1.00
EROTON EXPLORATION AND PRODUCTION CO LTD	14	2	12	1.00
ESSO EXPLORATION AND PRODUCTION NIGERIA LTD	42	36	6	1.00
FIRST HYDROCARBON NIGERIAN LTD	14	1	13	1.00
MIDWESTERN OIL & GAS CO LTD	15	1	14	1.00
ND WESTERN	11	3	8	1.00
NECONDE ENERGY LTD	19	5	14	1.00
NETWORK EXPLORATION AND PRODUCTION NIGERIA LTD	9	0	9	1.00
NEWCROSS PETROLEUM LTD	9	0	9	1.00
NIGERIAN AGIP OIL CO LTD	14	11	3	1.00
ORIENTAL ENERGY	10	0	10	1.00
SHORELINE NATURAL RESOURCES LTD	7	4	3	1.00
SOUTH ATLANTIC PETROLEUM LTD	17	7	10	1.00
TEXACO OVERSEAS (NIGERIA) PETROLEUM CO LTD	29	24	5	1.00
YINKA FOLAWIYO ENERGY LTD	4	0	4	1.00
MOBIL PRODUCING NIGERIA UNLTD	55	30	25	0.96
ENERGIA LTD	20	1	19	0.95
PRIME EXPLORATION AND PRODUCTION LTD	16	0	16	0.94
TOTAL NIGERIA PLC	94	72	22	0.91
PLATFORM PETROLEUM LTD	20	0	20	0.90
SHELL PETROLEUM DEVELOPMENT CO OF NIGERIA LTD	90	58	32	0.88
STAR DEEP WATER PETROLEUM LTD	35	19	16	0.88
CHEVRON NIGERIA LTD	59	38	21	0.86
NEWCROSS EXPLORATION AND PRODUCTION LTD	7	0	7	0.86
NIGER DELTA PETROLEUM RESOURCES LTD	6	0	6	0.83
PAN OCEAN OIL CORPORATION (NIGERIA) LTD	20	14	6	0.83
SHELL NIGERIA EXPLORATION AND PRODUCTION CO LTD	50	26	24	0.83
SEPLAT ENERGY PLC	25	10	15	0.80
FRONTIER OIL LTD	14	0	14	0.79
ADDAX PETROLEUM	37	28	9	0.78
AMNI INTERNATIONAL PETROLEUM CO LTD	17	5	12	0.75
TOTAL UPSTREAM NIGERIA LTD	58	50	8	0.75
EXPRESS PETROLEUM & GAS CO. LTD	7	0	7	0.71
BRITTANIA-U NIGERIA LTD	21	4	17	0.71
NIGERIAN PETROLEUM DEVELOPMENT CO LTD	53	0	53	0.68
CONTINENTAL OIL AND GAS LTD	9	0	9	0.67
UNIVERSAL ENERGY RESOURCES LTD	34	3	31	0.65
WALTER SMITH PETROLEUM OIL LTD	24	5	19	0.63
CONOIL PLC	99	48	51	0.61
PILLAR OIL LTD	24	0	24	0.58
ALLIED ENERGY PLC	30	6	24	0.54
AITEO EXPLORATION AND PRODUCTION CO LTD	6	0	6	0.50
MONI PULO LTD	20	0	20	0.50
CAMAC NIGERIA LTD	17	4	13	0.38
DUBRI OIL CO LTD	16	0	16	0.31
STERLING OIL EXPLORATION & ENERGY PRODUCTION CO LTD	5	5	0	
ALL	1230	524	706	0.78

Table shows counts of total unique identified officers (boardmembers and shareholders), foreign officers, Nigerian officers, and the coverage rate, by firm. The coverage rate is the share of Nigerian officers for whom biographical information can be found.

Table A4—Summary statistics

				Unconditional		Onshore control	
	AT	D	NT	D-AT	D-NT	D-AT	D-NT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Covariates							
Max analog well depth (m)	2739.743	2879.298	2602.357	139.555	276.942	-3.070	77.792
	(1144.686)	(906.545)	(870.252)	(221.355)	(136.793)	(221.524)	(130.255)
Field latitude	5.410	5.102	4.899	-0.308	0.203	-0.459	-0.020
	(0.794)	(0.527)	(0.599)	(0.146)	(0.081)	(0.133)	(0.083)
Distance to nearest militant camp (km)	48.032	23.825	32.642	-24.207	-8.817	-17.031	1.779
	(38.010)	(17.572)	(25.398)	(6.545)	(2.901)	(5.988)	(2.615)
Distance to state capital (km)	79.341	80.896	87.017	1.555	-6.121	7.893	3.238
	(52.915)	(53.944)	(49.719)	(11.134)	(7.926)	(10.479)	(8.223)
Distance to Atlantic coast (km)	33.072	26.884	34.076	-6.188	-7.192	-6.312	-7.376
	(31.595)	(29.911)	(29.130)	(6.455)	(4.441)	(6.410)	(4.690)
Distance to Niger river (km)	83.243	49.373	90.199	-33.870	-40.825	-11.135	-7.254
	(67.278)	(41.834)	(79.741)	(12.169)	(7.748)	(10.176)	(7.115)
Field age (2016)	34.459	46.774	42.072	12.314	4.701	10.610	2.445
	(11.987)	(10.441)	(11.998)	(2.420)	(1.655)	(2.260)	(1.608)
Number of wells	5.921	20.185	20.976	14.264	-0.791	15.774	1.318
	(7.642)	(21.659)	(34.087)	(3.182)	(3.765)	(3.341)	(3.634)
Onshore	0.737	0.929	0.645	0.192	0.283	.	.
	(0.446)	(0.260)	(0.479)	(0.080)	(0.047)	(.)	(.)
Outcomes							
Oil spills, 15 km	2.500	3.946	4.995	1.446	-1.049	1.258	-1.328
	(4.005)	(5.829)	(5.608)	(1.008)	(0.863)	(1.014)	(0.886)
Piracy attacks, 15 km	0.000	0.179	0.086	0.179	0.092	0.187	0.104
	(0.000)	(0.508)	(0.435)	(0.068)	(0.074)	(0.068)	(0.071)
Annual oil production (million barrels)	2.374	1.991	3.509	-0.383	-1.518	0.638	-0.080
	(6.120)	(3.817)	(8.344)	(1.119)	(0.760)	(1.005)	(0.613)
Oil theft events, 15 km	3.816	2.179	2.032	-1.637	0.147	-2.263	-0.777
	(8.311)	(3.829)	(5.260)	(1.431)	(0.621)	(1.425)	(0.683)
Shut-in field	0.054	0.179	0.265	0.125	-0.086	0.090	-0.135
	(0.229)	(0.386)	(0.442)	(0.064)	(0.060)	(0.066)	(0.064)
Total conflict deaths, 15km	0.842	0.768	0.455	-0.074	0.313	-0.221	0.097
	(2.400)	(2.750)	(1.952)	(0.532)	(0.389)	(0.533)	(0.400)
NNPC share	29.459	50.536	49.581	21.076	0.954	18.001	-3.552
	(36.396)	(16.031)	(20.532)	(6.303)	(2.554)	(6.190)	(2.382)
Operator share	74.527	38.977	41.073	-35.550	-2.096	-31.617	3.667
	(29.883)	(16.391)	(23.627)	(5.336)	(2.714)	(5.571)	(2.601)
Number of clusters	38	56	220				

Columns (1)-(3) display means of variables with standard deviations in parentheses, by group. AT = always treated, D=Divested, NT = never treated. Columns (4) and (5) display differences in means with standard errors in parentheses. Columns (6) and (7) display differences in means, controlling for an onshore indicator variable. Sample is 314 oilfields. Panel A gives summary statistics of field-level covariates while Panel B gives time-invariant outcomes measured in the year that the field first enters the sample. Sample sizes indicate the number of unique oilfields in each group. Treated refers to all oilfields that have any local operator over the sample period.

THEORETICAL APPENDIX

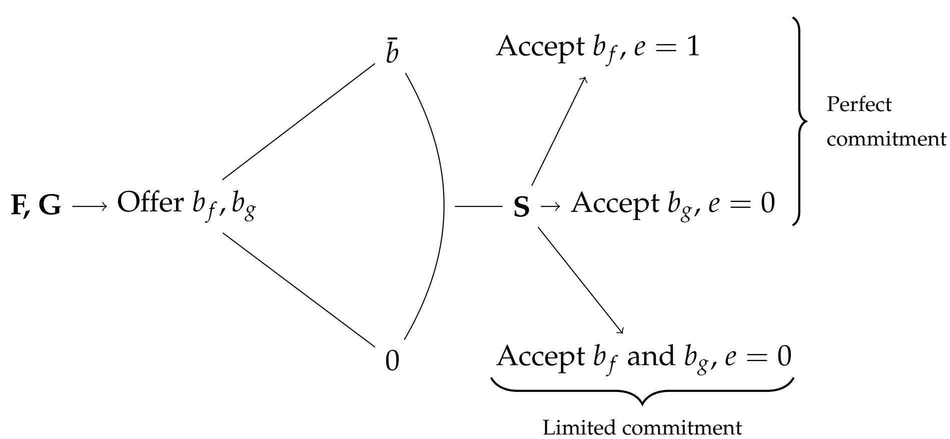
B1. Set-up

GAME STRUCTURE

The game is a sequential-move, one-shot interaction between firms f, gangs g, and state security agents s. The firm produces a fixed quantity of oil $\bar{Q}$, sold at the international oil price p^*, at cost C. Firms and gangsters simultaneously offer bribes b_f and b_g to law enforcement. Law enforcement observes the offers and decides on a level of enforcement. If law enforcement accepts the gang's bribe, oil theft is allowed, and enforcement $e = 0$. The gang steals a constant quantity $q < \bar{Q}$ at fixed cost $c - \epsilon_g$, where ϵ_g is private information. Stolen output is sold on the black market at the discounted price $\tilde{p}$. I assume theft is always profitable, $\tilde{p}q > c$, and so it will always occur in the absence of enforcement.

If law enforcement instead accepts the bribe b_f, then under *perfect commitment* they must enforce the law ($e = 1$), while under *limited commitment* they retain the option of reneging. The enforcement technology reduces the losses to oil theft (and its profits to gangs) by a factor $\alpha < 1$ at cost η. Theft is inefficient for three reasons: *i*) gangs incur costs that duplicate firms' extraction costs, $c + \epsilon_g > 0$, *ii*) it directly destroys output, denoted by $\kappa > 0$, and *iii*) since $\tilde{p} < p^*$, total oil revenue falls when the quantity q is transferred to gangs. As such, theft is therefore not merely a transfer from the firms to gangs, but entails welfare losses. Total output to the legal market Q is $\bar{Q}$ less expected transfers and spillage losses to the black market. Figure A9 diagrams the sequence of the stage game.

Figure A9. Stage game timing



Note: This figure represents stage game timing. **F**, **G**, and **S** denote firms, gangs, and security forces.

FIRM HETEROGENEITY

Firms have two types, $f \in \{m, \ell\}$, multinational and local, which may differ in several ways. If a bargain is consummated, firm f pays a penalty λ_f. This captures the fact that different firms may be subject to different legal or reputational costs of corrupt payments. In addition, firms only receive a share γ_f of $\bar{Q}$, to capture the important role of joint-ventures in Nigeria (see Figure A3). Importantly, law enforcement may internalize firm f's output up to the parameter μ_f, measuring the strength of political connections. If a firm is unconnected, then $\mu = 0$. Note that $\gamma + \mu \leq 1$. Lastly, firms may also differ in extraction cost C_f.[41]

B2. *First-best: no theft or corruption*

In the simplest case of the model, there is no threat of theft from the gangs, the state fully internalizes the cost of theft ($\mu = 1$), cannot accept any bribe, and perfectly commits to enforcement. Trivially, there is no strategic behavior, since there is no theft or bribery. Legal output is $Q = \bar{Q}$, and total surplus is $p^*\bar{Q} - C_f$. The lowest-cost firm maximizes total surplus.

B3. *First-best with theft*

In the first-best case, gangs may still steal oil. However, the state fully internalizes the cost of theft ($\mu = 1$), cannot accept any bribe, and perfectly commits to enforcement. The state's payoffs under $e = 0$ and $e = 1$ are:

$$U_s^0 = p^*(\bar{Q} - q - \kappa) \qquad\qquad U_s^1 = p^*\bar{Q} - \alpha p^*(q + \kappa) - \eta$$

Definition 1. ***Total surplus.*** *Define total surplus as a function of enforcement, $S(e)$:*

$$S(e) = p^*\bar{Q} - C + e[\alpha(q(\tilde{p} - p^*) - p^*\kappa - c) - \eta] + (1 - e)[q(\tilde{p} - p^*) - p^*\kappa - c]$$

$p^*\bar{Q} - C$ is total available surplus (value-added) from official production sold at the international price. However, the black market erodes this surplus: $q(\tilde{p} - p^*) < 0$ is the efficiency loss from the black market discount, $p^*\kappa$ is the value of spillage losses from theft, and c is the redundant black-market extraction cost. The extent of these inefficiencies is determined by the presence of enforcement e and the state's monopoly of force α. Lastly, the cost η represents real resources mobilized for enforcement when the state chooses $e = 1$. For enforcement to be socially optimal, η must be sufficiently low, yielding our next assumption.

Assumption 1. ***Efficient enforcement.*** *If enforcement is sufficiently cheap, productive, or theft sufficiently wasteful, then stopping crime is socially optimal, $S(1) > S(0)$:*

$$(1 - \alpha)(q(p^* - \tilde{p}) + p^*\kappa + c) > \eta$$

This assumption rules out pathological cases where crime is socially valuable.

Proposition 1. ***No corruption equilibrium.*** *Under Assumption 1, the state's equilibrium strategy is to enforce the law for any bribe offers. Further, total surplus is maximized by the lowest-cost producer.*

Since the state cannot be swayed by bribes, enforcement occurs if and only if the state finds it optimal, $U_s^0 < U_s^1$; that is, when the value of the oil saved by enforcement is greater than the cost of enforcement

$$(1 - \alpha)p^*(q + \kappa) > \eta$$

[41] For ease of exposition, in what follows I suppress f subscripts except when necessary for analyzing the impact of firm type.

Note that when Assumption 1 holds, then equilibrium for the first-best case is $e = 1$, since $(1-\alpha)p^*(q+\kappa) > (1-\alpha)(q(p^* - \tilde{p}) + p^*\kappa + c)$ as long as $\tilde{p}q > c$, or when theft is profitable. Since $e = 1$ in the first-best case, expected surplus is:

$$S(1) = p^*\bar{Q} - C_f - \eta + \alpha(q(\tilde{p} - p^*) - c - p^*\kappa)$$

Clearly, the welfare-maximizing choice is to allocate extraction rights to the lowest-cost producer, which is the MNC under the reasonable assumption that $C_m < C_\ell$. □

In no corruption case, $Q = \bar{Q} - \alpha(c + \kappa)$. Since we have introduced theft and the state's monopoly of violence is incomplete, this is lower than in the first-best, where all output was appropriated by legal participants. However, legal output does not depend on firm identity, since enforcement is non-discriminatory.

B4. Second-best with corruption

Assume now that the state can be corrupted by bribes, but it must perfectly commit to an enforcement strategy. The payoffs under $e = 0$ and $e = 1$ are as follows:

$$\begin{aligned} U_f^0 &= \gamma p^*(\bar{Q} - q - \kappa) - C_f & \qquad U_f^1 &= \gamma p^*\bar{Q} - \alpha\gamma p^*(q+\kappa) - \lambda - b_f - C_f \\ U_g^0 &= \tilde{p}q - c + \epsilon_g - b_g & U_g^1 &= \alpha(\tilde{p}q - c + \epsilon_g) \\ U_s^0 &= b_g + \mu p^*(\bar{Q} - q - \kappa) & U_s^1 &= b_f + \mu p^*[\bar{Q} - \alpha(q+\kappa)] - \eta \end{aligned}$$

Definition 2. ***Bargaining Range.*** *The bargaining range B is the set of firm bribes b_f for which enforcement can be sustained in equilibrium, defined as the interval $[\bar{b}_g, \bar{b}_f]$.*

$\bar{b}$ are the reservation points of gangster and firm. If $b_f < \bar{b}_g$, then the gangster is willing to pay more than the firm offers, and crime occurs with probability one. Similarly b_f must be individually rational and therefore cannot exceed $\bar{b}_f$. Using the utilities for f and g yields the reservation points

$$\bar{b}_g = (1-\alpha)(\tilde{p}q - c + \epsilon_g) \qquad \bar{b}_f = (1-\alpha)\gamma p^*(q+\kappa) - \lambda$$

Note that government rents stem directly from their partial monopoly of violence. When enforcement is ineffective, $\alpha = 1$, neither party has any incentive to bribe the security forces.

The government prefers to enforce whenever $U_s^0 < U_s^1$. This yields the reservation point

$$b_g + \mu p^*(\bar{Q} - q - \kappa) = b_f + \mu p^*[\bar{Q} - \alpha(q+\kappa)] - \eta$$

Definition 3. ***Bribe offers.*** *Assume that law enforcement extracts all of the surplus from gangsters, so that $b_g = \bar{b}_g$.*[42] *Then the threshold bribe for which government enforces is given by:*

$$b^* = (1-\alpha)(\tilde{p}q - c + \epsilon_g) + (\alpha - 1)\mu p^*(q+\kappa) + \eta$$

This expression gives us our first key prediction: since $\alpha - 1 < 0$ an increase connections μ reduces the bribe required for security agents to enforce the law. Note also that setting $\mu = \alpha = \eta = 0$ reflects the situation where firm and gangster bargain directly with each other

[42] This is without loss of generality. We could allow some fraction of the surplus to be retained by gangsters, in which case we would simply have another fractional parameter to carry around.

and gangs receive a take-it-or-leave-it offer. μ introduces a friction in favor of firms, while costs of enforcement η and its incomplete nature α introduce wedges in favor of theft.

Assumption 2. ***Information structure.*** *Assume that the firm does not observe ϵ_g until the bargaining phase, and that ϵ_g is distributed uniformly on the interval $[0, c]$.*

Enforcement occurs in equilibrium when the firm is willing to pay what the state demands, $b^* < \bar{b}_f$. Define the probability of enforcement as $\varrho = Pr(e = 1)$. Using the uniform distribution of ϵ_g, we have:

$$\begin{aligned}\varrho &= Pr(b^* < \bar{b}_f) \\ &= \frac{1}{c}[(\gamma + \mu)p^*(q + \kappa) - \tilde{p}q] - \frac{\lambda + \eta}{(1 - \alpha)c} + 1\end{aligned}$$

Proposition 2. ***Comparative statics: enforcement and theft.*** *Given Definition 3 and Assumption 2, ϱ is decreasing in $\eta, \lambda, \tilde{p}, \alpha$, and increasing in μ, γ, κ, p^*. ϱ is increasing in q whenever $\gamma + \mu > \frac{\tilde{p}}{p^*}$. Since the expected incidence of theft is simply $n = \alpha\varrho + (1 - \varrho)$, it has the same predictions in the opposite direction.*

Proof:

$$\begin{aligned}\frac{\partial\varrho}{\partial\eta} &= -\frac{1}{(1-\alpha)c} < 0 \\ \frac{\partial\varrho}{\partial\lambda} &= -\frac{1}{(1-\alpha)c} < 0 \\ \frac{\partial\varrho}{\partial\alpha} &= -\frac{\lambda + \eta}{c(1-\alpha)^2} < 0 \\ \frac{\partial\varrho}{\partial\gamma} &= \frac{p^*(q+\kappa)}{c} > 0 \\ \frac{\partial\varrho}{\partial\kappa} &= \frac{p^*(\gamma+\mu)}{c} > 0 \\ \frac{\partial\varrho}{\partial\mu} &= \frac{p^*(q+\kappa)}{c} > 0 \\ \frac{\partial\varrho}{\partial p^*} &= \frac{(\gamma+\mu)(q+\kappa)}{c} > 0 \\ \frac{\partial\varrho}{\partial\tilde{p}} &= -\frac{q}{c} < 0 \\ \frac{\partial\varrho}{\partial q} &= \frac{p^*(\gamma+\mu) - \tilde{p}}{c} > 0 \text{ whenever } \gamma + \mu > \frac{\tilde{p}}{p^*}\end{aligned}$$

□

Legal output in the second-best case is

$$Q = \bar{Q} - \varrho\alpha(q+\kappa) - (1-\varrho)(q+\kappa) = \bar{Q} - (\varrho\alpha + (1-\varrho))(q+\kappa) = \bar{Q} - n(q+\kappa)$$

This is lower than in the no-corruption case, since $n = \varrho\alpha + (1 - \varrho) > \alpha$, as $\alpha, \varrho < 1$. Additional surplus is now appropriated by the black market because of the bargaining frictions, which lead to incomplete enforcement. Since enforcement is discriminatory, total legal output now depends on firm identity f, as $\frac{\partial \varrho}{\partial \lambda} < 0$, $\frac{\partial \varrho}{\partial \kappa} > 0$, and $\frac{\partial \varrho}{\partial \gamma} > 0$.

B5. Second-best without commitment

The previous environment assumes that all contracts can be perfectly enforced. In violent, anarchic environments like the Niger Delta, a no-commitment assumption is more plausible. To economize on notation assume that $\epsilon_g = 0$ for all g and that $\tilde{p} = p^*$. Now, the security agent has a third action available: accept a bribe from both parties and renege on the agreement with the firm.

For illustration, consider the case when $\mu = 0$. Then $b_g + b_f > b_f - \eta$ and $b_g + b_f > b_g$, so accepting both bribes and allowing theft is the dominant strategy for the government at any bribe. As such, the firm will always to obtain payoff U_f^0 for any offer. Therefore, setting $b_f > 0$ and incurring the cost of corruption λ can never be optimal for the firm, since $U_f^0 - \lambda - b_f < U_f^0$. Therefore, without commitment, bribes are ineffective and political connections are a necessary condition to sustain enforcement. This leads to a more general proposition.

Proposition 3. ***Enforcement without commitment***. *Assume a no-commitment environment, and assume that the behavior of the gang is fixed at $b_g = \bar{b}_g$. Then there are two possible outcomes of the stage game, each a unique Nash equilibrium. Let $\bar{\mu} = \frac{(1-\alpha)(pq-c)+\eta}{(1-\alpha)p(q+\kappa)}$. When $\mu \geq \bar{\mu}$, the government accepts any firm bribe offer and sets $e = 1$, and the firm sets $b_f = 0$. When $\mu < \bar{\mu}$, the government accepts both firm and rebel bribe offers and sets $e = 0$, and the firm sets $b_f = 0$.*

Proof: When the firm is politically connected, the incentives can align for sufficiently large μ. In particular, for $e = 1$ to be a dominant strategy, the payoff to the security forces from following the agreement must exceed that of reneging and accepting both bribes:

$$U_s^1 \geq U_s^0 + b_f > U_s^0$$

Which yields the condition

$$\mu \geq \frac{(1-\alpha)(pq-c)+\eta}{(1-\alpha)p(q+\kappa)} = \bar{\mu}$$

When this condition is met, the government has a dominant strategy. For any $b_f \geq 0$, accepting the bribe and enforcing is a best response, since μ is such that that the government sufficiently internalizes theft losses. Knowing this, the firm will set $b_f = 0$ to maximize its payoff. Therefore, the Nash equilibrium is unique.

Clearly, when $\mu < \bar{\mu}$ then we have $U_s^1 < U_s^0 + b_f$ and of course $U_s^0 + b_f \geq U_s^0$. So $e = 0$ is a dominant strategy for the government for any b_f. Again, the firm must set $b_f = 0$ because $U_f^0 - \lambda - b_f < U_f^0$. The profitability of theft is a sufficient condition for political connections to be a binding constraint on enforcement, because $pq - c > 0 \implies \bar{\mu} > 0$. □

One implication is that when no firms are not politically connected, $\mu_f = 0$ for all f, no security will be offered to any firm, and so there is no local advantage. But if $\mu_\ell > \bar{\mu}$ while $\mu_m = 0$, local advantage arises, and is driven entirely by connections.

Adding back in $\tilde{p}$, p^* and ϵ_g and evaluating the over uniform distribution yields a probability of enforcement:

$$\varrho^{NC} = \frac{\mu p(q+\kappa) - \tilde{p}q}{c} - \frac{\eta}{(1-\alpha)c} + 1$$

As before, legal output is now $Q = \bar{Q} - (\varrho^{NC}(\alpha - 1) + 1)(q + \kappa)$. Whether enforcement is lower without commitment is ambiguous. This is because ϱ (and Q) no longer depends on γ and λ, but solely on political connections μ; $\varrho^{NC} = \varrho$ when $\lambda = \gamma = 0$. While $\lambda = 0$ increases enforcement relative to the commitment case, $\gamma = 0$ reduces it. Since the state is not responsive to bribes, bargaining frictions affecting the firm no longer drive outcomes.

Table A5—Summary of theoretical predictions

Outcome	First-best	No corruption	Second-best, commitment	Second-best, no-commitment
Equilibrium bribe offer b_f	0	0	$(1-\alpha)(\tilde{p}q - c + \epsilon_g) + (\alpha-1)\mu p^*(q+\kappa) + \eta$	0
Enforcement				
ϱ	0	1	$\frac{1}{c}[(\gamma+\mu)p^*(q+\kappa) - \tilde{p}q] - \frac{\lambda+\eta}{(1-\alpha)c} + 1$	$\frac{1}{c}[\mu p^*(q+\kappa) - \tilde{p}q] - \frac{\eta}{(1-\alpha)c} + 1$
$\frac{\partial\varrho}{\partial\mu}$	0	0	+	+
$\frac{\partial\varrho}{\partial\lambda}$	0	0	+	0
$\frac{\partial\varrho}{\partial\gamma}$	0	0	+	0
Theft				
n	0	α	$\varrho(\alpha-1)+1$	$\varrho^{NC}(\alpha-1)+1$
$\frac{\partial n}{\partial\mu}$	0	0	−	−
$\frac{\partial n}{\partial\lambda}$	0	0	−	0
$\frac{\partial n}{\partial\gamma}$	0	0	−	0
Legal output				
Q	$\bar{Q}$	$\bar{Q} - \alpha(q+\kappa)$	$\bar{Q} - (\varrho(\alpha-1)+1)(q+\kappa)$	$\bar{Q} - (\varrho^{NC}(\alpha-1)+1)(q+\kappa)$
$\frac{\partial Q}{\partial\mu}$	0	0	+	+
$\frac{\partial Q}{\partial\lambda}$	0	0	+	0
$\frac{\partial Q}{\partial\gamma}$	0	0	+	0

Table A5 summarizes the equilibrium outcomes of the four model cases and comparative statics with respect to firm heterogeneity. The key testable implications of the model are as follows: assuming that $\mu_\ell > \mu_m$, then, all else equal, local firms should see increased enforcement, reduced theft, and increased output in any second-best equilibrium. This is what I refer to as the "local advantage." In the no-commitment environment, local advantage is driven *only* by μ, while under commitment both γ and λ also play a role, assuming either $\gamma_\ell > \gamma_m$ and/or $\lambda_\ell < \lambda_m$. Note the bindingness of political connections when we move from commitment to no-commitment: $\mu > 0$ is a necessary condition for $\varrho^{NC} > 0$, but not for ρ.

Importantly, C_f does not determine the firm's bribe, and so does not enter the equilibrium outcomes. This may seem counter-intuitive, as more efficient firms should be able to withstand more theft and remain in business. The explanation is simple: costs are sunk and output is fixed; extraction costs are incurred whether or not enforcement occurs and so do not enter the firm's willingness-to-pay. Further, I do not model the extensive margin decision to operate; predictions are therefore *conditional* on the firm deciding to operate. This somewhat restrictive set of assumptions is driven by the empirical setting; I focus on quantities that are observed – connections, ownership structure – rather than those that aren't, like costs.

B6. Extension: dynamic bargaining

The assumption of a one-shot game without commitment makes sustaining cooperation impossible, and may be too extreme. Instead, consider the game with no commitment, repeated infinitely. The players have a common discount factor δ and for simplicity let $\alpha = 0$, $p^* = \tilde{p} = p$, and $\epsilon_g = 0$. Then enforcement may occur even when $\mu < \bar{\mu}$.

Proposition 4. ***Dynamic enforcement****. Let $\mu < \bar{\mu}$. Then for sufficiently large δ, law enforcement provision can be sustained in a subgame perfect equilibrium of the infinitely repeated game where government cannot commit in the stage game.*

Proof: First note that when $\mu \geq \bar{\mu}$, enforcement is sustained in subgame perfect equilibrium by playing the Nash equilibrium of the stage game in every period. When $\mu < \bar{\mu}$, enforcement is no longer a Nash equilibrium of the stage game. Nevertheless, it can be restored with a simple trigger strategy profile: the firm begins by offering $b_f = b^*$ and continues to do so in every period until the cooperative outcome is not played, after which the firm sets $b_f = 0$ forever. The government accepts all bribes $b_f \geq b^*$ and responds with $e = 1$. After any period in which the cooperative outcome is not played, government sets $e = 0$ forever.

Since $\mu < \bar{\mu}$, the punishment is the stage game Nash and so is subgame perfect after a deviation. The value to the security forces of playing the punishment equilibrium is:

$$r_s = \sum_{t=0}^{\infty}(pq - c + \mu p(\bar{Q} - q - \kappa))^{\delta} = \frac{pq - c + \mu p(\bar{Q} - q - \kappa)}{1 - \delta}$$

Given a bribe b_f, security forces are willing to enforce the law rather than deviate and allow theft whenever:

$$b_f + (1 - \delta)r_s + \delta r_s \leq \frac{b_f + \mu p\bar{Q} - \eta}{1 - \delta}$$

Solving for b_f gives us the minimal bribe that the government is willing to accept for the equilibrium to be sustained.

$$b^* = \frac{1}{\delta}(pq - c - \mu p(q + \kappa) + \eta)$$

Note that this is similar to the minimum bribe in the base case. However, in the dynamic game, the minimum per-period rent transferred to the state must be inflated by a factor of $\frac{1}{\delta}$ relative to the minimal transfer in the one shot game with commitment, since now it must be enforced with dynamic incentives. The firm's value of punishment:

$$r_f = \sum_{t=0}^{\infty}(\gamma p(\bar{Q} - q - \kappa))^{\delta} = \frac{\gamma p(\bar{Q} - q - \kappa)}{1 - \delta}$$

The firm must be willing to set $b_f > b^*$ rather than set $b_f = 0$ and induce punishment. So the firm's incentive condition is

$$r_f \leq \frac{\gamma p\bar{Q} - \lambda - b_f}{1 - \delta}$$

Yielding the same maximal willingness to pay as the 2nd-best case with commitment:

$$b_f = \bar{b}_f = \gamma p(q + \kappa) - \lambda$$

Importantly, note that this condition is identical because the firm has no commitment problem, given the structure of the stage game.[43]

As before, efficient corruption occurs whenever $\bar{b}_f \geq b^*$. This implies the condition:

$$\delta \geq \bar{\delta} = \frac{(pq - c - \mu p(q+\kappa) + \eta)}{\gamma p(q+\kappa) - \lambda}$$

Note that $\mu < \bar{\mu}$ implies that $\bar{\delta} > 0$, so the incentive constraint binds. □

Now we can slightly revise the predictions of Proposition 2 to say that the enforcement equilibrium becomes *more likely* and theft becomes *less likely* as $\bar{\delta}$ falls.

Proposition 5. ***Comparative statics: dynamic enforcement***. *Let $\mu < \bar{\mu}$. Say that the likelihood of enforcement is decreasing in $\bar{\delta}$. Then the comparative statics from Proposition 2 all hold in the no-commitment dynamic bargaining game.*

The proof is immediate, since $\delta > \bar{\delta} \iff \bar{b}_f > b^*$. Similarly, Proposition 2 relies on the condition that $\bar{b}_f > \delta b^*$ and $\delta > 0$. □

B7. *Indigenization and welfare in the second-best case*

Following Definition 1, expected surplus in the second-best case is:

$$S = p^*\bar{Q} - C + \varrho[\alpha(q(\tilde{p} - p^*) - p^*\kappa - c) - \eta] + (1-\varrho)[q(\tilde{p} - p^*) - p^*\kappa - c]$$

Collecting terms, we have:

$$S = p^*\bar{Q} - C + (\varrho\alpha + 1 - \varrho)(q(\tilde{p} - p^*) - p^*\kappa - c) - \varrho\alpha\eta$$

Now recall expected theft is $n = \varrho\alpha + (1-\varrho)$. Furthermore note that n and ϱ are both functions of the firm-specific variables in the model. Finally, allow $\bar{Q}$ to vary by f as well, to capture local-multinational output differences that may be "outside" the model. This yields

$$S_f = p^*\bar{Q}_f - C_f + n_f(q(\tilde{p} - p^*) - p^*\kappa - c) - \varrho_f\alpha\eta$$

Allocating assets to the lowest-cost firm is no longer necessarily optimal, because the firm's identity affects not only costs, but also the equilibrium level of oil theft and enforcement. There are also externalities in the sector. $V^1 = n_f e_n + m_f e_m + g_f e_g$ is a local environmental externality, which depends on oil spills and gas flares. Gas flares depend on the firm's type g_f, with per-flare cost e_g while oil spills depend on the firm directly in operational malfunctions m_f and indirectly via the level of oil theft n. Note that the per-spill, externality cost is allowed to vary by spill type. $V^2 = Q_f e_c$ is a global environmental externality that depends on the level of oil output and the social cost of carbon. $V^3 = v_f e_v$ is a violence externality that depends on the level of violence in the black market and the cost of violence. Lastly, $VA = VA_n n_f$ captures value-added from local refining activities, multiplied by the size of the black

[43]If the firm can deviate in the stage game and enjoy a single period of bribe-free enforcement, the the incentive condition becomes: $\gamma p\bar{Q} + \delta r_f \leq \frac{\gamma p\bar{Q} - \lambda - b_f}{1-\delta}$, yielding a maximal willingness to pay of $\tilde{b}_f = \delta\gamma p(q+\kappa) - \lambda$

market.[44] Total surplus for firm f is:

$$S_f = p^* \bar{Q}_f - C_f + n_f(q(\tilde{p} - p^*) - p^*\kappa - c) - \varrho_f \alpha\eta + V_f^1 + V_f^2 + V_f^3 + VA_f$$

With some abuse of notation, for any function of firm type y_f, define the change due to indigenization, $\frac{\partial y}{\partial f} = y_\ell - y_m$. The change in welfare due to the indigenization policy will be:

$$\Delta S = \sum_{f=\ell} \frac{\partial Q_f}{\partial f}(p^* + e_c) - \frac{\partial C_f}{\partial f} + \frac{\partial n_f}{\partial f}[q(\tilde{p} - p^*) - p^*\kappa - c + e_n + VA_n] - \frac{\partial \varrho_f}{\partial f}\alpha\eta + \frac{\partial m_f}{\partial f}e_m + \frac{\partial g_f}{\partial f}e_g + \frac{\partial v_f}{\partial f}e_v$$

Table A6 provides a summary of welfare effects and the expected signs.

Table A6—Summary of welfare effects

Description	Parameter	Welfare effect
Output effect	$\frac{\partial Q_f}{\partial f}p^*$	$+$
Carbon externality	$\frac{\partial Q_f}{\partial f}e_c$	$-$
Black market crude discount	$\frac{\partial n_f}{\partial f}q(\tilde{p} - p^*)$	$+$
Oil theft spillage loss	$\frac{\partial n_f}{\partial f}p^*\kappa$	$+$
Black market extraction cost	$\frac{\partial n_f}{\partial f}c$	$+$
Oil theft spillage externality	$\frac{\partial n_f}{\partial f}e_n$	$+$
Black market refining value added	$\frac{\partial n_f}{\partial f}VA_n$	$-$
Enforcement cost	$\frac{\partial \varrho_f}{\partial f}\alpha\eta$	$-$
Gas flare externality	$\frac{\partial g_f}{\partial f}e_g$	$-$
Violence externality	$\frac{\partial v_f}{\partial f}e_v$	$+$

[44]Note that value-added is external to the participants in the model, as qualitative research suggests that the illicit value chain is not vertically integrated (SDN, 2019*b*). In theory, value-added in the local illicit refining sector might affect $\tilde{p}$. However, such general equilibrium effects are beyond the scope of this paper.

Robustness tests: main outcomes

C1. *Oil spills output adjustment*

The results of Table 1, Panel B may not be driven by lower operating quality among local firms. Instead, the increase in per-field output may mechanically be driving up equipment failure-related oil spills. As a robustness test, I consider whether the increase in output is large enough to explain the effect on oil spills. Let γ denote the marginal effect of an additional (million) barrels of oil annually on recorded oil spills. In column (1) of Table A7, I estimate this quantity using a simple fixed effects regression of oil spills on output, controlling for field and year fixed effects.

Table A7—Divestment and oil spills: oil production adjustment

	γ	ψ_y	ψ_s	$\psi_s - \psi_y\gamma$
Estimate	0.130***	0.840***	1.380*	1.271*
	(0.042)	(0.297)	(0.765)	(0.759)

Standard errors in parentheses are clustered at the field level. Sample is the same as in Table 1. Parameters are in the table header, with y and s indexing the output and oil spills outcomes, respectively. Estimation of the three-equation system is conducted jointly with seemingly unrelated estimation for nonlinear hypothesis testing across equations. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Now recall the treatment effect of localization, ψ and index it by y for the output outcome and s for the oil spills outcome. Columns (2) and (3) display these effects, taken from Table 1. Finally, column (4) subtracts from ψ_s the implied increase in spills that would result only from the output gains, $\psi_y\gamma$. The three-equation system is estimated jointly to enable nonlinear hypothesis testing. In (4), once we account for this effect localization is still followed by increases in oil spills – the estimate falls only slightly and remains significant at the 10% level.

C2. *State-owned vs. private firms*

The main results in Table 1 include all non-multinational firms in "local." In Table A8 I disaggregate separate treatment indicators for fields operated the NPDC – the state oil company – and those operated by independent local firms. I find that the positive effect on output is driven almost exclusively by private firms. In contrast, the efficiency costs of localness in terms of greater malfunctions essentially vanishes when we disaggregate the treatment, with a small insignificant point estimate, while the effect size rises to 3.9 for state-run fields. At the same time, the reductions in oil theft is also large and significant for private firms but insignificant for the government. Private local firms appear to have no efficiency disadvantage, magnifying the output benefits of localness. In contrast, the efficiency costs of public production are quite large and the benefits minimal, resulting in a smaller output effect.

C3. *Theft-output elasticity*

The primary interpretation of the results in this paper is that improvements in the security situation and consequent reductions in oil theft lead to output advantage for local firms. This

Table A8—Divestment and field-level outcomes: state-owned vs. private firms

Outcome	Output		Oil spills		Oil theft	
	(1)	(2)	(3)	(4)	(5)	(6)
Private local operator	0.901***		0.390		-5.619***	
	(0.312)		(0.809)		(1.055)	
Government operator		0.145		3.939***		0.867
		(0.499)		(0.665)		(1.969)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2476	2476	3183	3183	3183	3183
R^2	0.878	0.878	0.649	0.650	0.756	0.753

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Private local operator is an indicator that the operator is a private Nigerian firm in a given field-year. Government operated is an indicator that the operator is the NPDC/NNPC in a given field-year. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

is supported by both the time-path of effects in the main event-study and the analysis of heterogeneous effects by asset type. Still, it is possible that different features of local firms' operations cause both increased output and reduced theft simultaneously. One important falsification test for the causal mechanism is to determine whether the the increase in output quantitatively consistent with the increase oil theft. That is, given the elasticity of oil output to theft, how much of the increase in output can be "explained" by the fall in oil theft.

First, define the following three parameters of interest: $\eta_{y,b}$ is the impact of black-market theft b on output y, ψ_y is the ATE of local ownership on output, and ψ_b is the ATE of ownership on oil theft. Then the "residual" increase in output that cannot be quantitatively explained by the reduction in theft is $\psi_y - \eta_{y,b}\psi_b$, and the total share of ψ_y explained by oil theft is $\frac{\eta_{y,b}\psi_b}{\psi_y}$. Of course, we have estimates for ψ_y and ψ_b from Table 1. However, we do not have reliable estimates for $\eta_{y,b}$, the causal effect of theft on oil output.

I estimate $\eta_{y,b}$ in Table A9 using an instrumental variables approach. Oil theft and output are equilibrium outcomes, likely exhibiting both reverse causality and omitted variables bias. Identification of $\eta_{y,b}$ requires an exogenous shock that alters incentives in the oil black market but does not directly affect oil production decisions. One such shock is the national energy market. Despite producing 2-2.5 million barrels of crude oil per day, Nigeria meets the vast majority of domestic fuel demand through imports. As Rexer and Hvinden (2022) show, the period between 2006-2016 was one of steadily worsening domestic fuel shortages, a result of shrinking domestic refining capacity, mismanagement and corruption in import market, and increasingly unsustainable fuel subsidies. This fuel crisis has coincided with the aggregate growth in oil theft *and* a shift toward supplying the domestic market SDN (2019*a*). Still, since Nigeria exports nearly 90% of its oil output, these domestic market conditions should not affect production incentives except by increasing black market oil theft.

I measure fuel shortages by the log of aggregate refined gasoline imports. However, this

Table A9—Oil theft and output: instrumental variables estimation

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: First stage						
Log gasoline imports × Distance to Atlantic coast (km)	-0.217***	-0.324***	-0.235**	-0.348***	-0.521***	-0.603***
	(0.075)	(0.111)	(0.103)	(0.081)	(0.117)	(0.162)
Observations	2476	2476	2476	2476	2476	2476
R^2	0.723	0.728	0.761	0.749	0.761	0.773
Panel B: Reduced form						
Log gasoline imports × Distance to Atlantic coast (km)	0.050**	0.058**	0.071***	0.038***	0.040**	0.039***
	(0.024)	(0.027)	(0.025)	(0.014)	(0.016)	(0.014)
Observations	2476	2476	2476	2476	2476	2476
R^2	0.862	0.862	0.873	0.869	0.869	0.879
Panel C: 2SLS						
Oil theft events, 15 km	-0.232	-0.180	-0.301*	-0.110**	-0.076**	-0.065**
	(0.155)	(0.116)	(0.179)	(0.047)	(0.034)	(0.028)
F-statistic	8.242	8.530	5.242	18.569	19.998	13.938
Observations	2476	2476	2476	2476	2476	2476
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Amnesty controls	No	Yes	Yes	No	Yes	Yes
Main controls × Year FE	No	No	Yes	No	No	Yes
Militant controls × Year FE	No	No	No	Yes	Yes	Yes

Standard errors in parentheses are clustered at the field level. Sample is the panel of 2476 field-years from 2006-2016 for which oil output information is available. Outcome in Panel A is oil theft, measured as the total number of sabotage spills within 15 km of the field. Outcome in Panels B and C is oil output, measured in millions of barrels of oil per year. Main controls are latitude of the field centroid, distance to Niger River, and distance to the capital. Amnesty controls is the interaction between a post-2009 indicator and distance to the coast. Militant controls includes distance to the nearest militant camp. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

quantity varies only at the national level over time, while the instrument must have field-specific variation. To generate cross-sectional variation, I interact the national trend in log gasoline imports with distance to the coast. Rexer and Hvinden (2022) show that distance to the coast is a reasonable proxy for black market costs, since coastal locations are proximate to coastal waterways used for illegal transport. As imports rise and fuel shortages are alleviated, black market margins shrink and oil theft becomes unprofitable first in higher-cost inland locations, while remaining profitable in low-cost coastal regions. As such, we should expect a negative first stage coefficient on the interaction between national import trends and distance to the coast (i.e., alleviation of fuel crises reduces oil theft more in further inland locations). This is exactly the result of Table A9, Panel A, which shows first stage estimates with different combinations of controls, all of which include two-way fixed effects. All estimates are significant at 1% suggesting a relevant instrument. The instrument is strongest in columns (4) and (5), achieving an F-statistic of 19-20. These specifications control for militant presence (4) and amnesty policy (5), both of which may be correlated with costs and trends in oil importation.

The reduced form shows consistent results in Panel B. Alleviating gasoline shortages is associated with differentially large increases in output in the low-coast inland locations where theft falls. The reduced form effects are significant at 5 or 1% in all specifications. Finally, Panel C estimates a 2SLS model using the interaction between log national gas imports and distance to coast as an instrument for theft, conditioning on year and field fixed effects. The results indicate a robustly negative relationship between theft and oil output, ranging 0.07-0.23 million fewer barrels annually per additional theft incident. These estimates are only significant at conventional levels for specifications in (4)-(6) with high first-stage F-statistics.

Table A10—Theft-output elasticity

	$\eta_{y,b}$	ψ_b	ψ_y	$\psi_y - \eta_{y,b}\psi_b$	$\frac{\eta_{y,b}\psi_b}{\psi_y}$
Estimate	-0.110**	-5.932***	0.929***	0.277	0.701*
	(0.046)	(1.110)	(0.323)	(0.418)	(0.381)

Standard errors in parentheses are clustered at the field level. Sample is the panel of 2476 field-years from 2006-2016 for which oil output information is available. $\eta_{y,c}$ is the effect of oil theft on oil output, from the 2SLS specification in column (2), Panel C in Table A9. $\eta_{c,l}$ is the effect of local ownership on oil theft. $\eta_{y,l}$ is the effect of local ownership on oil output. All equations include fixed effects for field and year and control only for militant presence interacted with year dummies. Estimation of the three-equation system is conducted jointly with GMM for nonlinear hypothesis testing. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

I choose column (4) of Table A9 as the preferred estimate of $\eta_{y,b}$ primarily because it sits in the middle of the range of magnitudes and obtains a strong first stage, however, one can easily repeat the exercise with different estimates. For Table A10, I jointly estimate all of the parameters using a GMM system with three moment conditions and the instrument for theft, on the sample of field-years with nonmissing output. These estimates are in columns (1)-(3).[45] I estimate standard errors for columns (4) and (5) using the delta method. The main output effect is 0.929, significant at 1%, while column (4) shows the residual effect after removing the role of theft is only 0.277, and not significantly different from zero. This implies that up to 70.1% of the output effect can be explained by theft (column 5), significant at the 10% level.

C4. *Output decomposition*

As mentioned in Section II, oil production is net of theft losses, and so the increase in output may represent either a transfer from black market to official, an output gain, or both. In Appendix C.C3, I use an instrumental variables approach to estimate that, quantitatively, the reduction in oil theft in Panel C is large enough to explain roughly 70% of the increase in output in (2). Therefore, at least 30% of the output gain is new production.

The remaining 70% has three components. First, there is spillage from theft, which is a pure loss, now recouped.[46] Second, firms may recoup production otherwise deferred in the face of oil theft and pipeline repairs.[47] While this deferred oil remains in the ground, it represents a deviation from optimal production. In expectation, this output will be deferred to lower price periods given the positive correlation between price and oil theft seen in Figure A5. Finally, there is output transferred back to firms and the state from gangs. Importantly, output transferred to the black market via theft will be sold at a substantial discount of the global price, reducing the total surplus (SDN, 2019*b*).

[45] Each equation controls for all of the variables indicated in Table A10 column (4).

[46] In this exercise, I ignore spillage losses as negligible; the NOSDRA data on oil spills shows that the average sabotage spill releases just 84 barrels of oil.

[47] In 2016, a total of 144 billion barrels of oil were deferred due to oil theft, representing nearly 20% of realized output (NEITI, 2016).

I decompose the per-field average increase in output as follows. First, I split the effect into the explained and unexplained shares. Second, I further decompose the explained component into transfers and deferred production. To identify this magnitude, I use data on the total quantity of theft losses, measured in 2016 from NEITI (2016). Let L be the aggregate loss and Q the total output quantity. Then total actual oil production is $Q + L$ and the share lost to theft is $\frac{L}{Q+L} = 13.5\%$. If average field-year production in millions of barrels is $\bar{q} = \frac{1}{NT}\sum_{i,t} q_{it} = 2.77$, the expected gain from completely eliminating oil theft is $\frac{\bar{q}}{1-\frac{L}{Q+L}} - \bar{q} = 2.76 \times \left(\frac{1}{0.865} - 1\right) = 0.432$. Finally, since the percentage treatment effect on oil theft is $\frac{\psi}{\bar{y}} = 0.53$, then the recovered transfer losses from reducing theft by 53% should be $0.53 \times 0.432 = 0.23$ million barrels. The residual of the explained effect is then assumed to comprise regained deferred production.

Table A11—Output decomposition

	Output (mbbl) (1)	Output (SE) (2)	Output (%) (3)	Rev (MUSD) (4)	Rev (%) (5)
Total effect	0.944	0.331	100.000		
Unexplained	0.318	0.340	33.677		
Explained	0.626	0.254	66.323		
Transfers	0.232	0.041	24.526	21.998	171.888
Deferred production	0.395	0.240	41.797	3.598	28.112

Table shows a decomposition of the effect of indigenization on output. Unexplained and explained are the output effects, estimated using the elasticity of output to theft, according to the method in Appendix C.C3. Transfers are technical losses from oil theft and sabotage, including both oil spillage and quantity transferred to the black market. Deferred production is the residual of explained output after subtracting transfers. Oil prices are in 2016 USD.

The results of this decomposition are in Table A11. Each row gives a separate component of the total effect in the top row. The columns are the number of barrels (1), the standard error of this estimate (2), the share of as a fraction of the total effect on output (3), the revenue effect (4), and the revenue effect as a percentage of average revenue (5). Revenue effects are included only for the components of the effect explained by oil theft. I value the cost of deferred production by estimating the model $p_t = \beta y_t + e_t$ where y is aggregate oil spills and p is the world oil price, measured monthly ($\beta = 0.237$, $se = 0.064$). I then calculate the expected price differential between high and low theft periods as $\beta \times IQR(y_t)$. Standard errors are clustered at the field level and account for covariances across models.

Transfers are smaller in quantity than deferred production, accounting for 25% and 42% of the total output gain, respectively. However, the value of the transfer is larger, at 22 million USD per field annually, or 10.4% of average annual field revenue. This is because deferred production is valued using the price differential between low and high theft periods, while transfers are valued at prevailing oil prices.[48] The results illustrate that incentive effects of oil theft on producers may be large, suggesting meaningful welfare gains from indigenization rather than simple transfers. A more extensive welfare analysis is provided in Section VII.B.

[48]All oil prices are normalized to 2016 US dollars.

C5. *Other outcomes*

I consider other outcomes in Table A12. This Table is described in detail in Section V.C.

Table A12—Divestment and other oil field outcomes

Outcome	Conflict deaths		Piracy		Shut-in		Gas flaring	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local firm	-3.043***	-3.203***	-0.111*	-0.097	-0.004	0.010	0.585**	0.443*
	(1.052)	(0.993)	(0.063)	(0.065)	(0.056)	(0.057)	(0.239)	(0.235)
Control group mean	2.006		0.154		0.232		1.092	
Observations	3183	3183	3183	3183	2476	2476	1503	1503
R^2	0.232	0.317	0.227	0.311	0.666	0.678	0.896	0.900
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls $\times$ Year FE	No	Yes	No	Yes	No	Yes	No	Yes

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Conflict deaths are the total number of conflict-related fatalities within 15 km of the field. Piracy is the pirate attacks within 15 km of the field. Shut-ins is an indicator for nonzero production in a field-year. Gas flaring is measured in million mscf. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C6. *Robustness to confounders*

It is possible that our results are driven by field life-cycle effects. As shown in in Figure A19, oilfields have a standard life-cycle of production. If firms are more likely to divest earlier in the cycle as production is increasing, this might drive spurious output results. This is unlikely, since optimal firm behavior would suggest that firms exploit an asset at least up to its peak production to cover fixed costs. Table A4 confirms this, showing that divested assets tend to be older on average. However, to dispel any lingering concerns, in Table A13 I re-estimate the main results including field age effects interacted with year dummies. The results are, if anything, stronger.

Another possible threat to identification is that there may be selection into field takeover based on field characteristics. If multinationals abandoned fields with these characteristics because they were experiencing differential trends in output and theft over the sample period, this could contaminate the results. In Table A14, I test robustness to including interactions between fixed field characteristics and time dummies in the main TWFE equation. Note that the sample size falls to 2,392 field-years for output and 3,038 for other outcomes because 17 fields have missing characteristics. Despite this, the results are unchanged.

Multinationals may be subject differential firm-specific trends in management, corporate practices, or macro shocks that generate incentives for divestment. These multinational-specific effects may spuriously generate observed treatment effects if these changes at the corporate level correlate with both divestment behavior and field-level outcomes over time. One strategy to address this potential source of endogeneity is to allow for unrestricted multinational-by-year trends, where the multinational is defined as a fixed, field-level characteristic measuring the MNC that operates the asset at the time of divestment, or for its lifetime in the

Table A13—Divestment and oil field outcomes: field lifecycle effects

Outcome	Output		Oil spills		Oil theft		Conflict deaths	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local firm	1.287***	1.560***	2.037**	1.354	-5.558***	-5.671***	-4.008***	-3.680***
	(0.382)	(0.369)	(0.928)	(0.920)	(1.395)	(1.319)	(1.270)	(1.181)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Age FE × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2392	2392	3038	3038	3038	3038	3038	3038
R^2	0.908	0.919	0.665	0.719	0.765	0.798	0.350	0.434

Standard errors in parentheses are clustered at the field level. Sample is the panel of 297 oilfields from 2006-2016 for which we have data on the date of first drilling in order to calculate field age. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Conflict is the total number of conflict deaths within 15 km of the field reported by local news media. Main controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A14—Divestment and oil field outcomes: field-level covariates

Outcome	Output		Oil spills		Oil theft	
	(1)	(2)	(3)	(4)	(5)	(6)
Local firm	0.687**	0.858**	1.735**	1.524*	-6.831***	-6.579***
	(0.291)	(0.356)	(0.732)	(0.785)	(1.038)	(1.101)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Main controls × Year FE	No	Yes	No	Yes	No	Yes
Field controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2392	2392	3038	3038	3038	3038
R^2	0.870	0.885	0.624	0.669	0.737	0.763

Standard errors in parentheses are clustered at the field level. Sample is the panel of 297 oilfields from 2006-2016 for which field-level covariates are available. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Main controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. Field controls are number of wells, initial year, onshore dummy, and maximum well depth. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

sample, if never divested. As such, this variable can only be defined for two groups – switchers originally owned by an MNC, and never-treated fields. The analysis in Table A15 therefore comprises only the 276 fields in these groups, dropping the always-local fields.

Panel A uses the sample of 276 oilfields. Columns (1) and (2) reveal that controlling for MNC-by-year effects cuts the output effect by more than half, and it is no longer significant. However, we observe no such difference for the oil spills, theft, and conflict outcomes in columns (3)-(8); the magnitude and significance of the estimates does not change with MNC-specific trends. Panel B shows that when we restrict to the onshore sample – where local advantage is the strongest – we see a much smaller difference between the estimates in columns

Table A15—Divestment and oil field outcomes: multinational effects

Outcome	Output		Oil spills		Oil theft		Conflict deaths	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Full sample				
Local firm	0.960***	0.431	1.565**	1.890**	-5.136***	-7.104***	-3.232***	-2.958***
	(0.317)	(0.272)	(0.777)	(0.792)	(0.837)	(1.030)	(1.060)	(1.104)
Observations	2282	2282	2900	2900	2900	2900	2900	2900
R^2	0.867	0.874	0.583	0.673	0.712	0.767	0.236	0.271
				Panel B: Onshore fields				
Local firm	0.736***	0.698***	1.454*	2.079**	-7.233***	-7.320***	-4.011***	-3.074**
	(0.275)	(0.262)	(0.797)	(0.840)	(1.039)	(1.067)	(1.166)	(1.184)
Observations	1598	1598	2097	2097	2097	2097	2097	2097
R^2	0.792	0.797	0.625	0.687	0.708	0.751	0.233	0.263
				Panel C: Offshore fields				
Local firm	0.255	-0.738	2.888***	2.222	0.054*	0.040	-0.011	-0.031
	(0.730)	(0.824)	(0.905)	(1.380)	(0.027)	(0.043)	(0.012)	(0.036)
Observations	684	684	803	803	803	803	803	803
R^2	0.867	0.884	0.572	0.671	0.184	0.335	0.102	0.161
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No	Yes	No	Yes	No
MNC FE $\times$ Year FE	No	Yes	No	Yes	No	Yes	No	Yes

Standard errors in parentheses are clustered at the field level. Sample is the panel of 276 oilfields from 2006-2016 which are either never-treated or switch from MNC to local control during the sample period. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Conflict is the total number of conflict deaths within 15 km of the field reported by local news media. Main controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

(1) and (2); the output effect remains large and significant in the presence MNC-by-year effects. This implies that the reduction in Panel A is driven by offshore fields, where we have already established a local disadvantage (see Table 2). Indeed, Panel C estimates the model with MNC-year effects for the offshore sample. As in the main results, there are positive effects for oil spills, but no effects on oil theft or conflict. Interestingly, in column (2), coefficient on output is becomes large and *negative* after conditioning on MNC trends, though imprecisely estimated. This negative effect clearly drives the insignificant output response in the full sample (Panel A). Taken together, the results suggest that controlling for MNC-by-year fixed effects does not materially affect the overall story, but it perhaps further underscores the local disadvantage offshore.

Another source of endogeneity is that firm rather than field-level characteristics may be highly correlated with localness, and thus drive the results. The clearest example here is firm size – oil theft gangs may target the assets of deep-pocketed larger firms because sabotage threats are more likely to generate direct payments, or because larger firms can finance operations under difficult conditions for longer. At the same time, we know that multinationals are much larger, on average, than local firms. The local advantage may therefore be a “small firm effect" rather than a local effect. This is inherently difficult to test, since there is a high degree of correlation between overall firm size and indigeneity. As such, in Table A16 I control for firm size using measures that capture the size of multinational *subsidiaries*. I calculate the (log of) total fields or wells operated by the operating firm of field i at time t. This is a time-varying, field-specific characteristic and so is not absorbed by fixed effects.

Table A16—Divestment and oil field outcomes: firm size

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Output						
Local firm		0.941***	0.890***		0.858**	0.832**
		(0.344)	(0.337)		(0.355)	(0.351)
Log number of wells (firm)	-0.033	0.125	-0.102			
	(0.193)	(0.209)	(0.149)			
Log number of fields (firm)				-0.224	-0.040	-0.226
				(0.269)	(0.293)	(0.193)
Observations	2476	2476	2476	2476	2476	2476
R^2	0.861	0.861	0.878	0.861	0.861	0.878
Panel B: Oil theft						
Local firm		-5.160***	-6.256***		-5.303***	-6.182***
		(0.863)	(1.086)		(0.877)	(1.083)
Log number of wells (firm)	-0.190	-0.869***	-1.275***			
	(0.229)	(0.286)	(0.412)			
Log number of fields (firm)				-0.246	-1.252***	-1.221***
				(0.310)	(0.419)	(0.416)
Observations	3183	3183	3183	3183	3183	3183
R^2	0.710	0.713	0.756	0.710	0.713	0.756
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	No	Yes	No	No	Yes

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Main controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital. Firm size at the field-level is measured as the log of the total number of fields or wells owned by the operating firm. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A uses output as the outcome while Panel B looks at oil theft. Columns (1)-(3) use the log number of wells, while columns (4)-(6) use the log number of fields, which is a more aggregate measure. The main results all hold. Furthermore, columns (2)-(3) and (5)-(6) of Panel B suggest that, after conditioning on local status, the assets of larger firms actually appear to be targeted relatively *less*, an association which is significant at 1%. This may be because, all else equal, larger firms have more capital to invest in security for their assets, either directly or through protection rackets.

A central threat to the assumption of parallel trends is the existence simultaneous shocks

Table A17—Divestment and oil field outcomes: oil prices

Outcome	Output		Oil spills		Oil theft	
	(1)	(2)	(3)	(4)	(5)	(6)
Local firm	0.696*	0.583*	1.047	0.287	-5.526***	-6.793***
	(0.360)	(0.333)	(0.824)	(0.890)	(0.891)	(1.173)
Treated × Oil price (USD/barrel)	-0.015	-0.028***	-0.031**	-0.044***	-0.068**	-0.097***
	(0.010)	(0.011)	(0.013)	(0.013)	(0.029)	(0.032)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	Yes	No	Yes	No	Yes
Observations	2476	2476	3183	3183	3183	3183
R^2	0.862	0.879	0.590	0.651	0.713	0.757

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Main controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. Oil prices are measured as the annual average world crude oil price in dollars per barrel. Field controls are number of wells, initial year, onshore dummy, and maximum well depth. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

that have differential effects on local and multinational firms. The most obvious of these is oil price shocks, which may have outsized impacts on the production of more capital-constrained local firms. I test robustness of the main results to differential oil price effects in Table A17 by including the interaction between the time-invariant localization treatment indicator and the time-varying oil price series p_t. I find no evidence that differential responses to oil price shocks among localized fields are driving the results.

Rexer and Hvinden (2022) show that the 2009 amnesty for Niger Delta militants reduced violence and increased oil theft differentially in conflict-affected regions. If multinationals divested of onshore oilfields in militant-controlled areas during and after the conflict period, then it may be the case that the amnesty policy is contaminating our estimate of the effect of localization on violence and theft. I test robustness to this concern in Table A18. Columns (1), (3), and (5) interact the distance to the nearest militant camp with an indicator for the post-2009 period. Columns (2), (4), and (6) include distance to the nearest militant camp in the standard set of interacted controls, allowing for flexible differential trends by exposure to militant control. The results are unaffected.

C7. Inference

Some divestments occur not necessarily at the field but rather the oil block-level. An oil block is a geographic ownership unit (concession) that typically contains several fields.[49] As such, treatment status may be correlated across fields within blocks, though not perfectly so. More broadly, assets may be correlated in their outcomes and treatment status across space, or within divesting firms. Within-block, firm, or spatial correlation may bias our standard errors downward when clustered at the field level. Table A19 estimates p-values for various different methods of inference: block-level clustering, municipality (LGA) clustering, LGA clustering with a wild-cluster bootstrap, two-way LGA and MNC clustering (standard and wild-cluster), and Conley (2010) spatial standard errors at radii from 25 to 100km. The output

[49] This is a substantially higher level of clustering – there are 314 fields in our data, and only 80 oil blocks.

Table A18—Divestment and oil field outcomes: amnesty policy

Outcome	Output		Oil spills		Oil theft	
	(1)	(2)	(3)	(4)	(5)	(6)
Local firm	0.951***	0.925***	0.938	0.969	-5.196***	-5.133***
	(0.300)	(0.293)	(0.804)	(0.798)	(1.026)	(1.091)
Distance to nearest militant camp (km) × Post-amnesty	-0.003		-0.073**		-0.246***	
	(0.012)		(0.032)		(0.044)	
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2476	2476	3183	3183	3183	3183
R^2	0.878	0.880	0.651	0.665	0.761	0.776

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Main controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. Distance to nearest militant camp is also included in the set of controls for columns (2), (4), and (6). Amnesty date is 2009. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

and oil theft effects remain significant at the 5% under all specifications except two-way LGA and MNC clustering, while the effect on armed conflict is only significant under field-level clustering and 25km spatial errors. The oil spill effects are only significant at the 10% level without controls in field clustering and 25km spatial errors.

C8. Main estimates by distance

The main theft outcome variable is defined as the number of theft incidents within 15km of the field centroid. This relatively wide radius accounts for the fact that oilfields can be large, and even beyond the boundaries of the field there may be significant pipeline infrastructure vulnerable to theft. However, using a wide radius also introduces estimation challenges. First, this may induce mechanical spatial correlation between nearby fields, biasing inference. Second, we may be capturing theft on infrastructure not owned/operated by the firm. If this measurement error is systematically related to localization, it risks biasing the results.

An alternative is to use data on oilfields polygons to define outcomes. This is restrictive because i) these polygons are small (on average 15km^2 in area, or a radius of 2.18) and therefore don't capture theft on outlying pipelines, and ii) because the data is only available for 258 of 314 fields. However, this constitutes a more stringent robustness test, requiring that outcomes are affected in the immediate field area. The results are in Table A20, Panel A. In general, the results for both oil theft and oil spills remain significant, and the patterns of onshore and offshore heterogeneity are evident. In Panel B, I expand the sample back to all 314 oilfields, but restrict the outcome radius to 2km, to generate circular fields with the same average area as the polygons. Panels C and D further extend to 5 and 10 km, respectively.

Since the outcome variables scales change depending on the radius, we have to compare coefficient magnitudes relative to the control group outcome mean. In the main results of Table 1, divestment is associated with a 45% reduction in oil theft from the mean. In Table A20, this effect ranges is 59.9%, 35%, 41% and 46.6% for the polygon, 2km, 5km, and 10km estimates, respectively. The chosen 15km estimate reflects the median effect size across radii, and is not cherry-picked. For the oil spills outcome, the main estimate represents an 18% gain in spills. In contrast, the robustness tests show 35.5, 34.1, -28.5, and 1.8% changes in oil spills, with the latter two (5 and 10km) not significant. Again, the chosen estimate is the median of

Table A19—Divestment and oil field outcomes: methods of inference

Outcome	Output		Oil spills		Oil theft		Conflict deaths	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local firm	0.878	0.944	1.380	0.788	-4.805	-5.703	-3.043	-3.203
Observations	2476	2476	3183	3183	3183	3183	3183	3183
R^2	0.861	0.878	0.590	0.649	0.712	0.756	0.232	0.317
Clustering method								
Field	0.005	0.002	0.073	0.343	0.000	0.000	0.004	0.001
Block	0.007	0.002	0.252	0.515	0.001	0.004	0.188	0.107
LGA	0.023	0.003	0.264	0.619	0.000	0.000	0.159	0.097
LGA (Wild boot)	0.033	0.008	0.349	0.755	0.004	0.002	0.229	0.079
LGA and MNC	0.067	0.041	0.145	0.481	0.044	0.006	0.059	0.062
LGA and MNC (Wild boot)	0.203	0.080	0.325	0.690	0.151	0.064	0.211	0.138
Spatial errors								
25 km	0.002	0.001	0.094	0.356	0.000	0.000	0.066	0.028
50 km	0.003	0.001	0.129	0.403	0.000	0.000	0.108	0.056
75 km	0.004	0.002	0.146	0.431	0.000	0.000	0.122	0.072
100 km	0.006	0.002	0.153	0.442	0.000	0.000	0.129	0.081
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	Yes	No	Yes	No	Yes	No	Yes

Table gives estimates (top row) and p-values for different methods of clustering, indicated in sub-table headers. Sample is the panel of 314 oilfields from 2006-2016 for all methods except two-way LGA and MNC, where the sample is 276 never-treated or divested fields. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Main controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

these estimates; however, these are clearly less robust than the oil theft effects.

C9. Divestment data source

As mentioned in Section II, the data on local firm participation at the oilfield level comes from two sources. The first is the administrative records of the NNPC, which records the operating firm of each oilfield-year. The second source is the DrillingInfo data on corporate transactions. This data provides more detail on local firms' stakes in oilfields, and helps fill in the substantial gaps in the NNPC data. However, it does not distinguish between divestments of operatorship or ownership. Throughout the paper I use a conservative approach that leverages all of the information in both of these datasets, defining treatment as all field-years with any local participation in either dataset.

Table A21 investigates the implications of different treatment definitions for the results. Column (1) reprints the main results for reference. Column (2) uses a similar treatment definition

Table A20—Divestment and oil field outcomes: robustness to distance radius

Outcome	Oil theft			Oil spills		
Sample	All	On	Off	All	On	Off
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Polygons						
Local firm	-0.430**	-0.497**	0.009	0.345**	0.304	1.025**
	(0.195)	(0.243)	(0.031)	(0.168)	(0.190)	(0.401)
Control mean	0.717	1.080	0.007	0.973	0.891	1.134
Observations	2662	1805	857	2662	1805	857
R^2	0.664	0.662	0.210	0.520	0.558	0.505
Panel B: 2km radius						
Local firm	-0.184*	-0.212*	0.008	0.269*	0.213	1.031***
	(0.095)	(0.116)	(0.030)	(0.141)	(0.156)	(0.388)
Control mean	0.525	0.740	0.006	0.789	0.692	1.026
Observations	3183	2296	887	3183	2296	887
R^2	0.615	0.610	0.208	0.421	0.469	0.400
Panel C: 5km radius						
Local firm	-0.735***	-0.874**	0.081*	-0.471	-0.638	1.441***
	(0.281)	(0.348)	(0.047)	(0.611)	(0.615)	(0.485)
Control mean	1.768	2.496	0.008	1.655	1.633	1.707
Observations	3183	2296	887	3183	2296	887
R^2	0.629	0.619	0.222	0.537	0.577	0.512
Panel D: 10km radius						
Local firm	-2.488***	-2.873***	0.111**	0.077	-0.118	3.019***
	(0.599)	(0.716)	(0.051)	(0.662)	(0.642)	(1.044)
Control mean	5.334	7.530	0.027	4.196	4.042	4.569
Observations	3183	2296	887	3183	2296	887
R^2	0.730	0.722	0.308	0.599	0.682	0.540
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses are clustered at the field level. Sample in Panels A and B is the panel of 258 oilfields for which we have polygon data from 2006-2016, while in Panels C and D it is all 314 oilfields. Oil theft / spills is the total number of sabotage / malfunction spills within the field polygon boundaries (A and B) or within 2km of the field centroid (C and D). Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A21—Divestment and oil field outcomes: treatment definition

Treatment	Main	DI first	DI only	NNPC only
	(1)	(2)	(3)	(4)
Panel A: Output				
Local firm	0.878***	0.858***	0.844**	2.151***
	(0.312)	(0.309)	(0.327)	(0.610)
Observations	2476	2476	2476	2476
R^2	0.861	0.861	0.861	0.862
Panel B: Oil spills				
Local firm	1.380*	1.395*	1.600**	1.431
	(0.767)	(0.758)	(0.791)	(0.927)
Observations	3183	3183	3183	3183
R^2	0.590	0.590	0.590	0.589
Panel C: Oil theft				
Local firm	-4.805***	-4.726***	-4.370***	-2.972**
	(0.804)	(0.791)	(0.728)	(1.170)
Observations	3183	3183	3183	3183
R^2	0.712	0.712	0.712	0.711
Field FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

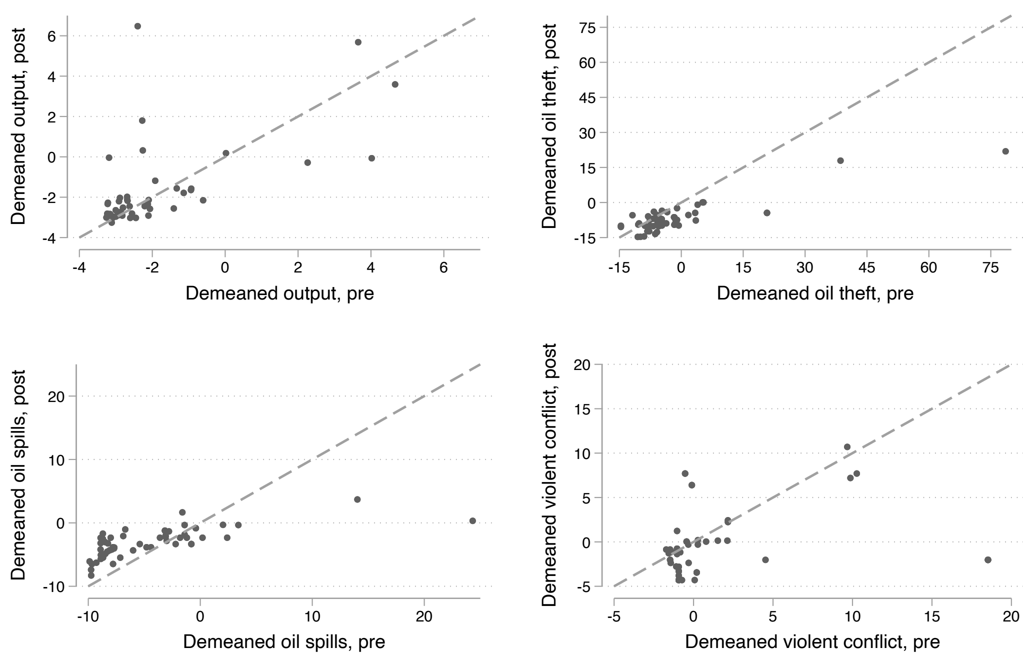
Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Treatment definition is given in table header. Main is the primary treatment indicator used throughout the paper, which defines the localization event year as the first year of treatment in either dataset. DI first takes both DI and NNPC treatments, but uses the DI event year if a field is treated in both datasets. DI only uses only localizations that occur in the DI data. NNPC only uses only localizations that occur in the NNPC data. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.s

as the main specification, leveraging information from both sources. However, when the two sources disagree on the event-year of a given divestment, it takes the DrillingInfo date, since companies may be updated in the NNPC data with a lag whereas DI contains precise dates. This changes treatment status for only a handful of field-years and so does not materially affect any of the results. Column (3) ignores all NNPC data and uses only divestments mentioned in DI from 2006-2016, which yields 56 ever-treated fields. The results remain similar, though the local output advantage falls slightly and the quality disadvantage rises. Finally, column (4) uses only changes in operatorship that are identified in the NNPC administrative data, which gives 70 ever-treated. The results here are still qualitatively similar. However, the output advantage is over twice as large as in (1), while oil theft reduction falls slightly. Both remain significant at 1%.

C10. Outliers

Figure A8 shows that all of the main outcome variables are long-tailed and right-skewed. This suggests that outliers could be driving the main results. I assess the role of outliers as follows. I begin with the stacked event-wise dataset from Section D.D2. For each stack (treated cohort), I calculate the average value of the outcome variable for all untreated (clean control) units in that stack annually, and subtract it from the annual outcomes of the treated units to form the first difference. I then take the average of this difference across years for the pre and post periods for each treated observation. The resulting two quantities are, for each treated field i, $y_{i,pre} = \frac{1}{T_{pre}} \sum_{t<0} y_{it} - \bar{y}_t$ and $y_{i,post} = \frac{1}{T_{post}} \sum_{t \geq 0} y_{it} - \bar{y}_t$, where t is measured in event-time, 0 is the divestment year, and $\bar{y}_t$ is the average of y in the control group at event-time t. In Figure A10, I plot $y_{i,post}$ against $y_{i,pre}$ for the four major outcomes, with the 45 degree line overlaid to indicate no change. The objective is to identify whether there are observations far from the 45-degree line that exhibit a large pre-post deviation, and may therefore be driving the result. For outcomes with a positive effect – output and oil spills, most of the treated fields lie above the 45 degree line. Similarly, for the outcomes with a negative estimate – theft and conflict – the observations generally lie below the line.

Figure A10. Outlier plot



Note: Figure plots $y_{i,post}$ against $y_{i,pre}$ for all divested assets across all four outcomes of the study – output, oil spills, oil theft, and conflict deaths. $y_{i,pre} = \frac{1}{T_{pre}} \sum_{t<0} y_{it} - \bar{y}_t$ and $y_{i,post} = \frac{1}{T_{post}} \sum_{t \geq 0} y_{it} - \bar{y}_t$, where t is measured in event-time, 0 is the divestment year, and $\bar{y}_t$ is the average of y in the control group at event-time t. All estimates come from the stacked dataset with only clean (never-treated) control units and an event-window of five years before and after a divestment event. Dashed line indicates the 45-degree line.

The fact that most observations are of the correct sign is reassuring; however, outlier observations are immediately observable for all of the outcomes. To ensure that these outliers are not driving the results, I calculate for each i the absolute difference $|y_{i,post} - y_{i,pre}|$ and then drop the observations that exceed the 95th percentile of this distribution across treated assets, separately for each outcome. I then re-estimate the results without these outlier treated observations, presented in Table A22. The main results are not meaningfully changed.

Table A22—Divestment and oil field outcomes: dropping outliers

Outcome	Output		Oil theft		Oil spills		Conflict	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local firm	0.637***	0.671***	-3.988***	-4.942***	1.905***	1.335**	-0.797	-1.191**
	(0.226)	(0.247)	(0.676)	(0.936)	(0.512)	(0.604)	(0.505)	(0.589)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2447	2447	3150	3150	3150	3150	3128	3128
R^2	0.865	0.881	0.713	0.754	0.593	0.651	0.273	0.332

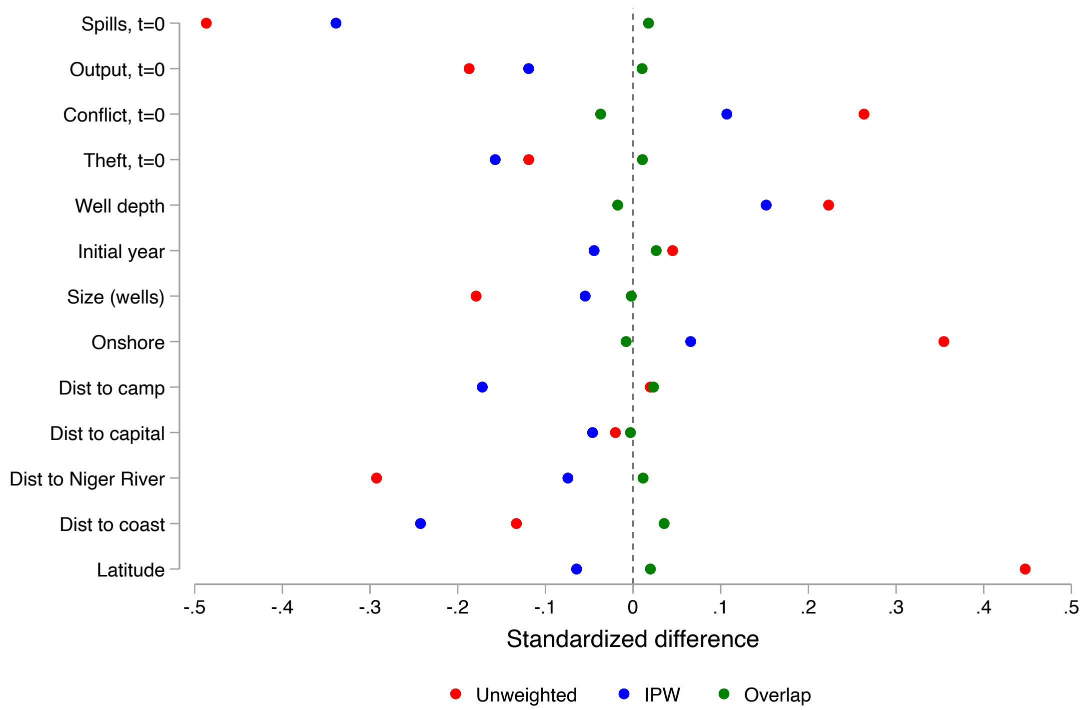
Standard errors in parentheses are clustered at the field level. Sample is all controls, and all treated fields for which $|y_{i,post} - y_{i,pre}|$ is below the 95th percentile for a given outcome. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. Main controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C11. *Propensity score weighting*

The summary statistics in Table A4 suggest some meaningful differences between treated and never-treated assets, and in particular between the divested and controls. Of course, these fixed differences are accounted for in the field fixed effects. Furthermore, throughout the paper and particularly in Appendix C.C6, I demonstrate extensive robustness to the inclusion of various linearly interacted controls.

An alternative nonparametric method of ensuring balance between treated and control units is propensity score re-weighting. I calculate a propensity score $p(\hat{\beta}, X_i)$, which gives the predicted probability that observation i is treated (divested), T_i, given the field covariates X_i, estimated via probit regression. I consider 3 different specifications for X_i: i) the parsimonious set of geographic control variables used throughout the paper, ii) these baseline controls plus field characteristics (as in Table A14), and iii) baseline controls, field characteristics, and initial conditions of the four outcome variables. I choose two re-weighting schemes: trimmed inverse propensity weighting (Crump et al., 2009) and overlap weights (Li and Thomas, 2018). In the former $w_i^{IPW} = T_i \frac{1}{p(\hat{\beta}, X_i)} + (1 - T_i) \frac{1}{1 - p(\hat{\beta}, X_i)}$, while in the latter $w_i^O = T_i(1 - p(\hat{\beta}, X_i)) + (1 - T_i) p(\hat{\beta}, X_i)$. In both cases, control units are up-weighted if they have a higher likelihood of treatment given their covariates (that is, if they "look" like treatment observations along X_i). Figure A11 estimates the average difference between treatment and control across fixed covariates and initial conditions in the unweighted sample, and with reweighting. Both re-weighting schemes substantially improve covariate balance, although overlap outperforms IPW. This is partially by construction, as overlap weights can be shown to achieve exact balance of covariates with a logistic propensity score (Li and Thomas, 2018).

Figure A11. Reweighting



Note: Figure shows differences in average characteristics between divested (switcher) and pure control (never-treated) assets under different weighting schemes. $t = 0$ variables are measured as the annual average of the outcome variable for all $t \leq 0$ in event time. Event window is ten years before and after divestment.

The results of the re-weighted estimation, using the stacked data setup, are in Table A23. Standard errors are estimated with a clustered bootstrap routine with 100 replications. Panel A reprints the unweighted stacked estimates for comparison. All of the estimates retain their original signs, and the vast majority remain significant. The primary exceptions are the more stringent specifications using overlap weights, where the estimates for output, oil spills, and conflict deaths are no longer significant.

Table A23—Divestment and oil field outcomes: propensity score re-weighting

Outcome	Output	Oil spills	Oil theft	Conflict
	(1)	(2)	(3)	(4)
Panel A: No weights				
Local firm	1.013***	1.632**	-5.001***	-3.332***
	(0.327)	(0.778)	(0.827)	(1.055)
Panel B: Inverse-propensity weights				
Base controls	0.754***	1.776**	-5.691***	-2.686***
	(0.255)	(0.794)	(0.884)	(0.787)
Base + field controls	0.627**	2.484***	-6.510***	-3.230***
	(0.259)	(0.545)	(1.161)	(0.967)
Base + field + initial conditions	0.694***	1.136	-6.237***	-2.892***
	(0.264)	(1.371)	(0.896)	(0.710)
Panel C: Overlap weights				
Base controls	0.941***	2.144***	-6.126***	-2.912***
	(0.318)	(0.722)	(0.903)	(0.899)
Base + field controls	0.796**	2.292***	-5.632***	-1.690
	(0.378)	(0.650)	(1.021)	(1.052)
Base + field + initial conditions	0.521	0.163	-4.879***	-0.944
	(0.352)	(0.497)	(1.165)	(0.816)
Field-by-Cohort FE	Yes	Yes	Yes	Yes
Year-by-Cohort FE	Yes	Yes	Yes	Yes

Standard errors in parentheses are clustered at the field level and bootstrapped with 100 replications. All estimates use the stacked event-wise dataset. Weighting scheme given in panel header. Base controls are latitude of the field centroid, distance to Niger River, and distance to the capital. Field controls are distance to the enarest militant camp, an onshore indicator, the number of wells, field age, and average well depth. Initial conditions are annual average oil theft, conflict, output, and oil spills averaged for years $t \leq 0$ in event-time.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

ROBUSTNESS TESTS: TWO-WAY FIXED EFFECTS

The main estimates of divestment effects use TWFE estimation methods. A growing literature has identified the fundamental challenges in interpreting TWFE estimates as average treatment effects (ATEs) in settings with staggered treatment timing and treatment effect heterogeneity across cohorts or over time. The central issue is that the TWFE estimate is a weighted average of 2x2 difference-in-difference comparisons across different treatment and control groups (Goodman-Bacon, 2021). Unfortunately, some of these 2x2 comparisons are quite bad in the presence of dynamic effects; for example, always-treated observations act as controls even though their previous treatment status should alter their trends. Similar logic applies to comparing units treated later to those earlier. These bad control have been alternatively expressed as negative weights on some unit-time-specific heterogeneous ATEs (de Chaisemartin and D'Haultfoeuille, 2020). Baker, Larcker and Wang (2021) show that these issues have real empirical implications, and can lead TWFE estimates to be quantitatively misleading and even wrong-signed relative to underlying ATEs. Fortunately, several diagnostic decomposition tools and alternative estimators can help address this problem (Goodman-Bacon (2021), Callaway and Sant'Anna (2021)).

D1. TWFE decomposition

As an initial diagnostic tool, Goodman-Bacon (2021) shows that the staggered-adoption TWFE estimate can be decomposed into a weighted average of all 2×2 difference-in-difference comparisons. These weights depend on the size of the groups and the variance of the treatment in each 2×2 comparison; TWFE will tend to place lower weight on 2×2 estimates for units treated early or late in the panel. The key insight is that these weights identify which comparisons are driving the overall TWFE results. Table A24 presents weights and average treatment effect estimates for each 2×2 DD comparison type.[50] Panel A presents the results for output, while Panel B decomposes the oil theft effect.

Because of the large sample of never-treated clusters, the TWFE estimate heavily weights the "treated vs. never treated" 2×2 comparison, which accounts for 85% of the treatment effect. In Panel B, every 2×2 comparison estimate is negative, ranging from -1.1 to -11.2. In Panel A, every 2×2 group estimate is positive and of comparable magnitude, except the "Later treated (T) vs. Earlier treated (C)" comparison, which uses earlier treated units as controls for fields that switch into treatment towards the end of the panel. Note, however, that this is a so-called "forbidden" comparison, since it uses already-treated units as controls, despite the fact that their treatment status has already affected their subsequent trend. To adjust for this, I report the "purged estimate", which removes the "Later treated (T) vs. Earlier treated (C)" and "Treated (T) vs. Already treated (C)" comparisons, both of which rely on already-treated fields to serve as controls, and reweights the estimate accordingly. Both purged estimates are even larger than the main estimate, suggesting that TWFE biases our estimates toward zero.

Table A24 also decomposes treatment effects by identifying variation, allowing us to probe the identification assumption. In Section IV, I argue that treatment timing may be more plausibly exogenous than treatment assignment. The comparison that leverages only timing variation among ever-treated units is "Earlier treated (T) vs. Later treated (C)," while the comparison that relies on variation between treatment and never-treated fields is "Treated (T) vs.

[50] The estimation is run on a subsample of 275 fields for which a balanced panel is available.

Table A24—Goodman-Bacon (2021) TWFE weights

Comparison	Weight	Estimate
Panel A: Output		
Earlier treated (T) vs. Later treated (C)	0.051	0.811
Later treated (T) vs. Earlier treated (C)	0.022	-0.854
Treated (T) vs. Never treated (C)	0.848	0.844
Treated (T) vs. Already treated (C)	0.079	0.572
TWFE estimate		0.783
Purged estimate		0.842
Panel B: Oil theft		
Earlier treated (T) vs. Later treated (C)	0.051	-11.202
Later treated (T) vs. Earlier treated (C)	0.022	-8.658
Treated (T) vs. Never treated (C)	0.848	-4.684
Treated (T) vs. Already treated (C)	0.079	-1.147
TWFE estimate		-4.824
Purged estimate		-5.054

Sample is the subset of 275 fields for which a balanced panel is available ($N = 3025$). Outcome variable in panel header. Output is measured in millions of barrels of oil per year, using GLM to impute missing output. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. All models estimate two-way fixed effects weights and ATEs for different 2x2 comparison groups using the method explained in Goodman-Bacon (2021). T and C in parentheses indicates which observations are used as treatment and which as control, respectively, for a given comparison. Purged estimate refers to the weighted ATE which removes Treated (T) vs. Already treated (C) and Later (T) vs. Earlier treated (C) comparisons. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Never treated (C).” Despite the fact that the latter drives much of our observed effect, the signs and magnitudes of the former are similar. Relying only on exogenous treatment timing does not weaken the results.

D2. Stacked DD

Moving beyond diagnostics, several recent papers propose estimators for addressing the issues in TWFE. In general, these methods amount to different ways of removing already-treated controls from the estimation. One alternative estimation method is the stacked DD,[51] as suggested by Goodman-Bacon (2021). In this method, treated units in each treatment-year cohort are paired with all not-yet-treated observations in the data as of year t. The cohorts are then “stacked” to obtain a dataset in which the control groups are always untreated, and event-time takes the place of calendar year. This ensures that already-treated observations are never used as controls. We then estimate the following equation, for unit i in cohort-stack c for event-time t

[51] See Gormley and Matsa 2011, Deshpande and Li 2019, and Baker, Larcker and Wang 2021 for examples.

$$y_{ict} = \alpha + \beta local_{ict} + \delta_{ct} + \gamma_{ic} + \epsilon_{ict}$$

Standard errors are clustered at the field level. The parameter β is a variance weighted average of cohort-specific treatment effects, where each cohort-specific comparison is only between newly treated and not-yet-treated groups. An additional robustness test is to further restrict the sample either to ever-treated or never-treated fields, in order to isolate the role of treatment timing, as in the Goodman-Bacon (2021) decomposition.

Table A25—Divestment and oil field outcomes: stacked-DD estimation

	(1)	(2)	(3)
Panel A: Output			
Local firm	1.013***	1.010***	1.157**
	(0.327)	(0.336)	(0.536)
Observations	20637	18695	2506
R^2	0.872	0.873	0.750
Panel B: Oil spills			
Local firm	1.632**	1.619**	2.859**
	(0.778)	(0.816)	(1.182)
Observations	26229	23675	3338
R^2	0.584	0.584	0.622
Panel C: Oil theft			
Local firm	-5.001***	-4.718***	-10.885***
	(0.827)	(0.784)	(2.514)
Observations	26229	23675	3338
R^2	0.715	0.716	0.729
Field-by-Cohort FE	Yes	Yes	Yes
Year-by-Cohort FE	Yes	Yes	Yes
Control group	All	Untreated	Treated

Standard errors in parentheses are clustered at the field level. Outcome variable in panel header. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. All models use the stacked difference-in-differences estimation method explained in Baker, Larcker and Wang (2021). All models use a symmetric event window of +/- 10 years. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The results of this analysis are given in Table A25 for the three main outcomes. Column (1) uses all possible control units, while column (2) uses only never-treated and column (3) uses only ever-treated. I find that full-sample stacked-DD estimates (columns 1-3) are robustly negative and significant for oil theft, and positive and significant for output and oil spills. The magnitude of effects is in fact somewhat larger than the TWFE estimates in Table 1. The results indicate that using already-treated units as control is not a substantial source of bias in our main TWFE estimates, consistent with their low weights in Table A24. If anything, TWFE biases our main results toward zero.

D3. CSDID

Callaway and Sant'Anna (2021) propose a semi-parametric DD estimator to address the "negative weights" problem, which also corrects for the down-weighting of early and late-treated groups in the presence of cohort-specific heterogeneity. The estimator computes propensity-score-weighted ATT effects for each cohort-period, and then aggregates these estimates across various dimensions (cohort, time, or both). It is similar in spirit to the stacked model in that it emphasizes cohort-specific variation and uses only the untreated as controls. However, it does not rely on a linear parametric specification, and allows for more flexible re-weighting in the aggregation of cohort-and-time-specific ATT parameters.

Table A26—Divestment and oil field outcomes: Callaway and Sant'Anna (2021) estimation

Outcome	Output		Oil spills		Oil theft	
	(1)	(2)	(3)	(4)	(5)	(6)
Local firm	0.570*	0.561*	1.689***	1.623***	-4.169***	-4.366***
	(0.303)	(0.303)	(0.468)	(0.466)	(1.119)	(1.195)
Observations	2246	2246	2895	2895	2895	2895

Standard errors in parentheses are clustered at the field level. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Oil theft is the total number of sabotage spills within 15 km of the field. All models use the difference-in-differences estimation method for staggered adoption settings detailed in Callaway and Sant'Anna (2021). Columns (1), (3), and (5) use only never-treated observations as controls. Columns (2), (4), and (6) use both never and not-yet treated as controls. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

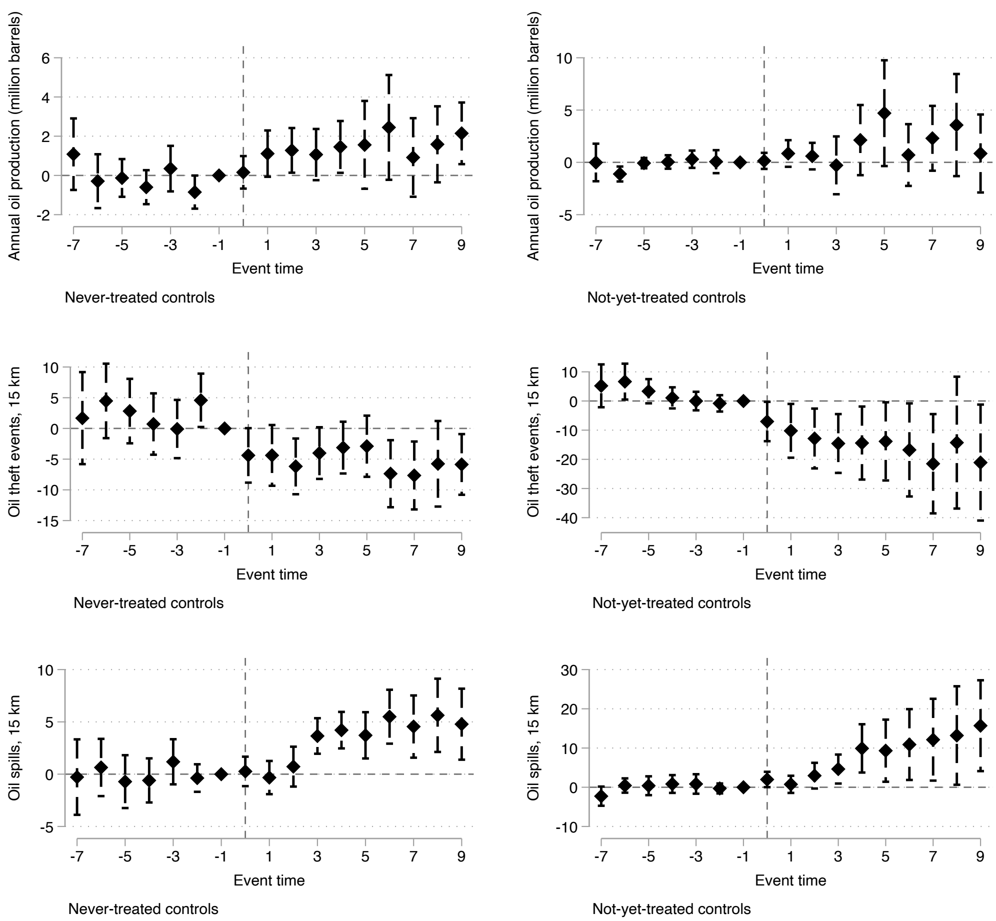
Table A26 provides results using doubly-robust inverse-probability weighting (**?**) for the three main outcomes. Columns (1), (3), and (5) use only never-treated observations as controls. Columns (2), (4), and (6) use both never and not-yet treated as controls. None of the specifications include control variables. All results are directionally robust and statistically significant at the 10% level or lower. The output effects are smaller than in Table 1 and only significant at the 10% level, while the oil spill effects are larger and more significant. Note, however, that the estimation routine drops fields that are not "pair-balanced", that is, observed in both $t = 0$ and $t = 1$ of event-time. This smaller sample may explain slightly different results and loss of significance.

D4. Event-study plots: main effects

The main event-study plot in Figure 3 employs the stacked DD configuration for the three main outcomes, using all yet-untreated fields as controls, and controlling for interacted covariates. In this section, I consider event-studies using different estimators, samples, control variables, and outcome variables, in order to verify that the parallel trends obtained in Figure 3 are robust.

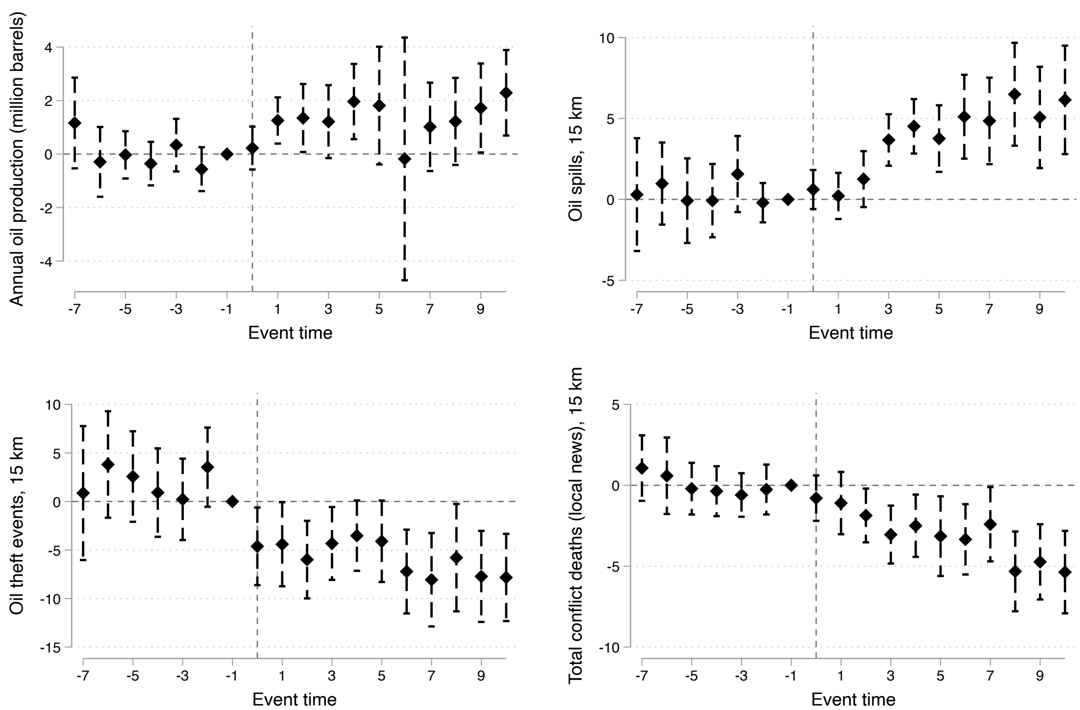
Figure A12 decomposes the stacked DD event-study by control group following the method in Table A25 for the three main outcomes of output, theft, and oil spills. The left panel uses only never-treated controls, while the right panel uses only ever-treated controls. There is evidence for parallel trends in both comparisons. Figure A13 uses TWFE for the event-study specification, for the three main outcomes as well as conflict deaths. All regressions include the main set of spatial controls used throughout the paper. The results are visually very similar to the main event-study, with the exception of a noisy zero coefficient estimated for year 6 after treatment for the output outcome. Figure A14 uses the Callaway and Sant'Anna (2021) estimates aggregated by event-year to generate event-study plots. Given the smaller sample, the results appear somewhat noisier than stacked and TWFE specifications, but the patterns are remarkably consistent. The only major difference is that the post-event coefficients for oil theft converge back to zero in years past five, suggesting smaller long-run effects on theft than the other specifications. Lastly, Figure A15 estimates the TWFE event study corresponding to Table 3, columns (3) and (6), which look at terminated divestments. In this case, the date in which a terminated divestment was initiated is taken as the event-date, yielding 36 treated fields. As expected, the plots indicate no measurable pre-trends and no post-treatment effects.

Figure A12. Stacked event-study: main outcomes by control group



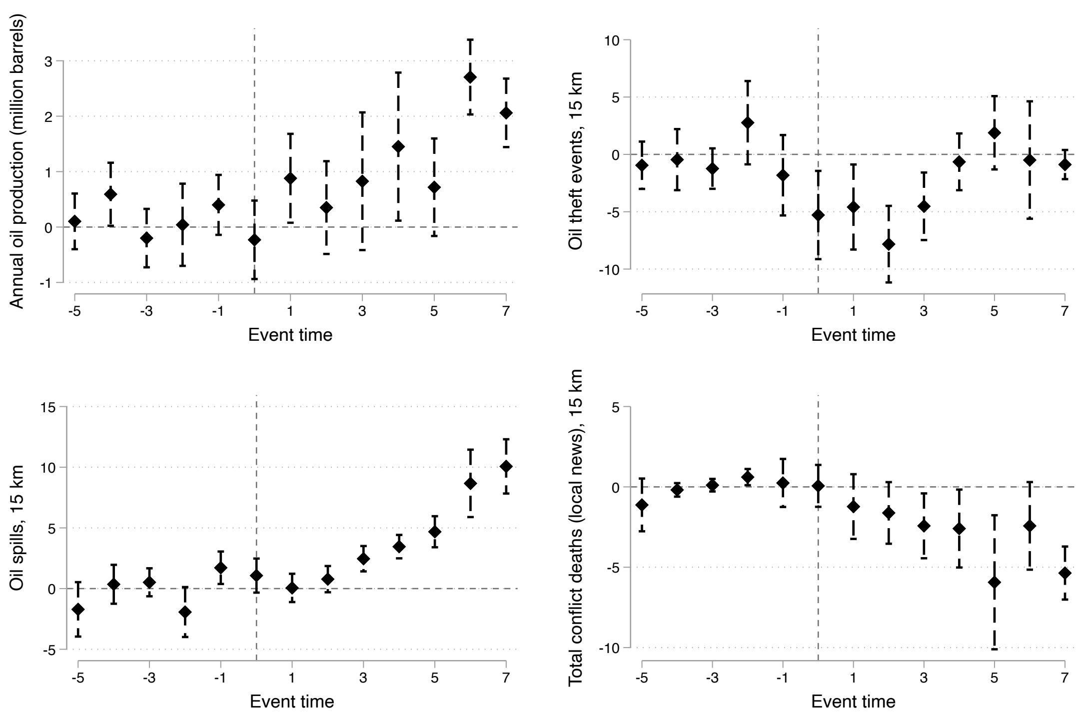
Note: Figure shows coefficients from stacked event-study regressions for oil production (top panel), oil spills (middle panel), and oil theft (bottom panel). Standard errors are clustered at the field-level. Left panel uses only never-treated controls, while right panel uses only ever-treated controls. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Theft is the total number of sabotage spills within 15 km of the field. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km.

Figure A13. TWFE event-study: main outcomes



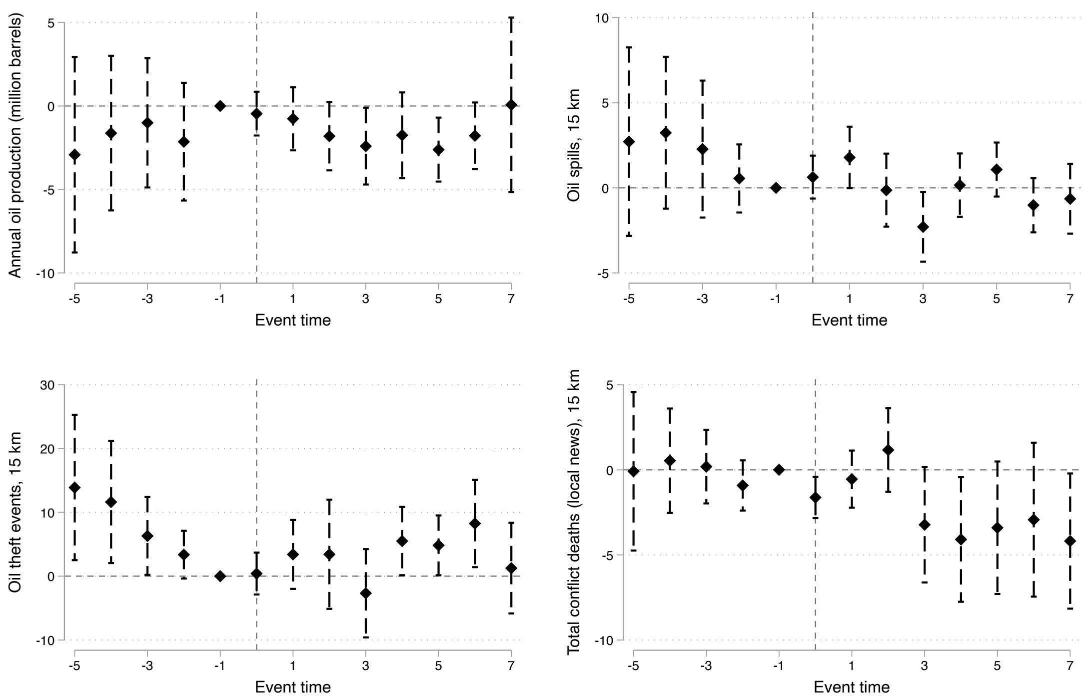
Note: Figure shows coefficients from TWFE event-study regressions for oil production, oil theft, and oil spills. Standard errors are clustered at the field-level. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Theft is the total number of sabotage spills within 15 km of the field. Conflict is the total number of conflict deaths within 15 km of the field reported by local news media.

Figure A14. Callaway and Sant'Anna (2021) event-study: main outcomes



Note: Figure shows coefficients from event-study estimation following Callaway and Sant'Anna (2021) for oil production, oil theft, and oil spills. Standard errors are clustered at the field-level. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Theft is the total number of sabotage spills within 15 km of the field. Conflict is the total number of conflict deaths within 15 km of the field reported by local news media.

Figure A15. Placebo event-study: main outcomes



Note: Figure shows coefficients from TWFE event-study estimation for oil production, oil theft, and oil spills. Event time is defined relative to the first year of a terminated divestment. Standard errors are clustered at the field-level. Output is measured in millions of barrels of oil per year. Oil spills are the total number of non-sabotage spills within 15 km of the field. Theft is the total number of sabotage spills within 15 km of the field. Conflict is the total number of conflict deaths within 15 km of the field reported by local news media.

Robustness tests: Other outcomes

E1. Robustness: conflict measurement

The results of Table A12 columns (1)-(2) may be biased by measurement error if multinational firms are considered more newsworthy, since ACLED data is derived from media reports. As such, a local takeover may reduce the media attention to a given field, rather than the underlying level of violent conflict. Table A27 subsets conflict events by news media source. Column (1) re-prints the main results, while column (2) includes only conflict events reported by local Nigerian news media. Local reports are less susceptible to media bias because indigenous Nigerian firms are likely to be newsworthy to a local audience. Column (3) shows the results for only internationally-reported events; they all remain statistically significant at the 1% level. Column (4) subsets conflict events to only those that include organized militant groups, while column (5) further restricts to militant events targeting the oil sector. The reduction in conflict is robust to restricting to explicitly oil-related organized violence.

Table A27—Divestment and conflict: robustness to measurement error

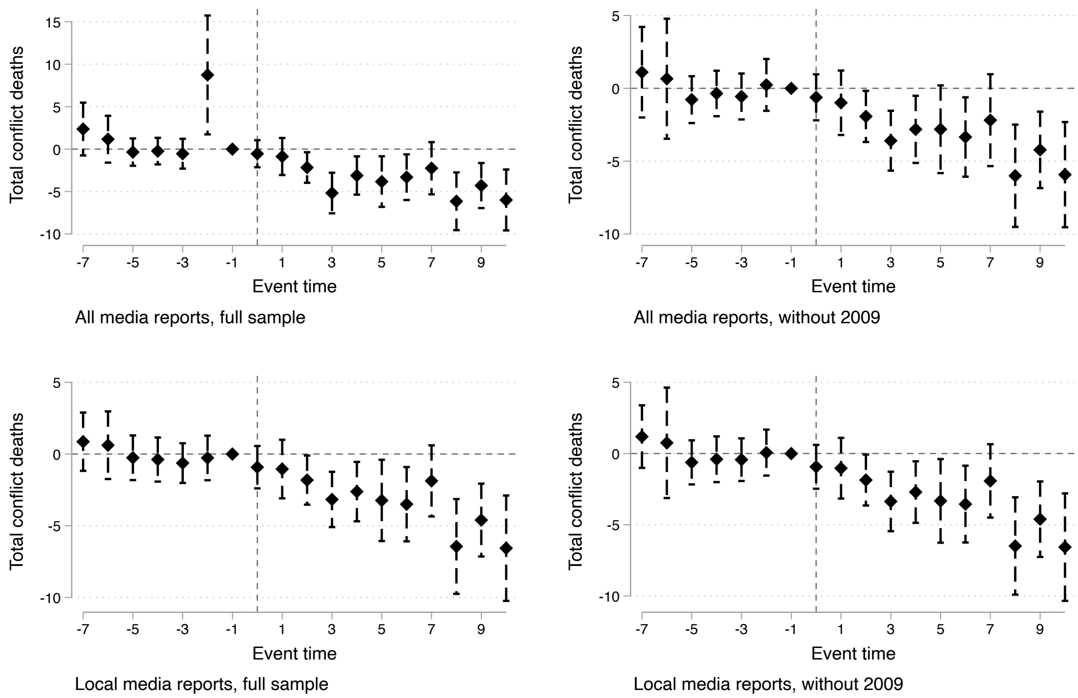
Outcome	All	Local	Int'l	Militant	Oil mil
	(1)	(2)	(3)	(4)	(5)
Local firm	-3.203***	-1.231***	-1.972**	-3.138***	-1.522*
	(0.993)	(0.461)	(0.902)	(0.953)	(0.897)
Field FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls $\times$ Year FE	Yes	Yes	Yes	Yes	Yes
Control group mean	2.006	1.099	0.907	1.720	0.621
Observations	3183	3183	3183	3183	3183
R^2	0.317	0.348	0.261	0.292	0.221

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Conflict deaths in (1) are the total number of conflict-related fatalities reported in news media within 15 km of the field. Columns (2) and (3) subset conflict events to only those reported by international and local news media, respectively. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure A16 assesses the role conflict measurement error in the stacked event-study framework. The top-left panel includes all conflict events and all years of data. While the trends are broadly parallel, there is a spike in conflict at event-time $t = -2$. This outlier coefficient is likely to be driven by fields that were divested in 2011, such that $t = -2$ corresponds to 2009, the year of the culmination of the Niger Delta conflict, which witnessed an unprecedented spike in violent conflict (Rexer and Hvinden, 2022). To account for this, I drop the year 2009 from the estimation in the top-right panel. Pre-trends become flat and insignificant.

The Niger Delta conflict primarily targeted multinational firms, and was highly publicized in international media. It is therefore possible that the 2009 spike in conflict is an artefact of the data, driven by over-reporting of the conflict among international news sources. The bottom panel of the figure uses only local news media reports, as in Table A27 column (2). The pre-divestment spike in conflict disappears, regardless of whether 2009 is included (left panel)

Figure A16. Stacked event-study: conflict deaths by measurement approach



Note: Figure shows coefficients from stacked event-study regressions for total conflict deaths within 15 km of the field. Top panel uses all news media reports in ACLED data. Bottom panel drops conflict events in ACLED reported by international news media sources. Left panel uses the full sample, while right panel drops 2009, the year of the Niger Delta amnesty. Standard errors are clustered at the field-level. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km.

or dropped (right panel). Conflict measurement using international news media is indeed sensitive to outliers, and the restriction to local news media sources may be more appropriate. I therefore use only local media reports in all of the event-study figures of Appendix D.D4.

E2. Robustness: enforcement outcomes

I argue that local law enforcement agents offer preferential protection to the assets of Nigerian firms. This protection is specific to the black market for stolen oil, the primary production risk faced by firms. However, a plausible alternative mechanism is that localization simply coincides with a generalized increase in law enforcement activity. I consider this hypothesis in Table A28, which estimates the impact of divestment on law enforcement actions against non-oil crime. Columns (1)-(2) aggregate all non-oil related criminal activities, while columns (3)-(6) disaggregate this category into two important crimes – kidnapping and gang activity. The results for all non-oil crime are quantitatively small and insignificant. Kidnapping produces somewhat larger positive point estimates, but remains noisy and insignificant. Columns (5)-(6) show, if anything, a reduction in law enforcement actions against gangs. Overall, there is no evidence of a generalized increase in law enforcement activity following divestment.

Table A28—Divestment and non-oil law enforcement activity

Outcome	All non-oil		Kidnapping		Gangs	
	(1)	(2)	(3)	(4)	(5)	(6)
Local firm	0.091	0.007	0.359	0.381	-0.013	-0.066*
	(0.606)	(0.595)	(0.243)	(0.254)	(0.035)	(0.039)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	Yes	No	Yes	No	Yes
Control group mean	1.739		0.568		0.077	
Observations	3183	3183	3183	3183	3183	3183
R^2	0.390	0.501	0.456	0.513	0.208	0.284

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Enforcement category is given in the table header, defined as the total number of enforcement actions reported in news media within 15 km of the field. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Additional results

F1. Political connections

Section VI.B shows that firms' connections to the Nigerian security forces are associated with substantially lower levels of oil theft, increased output, and increased enforcement. In Table A29, I disaggregate by type of political connection for the primary outcome of oil theft. Columns (1) shows that there is no association between political connections and oil theft, conditional on TWFE and interacted controls. Columns (2)-(5) further show that specific connections to technocrats, elected politicians, cabinet-level figures, and traditional leaders are not significantly associated with oil theft, although the point estimate on elected connections is negative and represents 17% of the outcome mean. Column (6), however, shows that connections to the Nigerian security forces are associated with a large reduction in theft, significant at 1% and equivalent to 42.7% of the outcome mean. Only connections to the security forces matter for reducing theft.

Table A29—Political connections on oil theft: type of connection

Connection	Any	Tech.	Elected	Cabinet	Chief	Security
	(1)	(2)	(3)	(4)	(5)	(6)
Connected	0.168	0.070	-1.635	-0.761	-0.190	-4.068***
	(0.811)	(0.782)	(1.033)	(0.905)	(0.954)	(1.492)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3183	3183	3183	3183	3183	3183
R^2	0.753	0.753	0.753	0.753	0.753	0.754

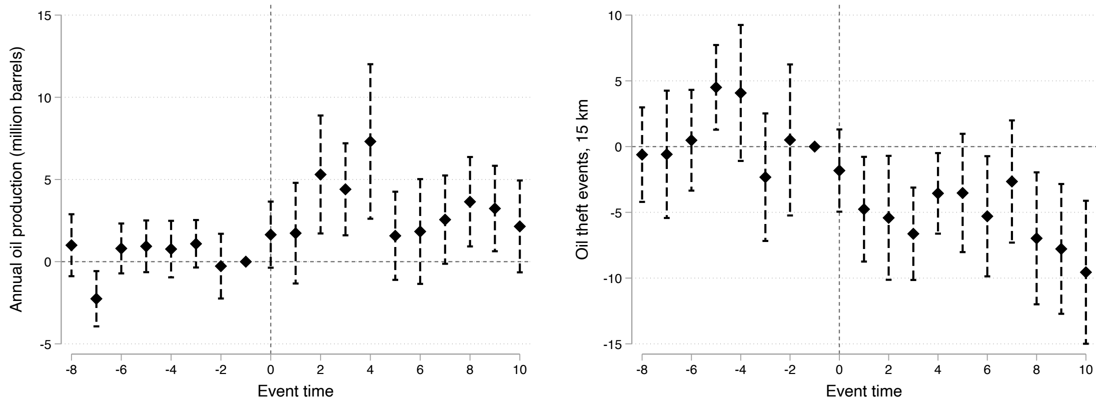
Standard errors, in brackets, are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016 for which political connections data is available. Outcome variable is oil theft, the total number of sabotage spills within 15 km of the field. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. Political connections variables are dummy variables indicating that the operator of a given field-year has a particular type of politician as a board member, shareholder, or manager. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure A17 estimates TWFE event studies, where the event-year is defined as the first year in which an oilfield obtains a boardmember or shareholder connected to the security forces. The results indicate that for both oil output (left panel) and oil theft (right panel), pre-trends are broadly parallel.

F2. Partial ownership

Partial ownership drives a wedge between the losses to the operating firm and criminal profits; operators with larger ownership stakes γ internalize a greater share of the losses from theft, increasing bargaining space. The Nigerian oil market exhibits substantial variation in ownership agreements (see Figure A3), and local operators may have greater ownership stakes for several reasons: *i*) multinational divestment may lead to consolidation of stakes in joint ventures, and *ii*) because of indigenization policies, local firms are more likely

Figure A17. TWFE event-study: security connections



Note: Figure shows coefficients from TWFE event-study regressions for oil production (left panel) and oil theft (right panel). Standard errors are clustered at the field-level. Treatment timing is defined as the year an oilfield obtains its first boardmember or shareholder connected to the security forces. Output is measured in millions of barrels of oil per year. Theft is the total number of sabotage spills within 15 km of the field.

to obtain sole-risk contracts than multinationals, who must provide mandated equity stakes to government. It is therefore plausible that greater ownership stakes allow local firms to more efficiently internalize losses.

Table A30—Divestment and asset ownership shares

Outcome	HHI	Op share	Gov share
	(1)	(2)	(3)
Local firm	0.044***	6.769***	-1.205
	(0.014)	(1.831)	(1.150)
Field FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	3082	3082	3082
R^2	0.984	0.983	0.984

Standard errors, in brackets, are clustered at the block level. Sample is the panel of 84 concession blocks from 2006-2016. Outcome variable is indicated in table header; either the block-level equity HHI, the equity stake of the operating firm, or the equity stake of the NNPC. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

To test whether localization increases consolidation, in Table A30 I re-estimate the main TWFE regression where the outcome variable is the concession equity Herfindahl-Hirschman Index, measured from 0 to 1 (1), the operator's stake (2), or the state's stake (3). Divestment increases the HHI by 0.044 points, a 6.6% increase on the multinational mean, significant at 1%. Local divestment also increases operator ownership by 6.8 p.p., a 10.7% increase, significant at 1%. However, there is no significant change in the NNPC share. The results indicate that divestment substantially increases ownership concentration in the hands local operators. However, this is driven by consolidation of MNC stakes in divestment, rather than by any reduction in the burden of government ownership.

F3. Corruption laws

Multinational firms may face higher expected costs of λ of engaging in corrupt behavior. In general, these costs are driven by home anti-corruption statutes that prohibit multinationals from improper payments to foreign officials, such as the Foreign Corrupt Practices Act (FCPA) in the United States. Given the relatively broad definitions of foreign officials contained in these laws, the prospect of legal liabilities could plausibly deter multinationals from bargaining with law enforcement, even at arms length. If this is the case, we should observe that among multinationals, exposure to these laws should explain variation in levels of theft. By restricting the sample to multinationals, I remain agnostic about the content, quality, and enforcement of Nigeria's own anti-corruption laws.[52]

Every multinational firm in Nigeria's oil sector currently falls under some form of foreign anti-bribery statute. In order to test this hypothesis in a TWFE model, I employ the staggered nature of law passages. The US FCPA was passed in 1977, but the UK Bribery Act, which covers Shell, was only passed in 2010. The Italian statute governing Agip was passed in 2012, the Swiss statute governing Addax (until its sale to SINOPEC in 2009) was passed in 2000, while the French law governing Total was not passed until 2017. Thus, there is variation in the timing of laws governing each oilfield over the sample period.

The results of this estimation are in Table A31 for oil output, theft, and local conflict outcomes. In general, foreign corruption laws have limited effect on the actual production decisions of the firm: the estimate with controls in column (2) is near zero and insignificant. However, in columns (3)-(4), we can see that increased corruption costs do impact the ability of multinational firms to mitigate theft on their assets. The passage of a home-country corruption law is associated with 3.8-7.6-incident increase in theft, significant at the 1% level. A similarly large increase of 0.7-1.3 conflict deaths is shown in columns (5)-(6).

ALTERNATIVE EXPLANATIONS

G1. Spatial spillovers

In a general equilibrium setting, gangs may optimally choose targets for theft across all possible oil fields. As such, divestment could increase targeting of surrounding multinational fields if local fields are politically protected but their multinational neighbors are not. In contrast, if security is non-excludable, increased anti-crime enforcement by security forces may have positive spillovers to nearby multinational firms. In either case, spatial spillovers will bias the treatment effect by violating the stable unit treatment value assumption (SUTVA) (Rubin 2005), since nearby untreated fields experience some impact of treatment.

To test for spatial spillovers, I follow the "ring method" common in the urban economics literature (see e.g. Autor, Palmer and Pathak 2014 and Diamond and McQuade 2019). In the stacked dataset (see Appendix D.D2), for each event date, I identify all untreated fields. For each untreated field, I calculate the distance from that field to the nearest treated field. I then re-estimate the stacked difference-in-differences specification including interactions between

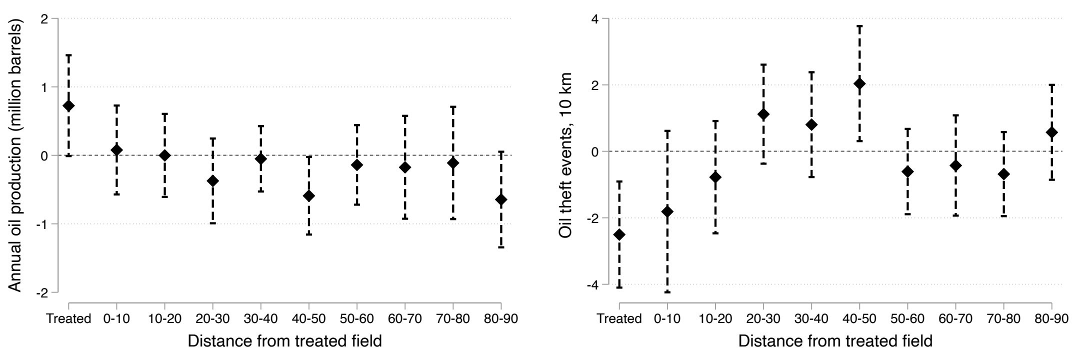
[52] This is preferable to assessing the effectiveness of these laws, which legal analysis suggest are basically ineffective (Aigbovo and Atsegbua 2013).

Table A31—Anti-corruption laws and oil field outcomes

Outcome	Output		Oil theft		Conflict deaths	
	(1)	(2)	(3)	(4)	(5)	(6)
Home-country corruption law	0.695**	0.066	7.593***	3.808***	1.260***	0.744***
	(0.278)	(0.329)	(0.925)	(0.908)	(0.217)	(0.265)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	Yes	No	Yes	No	Yes
Control group mean	2.072		8.826		0.503	
Observations	2111	2111	2679	2679	2679	2679
R^2	0.874	0.888	0.729	0.766	0.277	0.384

Standard errors in parentheses are clustered at the field level. Sample is all untreated field-years from 2006-2016 (i.e., operated by multinational firms). Output is measured in millions of barrels of oil per year. Oil theft is the total number of sabotage spills within 15 km of the field. Conflict deaths is the total number of violent conflict-related fatalities within 15 km of the field as reported by local news media sources. Home country corruption law indicates that a field is operated by a company under the jurisdiction of a foreign anti-corruption statute. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure A18. Spatial spillovers



Note: Figure plots coefficient estimates of treatment effect and spillover effects for output (left panel), and oil theft (right panel), defined as the total number of sabotage spills within 15 km of the field. Estimates are derived from a stacked difference-in-differences regression of the outcome on a dummy for post-treatment interacted with indicators for the treatment and "ring" distances from the nearest treated field. Omitted control group is untreated fields further than 90km from the nearest localized field. All specifications include stack, time, and field fixed effects and their interactions, as well as interacted controls for latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital. Standard errors clustered at the field level.

the post-treatment indicator and dummy variables for treated fields, as well as dummies for control fields within each ten-kilometer interval from 0 to 90. The omitted group of untreated fields greater than 90 kilometers away from a treated field acts as the "pure" control group. I define theft outcomes in a 10 km radius around the field in order minimize overlap which induces a mechanical spatial correlation in outcomes and therefore spurious spillover effects.

The results are in Figure A18, which plots the treatment effect, as well as coefficients at each ring from 0-10 to 80-90 km, for output and oil theft (left and right panels, respectively).[53] In both cases, the main treatment effects remain strong. Furthermore, the output effect does not appear driven by declines on control fields. In fact, there is minimal evidence of substantial spatial spillovers across either outcome for nearby or faraway fields. There *are*, however, statistically significant spillovers in the 40-50 km bin, suggesting a crime displacement effect resulting in less output. This is reasonable, since positive security externalities and increasing costs of transport might limit displacement effects for nearby and faraway fields, respectively. Nonetheless, they do not affect the main treatment effects.

G2. Discount rates

The local advantage in production may be driven not by organized criminal activity and law enforcement corruption, but rather by different optimal extraction profiles given underlying time preferences. This is a plausible mechanism if local companies have shorter time horizons than multinationals. I test this argument directly by estimating extraction profiles for local and multinational fields. In petroleum engineering, oil production typically follows what is called a "decline curve," which models oil output as an exponential decay function over time (Arps, 1945). The curvature of this extraction profile suggests an implied discount rate, given field characteristics – steeper declines suggest over time less patience. As such, impatient local firms should extract more oil earlier in the life cycle of the field.

I estimate and plot decline curves in Figure A19 for the subsamples of multinational and locally-operated fields separately. Instead of directly estimating the parameters of an exponential decline curve, I model output as a flexible nonlinear function of age, following

$$y_{it} = \alpha + g(a_{it}) + \epsilon_{it}$$

I estimate g using a cubic spline with 7 knots spaced evenly every ten years.[54] I cut the data below at 5 years, since there are no local field-year observations younger than this age.

The results indicate a clear decline curve for both multinational local firms. These curves both peak between 10-20 years of field life near 10 million barrels annually and decline steeply thereafter, approaching zero between 20 and 30. Both curves also suggest a small revival of in the later years of the field lifecycle, perhaps driven by new well drilling.[55]

G3. Grievance

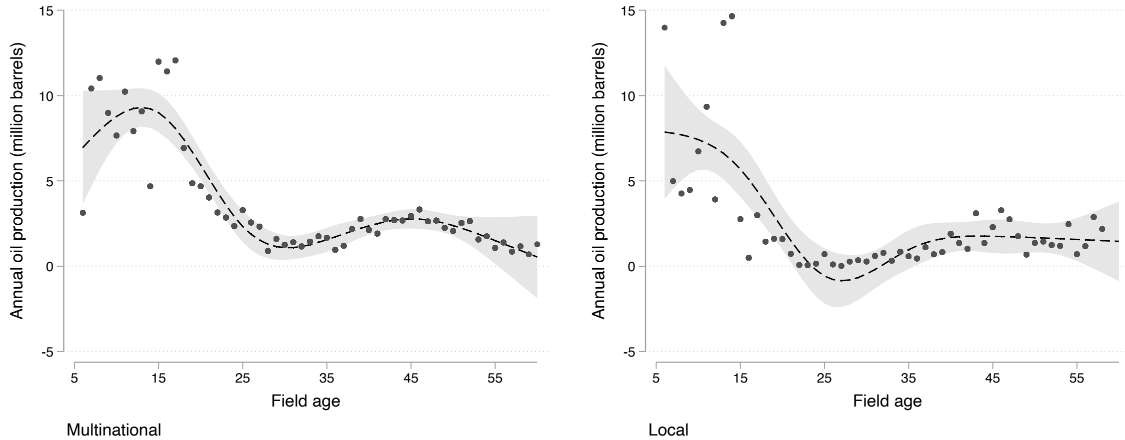
Pipeline vandalism may be driven by grievance rather than economic motives (Buhaug, Cederman and Gleditsch 2014). Niger Deltans retain longstanding, justified grievances against multinationals due to a long history of corporate malfeasance and environmental pollution (Obi and Rustad 2011). Sentiments toward local companies may be considerably better, resulting in reduction in grievance-driven attacks. If so, we should also expect to observe a reduction in community protest, the most direct expression of grievance. Protests against oil companies are common in host communities, affecting 21% of all fields at any point during the sample period. In Table A32, I re-estimate the main specification using the number of protests (columns 1-2), oil-related protests (columns 3-4), and riots (columns 5-6) within 15

[53]Note that the 0-10 km coefficient in the oil theft panel will exhibit mechanical spillovers because of spatial correlation and should be disregarded.

[54]The intervals for the cubic polynomial are $[0,5], [5,15], [15,25], [25,35], [35,45], [45,55], [55,60]$

[55]Decline curves are typically modeled at the well-level, though we only have output data at the field-level.

Figure A19. Extraction curves by type, cubic spline



Note: Figure plots extraction curves – the level of oil output by the age of field – for subsamples of multinational (left panel) and local fields (right panel). Dots indicate mean output by age for each subsample. Fit is estimated using a cubic spline with the following knots: $[0,5], [5,15], [15,25], [25,35], [35,45], [45,55], [55,60]$. Field age is defined as the difference between the current year and the year of the first well drilling.

kilometers of the field as the outcome variable. The point estimates are, if anything, positive, but generally insignificant. There is no evidence of a change in grievance as a result of localization.

Table A32—Divestment and riots and protests

Outcome	All protests		Oil protests		Riots	
	(1)	(2)	(3)	(4)	(5)	(6)
Local firm	0.273	0.144	0.004	-0.014	0.193	0.155
	(0.168)	(0.170)	(0.017)	(0.022)	(0.306)	(0.318)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls × Year FE	No	Yes	No	Yes	No	Yes
Observations	3183	3183	3183	3183	3183	3183
R^2	0.353	0.410	0.160	0.238	0.401	0.452

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Outcome variable is given in table header, defined as the total number of incidents within 15 km of the field. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

G4. Local employment and welfare

Part of the rationale behind indigenization is that local firms may generate positive spillover effects to local communities. If this is the case, then it's possible that the effects we see are driven by higher opportunity costs for attracting labor into the criminal sector. In particular, if spillovers improve employment opportunities for young men, then the gangster's cost c may rise as labor costs rise. All else equal, this increases the bargaining space between firms and law enforcement, since gangs have less profit with which to offer competing bribes.

To test this hypothesis, I use data from three rounds of Nigeria's General Household Survey, a 3-wave panel survey covering 16,211 working-age (age 15-60) Nigerians in 500 villages from 2010-2016. I link each village to its nearest oilfield in order to identify villages treated by localization of nearby fields. I then drop all villages further than 50 km to their nearest oilfield. For individual (or household) i in village v near to field f at time t, I estimate:

$$y_{ivft} = \alpha + \psi local_{ft} + \delta_t + \xi_f + X'_{ivft}\beta + \mu_{ivft}$$

y_{ivft} is a labor market outcome, including employment, employment outside the home, self-employment, and employment in household agriculture, as well as the log of overall per capita household consumption. Household-level controls in X are distances to roads, population centers, markets, borders, and state capitals; village-level controls are slope, altitude, annual temperature, annual rainfall, and a rural indicator. Each of these time-invariant conditions is interacted with year effects. Standard errors are clustered at the field level.

Results of this estimation are given in Table A33. Columns (1)-(4) estimate using the entire sample of fields with various combinations of year, month, field, and state-by-year fixed effects, as well as the interacted controls. Columns (5) and (6) exclude all individuals residing in a village whose nearest oilfield was offshore, where spillovers are less likely to occur. The results show no effect on the level of employment (Panel A). For the composition of employment, I do not find any statistically significant changes in employment outside the home (Panel B) or employment in household agriculture (Panel D). However, there does appear to be an increase in self-employment (Panel C) by 9-10 percentage points, significant at 1%. Since overall employment does not change, this effect offsets small and statistically insignificant reductions in other categories. Panel E, shows no change in log household per capita consumption evolves after divestment. Overall, there is no evidence that localization creates meaningful positive economic spillovers for nearby oil-producing villages.

G5. Corporate social responsibility

The most visible local benefits of oil extraction are host community investments in the form of corporate social responsibility (CSR). It may indeed be more efficient for firms to provide CSR benefits to troubled areas to dissuade militancy and theft than to negotiate with the security forces. If local firms have a greater propensity to target their investment toward volatile communities, this mechanism could plausibly drive the observed effects. In 2016, voluntary expenditures on CSR projects by oil companies in host communities totaled 92.6 million dollars, a tiny fraction of the annual profits from oil theft, suggesting that these projects are unlikely to dissuade violence. Using data on all CSR projects in 2016, I regress the number and value of multinational or local projects at the village level on the lagged level of conflict, measured as either the cumulative number of militant attacks from 1997-2015 or the number of militant attacks in 2015. If companies follow a targeting policy, we should observe a positive correlation between conflict and CSR expenditures.

Figure A20 plots coefficients from these regression models. The top panel uses standardized CSR projects as the outcome to account for the fact that local firms are generally smaller and therefore have fewer projects overall, while the bottom panel use total CSR expenditure

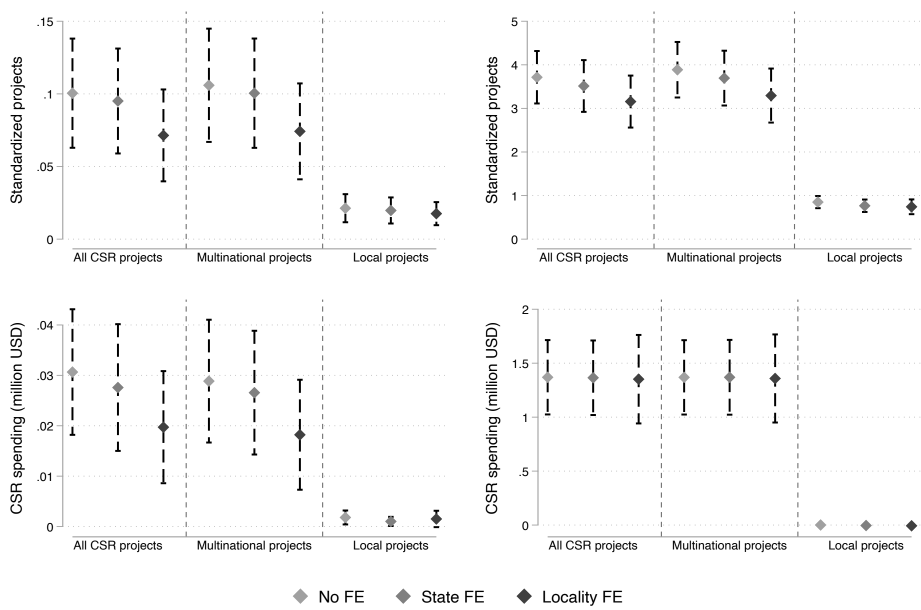
Table A33—Divestment and local employment

Sample	All				Onshore	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Employed						
Local firm	0.024	0.024	0.001	0.045**	0.022	0.000
	(0.028)	(0.030)	(0.025)	(0.019)	(0.032)	(0.029)
Observations	16827	16827	16827	16827	15616	15616
R^2	0.033	0.034	0.041	0.039	0.029	0.038
Panel B: Employed outside home						
Local firm	-0.020	-0.017	-0.033*	-0.023	-0.021	-0.014
	(0.020)	(0.019)	(0.017)	(0.018)	(0.018)	(0.020)
Observations	9225	9225	9225	9225	8551	8551
R^2	0.107	0.109	0.158	0.132	0.095	0.139
Panel C: Self-employed						
Local firm	0.098**	0.097**	0.108**	0.096***	0.109**	0.102**
	(0.037)	(0.037)	(0.044)	(0.033)	(0.044)	(0.049)
Observations	9225	9225	9225	9225	8551	8551
R^2	0.070	0.071	0.122	0.106	0.068	0.123
Panel D: Employed in household agriculture						
Local firm	-0.050*	-0.050	-0.025	-0.008	-0.078**	-0.055
	(0.029)	(0.031)	(0.037)	(0.037)	(0.031)	(0.038)
Observations	9225	9225	9225	9225	8551	8551
R^2	0.152	0.153	0.271	0.183	0.151	0.277
Panel E: Household consumption						
Local firm	0.035	0.038	0.038	0.028	0.038	0.030
	(0.072)	(0.069)	(0.073)	(0.045)	(0.079)	(0.069)
Observations	5119	5119	5119	5119	4750	4750
R^2	0.243	0.244	0.292	0.270	0.250	0.305
Field FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	Yes	No	No	No	No
Year × State FE	No	No	No	Yes	No	No
Controls × Year FE	No	No	Yes	No	No	Yes

Standard errors clustered at the field level in brackets. Outcome variable is given in the panel header. Sample is all individuals in the three waves of the GHS between the ages of 15-60 living in clusters within 50 km of an oilfield, except Panel E, which is at the household-level. All regressions use household-level sampling weights. GHS controls are cluster distance to road, population center, market, border, and administrative center, a rural dummy, slope, elevation, and mean annual temperature and precipitation. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

in millions of USD. The left panel use cumulative attacks up to 2015 on the righthand side, while the right panel uses attacks in 2015. For each specification, I estimate the bivariate relationship unconditionally, and with state or locality fixed effects. In general, there is evidence suggestive of targeting – prior local conflict is positively and significantly correlated with the number and value of CSR projects at the village-level. However, this aggregate relationship obscures substantial differences between local and multinational projects. Across all outcomes and independent variables, the correlation between CSR investments and conflict is much stronger for multinationals. This suggests that the main results are unlikely to be driven by superior targeting by local firms. If anything, the results are consistent with multinationals leaning more heavily on CSR to mitigate conflict risk that local firms because they are unable to leverage political connections to bargain directly with gangs.

Figure A20. CSR projects and local conflict



Note: Figure plots coefficient estimates of the village-level correlation between oil company expenditure on corporate social responsibility (CSR) in 2016 and lagged militant activity. The outcome is measured as either the standardized number of CSR projects (top panel) or total expenditure (bottom panel), either in total or disaggregated by local and multinational projects. The independent variable is measured as the number of oil-related militant attacks in 2015 (right panel) or the cumulative number oil-related militant attacks from 1997-2015 (left panel). Model specification is indicated in subfigure headers. Models are either unconditional or include state or locality fixed effects, as indicated in legend.

G6. *Monopoly of violence*

It is possible that local firms are connected to local gangs, and leverage violence to force multinational divestments. Once these divestments are complete, crime falls and profits are shared between gangs and local firms. There are several reasons that this explanation is unlikely. First, the use of violence to force divestments implies a violation of parallel trends, as violence should *rise* on localized assets relative to controls in the run-up to divestment. There is no evidence for this in Figure 3. Second, this explanation is not consistent with a rise in enforcement after divestment (Table 4), since enforcement should not be needed to reduce violence. Finally, this explanation depends on gangs' monopoly of violence, since firms would not collude ex-ante if gangs were unable to enforce peace ex-post. As such, the collusion model predicts that effects on output and theft should be largest for fields located in areas of monopolist control, where one (or a small number) of large gangs hold sway.

Table A34 investigates this hypothesis. To measure monopoly control, I use data from Rexer and Hvinden (2022) that maps the location of major armed groups in the Niger Delta conflict. For each field, I calculate the total number of organized armed groups operating within 50 kilometers of the field, and divide the sample into three groups: fields with 0, 1, or > 1 armed groups within 50 kilometers. I then interact the local treatment variable with indicators for these groups, with monopoly as the omitted group. There is no evidence that the effects are systematically larger in areas under monopolist control.

Table A34—Divestment and oil field outcomes: local monopoly of violence

Outcome	Output		Oil theft	
	(1)	(2)	(3)	(4)
Local firm	0.723*	0.798**	-4.252***	-5.401***
	(0.386)	(0.393)	(1.281)	(1.438)
Local firm $\times > 1$ armed groups in 50km	-0.214	0.375	-1.159	0.565
	(0.500)	(0.553)	(1.807)	(2.156)
Local firm $\times$ No armed groups in 50km	0.847	0.468	-1.203	-0.704
	(1.234)	(1.059)	(2.415)	(2.453)
Field FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls $\times$ Year FE	No	Yes	No	Yes
Observations	2476	2476	3183	3183
R^2	0.863	0.881	0.723	0.761

Standard errors in parentheses are clustered at the field level. Sample is the panel of 314 oilfields from 2006-2016. Outcome variable is the total number incidents or fatalities from state violence against civilians within 15 km of the field. Controls are latitude of the field centroid, distance to coast, distance to Niger River, and distance to the capital, all measured in km. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Welfare estimation

In Appendix B.B7, I use the model to derive an expression for the total welfare effect of indigenization. In this section, provide full estimation details for each of the welfare parameters. All dollar values are deflated relative to 2016 USD. The equation for the change in total surplus ΔS as a result of indigenization is:

$$\Delta S = \sum_{f=\ell} \frac{\partial Q_f}{\partial f}(p^* + e_c) - \frac{\partial C_f}{\partial f} + \frac{\partial n_f}{\partial f}[q(\tilde{p} - p^*) - p^*\kappa - c + e_n + VA_n] - \frac{\partial \varrho_f}{\partial f}\alpha\eta + \frac{\partial m_f}{\partial f}e_m + \frac{\partial g_f}{\partial f}e_g + \frac{\partial v_f}{\partial f}e_v$$

H1. Environmental externalities

Oil spills: Because it affects the quantity of oil spilled, gas flared, and overall output, indigenization generates environmental externalities. Oil spill costs are captured in the term $\frac{\partial n_f}{\partial f}e_n + \frac{\partial m_f}{\partial f}e_m$, where n_f is oil theft spills, with cost e_n and m_f are operational malfunction spills with cost e_m. I measure these costs using estimates of the impact of oil spills on child mortality and land values.

The estimates for $\frac{\partial n_f}{\partial f}$ and $\frac{\partial m_f}{\partial f}$ are simply the treatment effects of divestment on sabotage spills and malfunction spills, respectively. I use field polygon estimates from Table A20 Panels A and B, to eliminate double-counting of effects. I then take estimates of e_n and e_m from Bruederle and Hodler (2019), which studies the effect of oil spills on infant mortality in the Niger Delta. However, these estimates do not differentiate impacts of sabotage vs. malfunctions, and so I re-weight effects by average relative size of spills and relative frequency to recover spill type-specific estimates. I then multiply these coefficients to get an annual field-level average estimate of the impact of divestment on infant mortality. Then, following the method in Bruederle and Hodler (2019), I aggregate number of life-years gained by multiplying total live births over sample period (from the World Bank) $\times$ share of population within 10km of treated field (from the DHS) $\times$ life expectancy (from WHO). I then value these life-years at three times Nigerian GDP per capita in 2016.[56]

To this I also add estimates of the agricultural productivity costs of oil spills using plot-level data from the General Household Survey on land values. In this repeated cross section, I regress land values on a time-varying indicator for whether the plot is within 10 kilometers of an oil spill site after the spill occurs, as well as state-by-year, state-by-month of year, and village fixed effects. The results indicate that exposure to oil spills reduces land value by approximately 1.7 USD per square meter. I then re-weight these effects to disaggregate by type as above, and multiply by $\frac{\partial n_f}{\partial f}$ and $\frac{\partial m_f}{\partial f}$, to get the net field-level effect of divestment on agricultural land value. I then identify all GHS sample villages within 10 kilometers from a treated oil field, aggregate the agricultural land area exposed to divestment, and multiply by the net field-level on land value.

Gas flaring: The gas flaring health externality is captured in the term $\frac{\partial g_f}{\partial f} e_g$, where e_g is the local cost of gas flares via incidence of respiratory disease plus the social cost of CO2 emissions. For $\frac{\partial g_f}{\partial f}$ I use estimates from Table A12, column (8). To value e_g, I use estimates from Alimi and Gibson (2022) for the impact of gas flaring on child respiratory disease prevalence in the Niger Delta. Combining these estimates with DHS data from 2018 on the location of children in Nigeria relative to treated oilfields, I aggregate a distance-weighted exposure variable that implies a 0.7 percentage point increase in the prevalence of childhood respiratory illness as a result of indigenization. I value the costs of this increase using Nigeria-specific estimates of the burden of respiratory illness from Soriano et al. (2020). Alimi and Gibson (2022) also identify flaring impacts on stunting, and combining these estimates with the increase in flaring implies a 2% increase in the Nigerian stunting rate. I value this increase using country-specific estimates from Galasso and Wagstaff (2019) on the costs of stunting. Lastly, gas flaring increases carbon emissions itself, which has a global social cost. I use a conversion factor from the US Energy Information Administration to convert gas emissions into CO_2, and then multiply this by a social cost of carbon of USD 51/ton, the rate currently used by the US federal government.

Carbon cost of production: Finally, $\frac{\partial Q_f}{\partial f} e_c$ is a global externality that captures the carbon cost of new output. $\frac{\partial Q_f}{\partial f}$ is the unexplained output effect from Table A11, row 2, which represents "new" production (i.e., not a transfer or deferred production), and summed across all

[56]This simple rule of thumb is taken from research done by the philanthropy evaluator GiveWell.

504 treated field-years. e_n is measured by the social cost of carbon multiplied by a conversion factor from the US Environmental Protection Agency to convert barrels of oil into tons of CO_2.

H2. *Production effects*

Production effects: Indigenization generates new output unexplained by theft rather than simply transferring output back to firms and the state. In Scenario 1, $\frac{\partial Q_f}{\partial f} p^*$ captures the unexplained output effect of indigenization valued at current international prices from Table A11 row 2, plus the deferred production effect from Table A11, row 5 both summed over all treated field-years. In Scenario 2, we assume that the unexplained output effect is instead deferred production. Reading from Table A11, the entire residual output effect net of transfers, $0.963 - 0.236 = 0.712$ million barrels, is valued as deferred production. Therefore, the Scenario 2 production effect is $0.727/0.395 = 1.84$ times larger than the baseline deferred production effect.

But local firms may have greater extraction costs, counteracting some of this surplus-creation; $\frac{\partial C_f}{\partial f}$ is the cost effect. Unfortunately, we lack firm input data and are unable to estimate this quantity, so the final welfare estimate may be considered an upper bound. Lastly, I also exclude dynamic welfare costs arising from local firms reducing investment, since firm-level data on well drilling from 2010-2014 reveals that local firms' investment share matches their market share (25.2% of output and 27.8% of wells drilled).

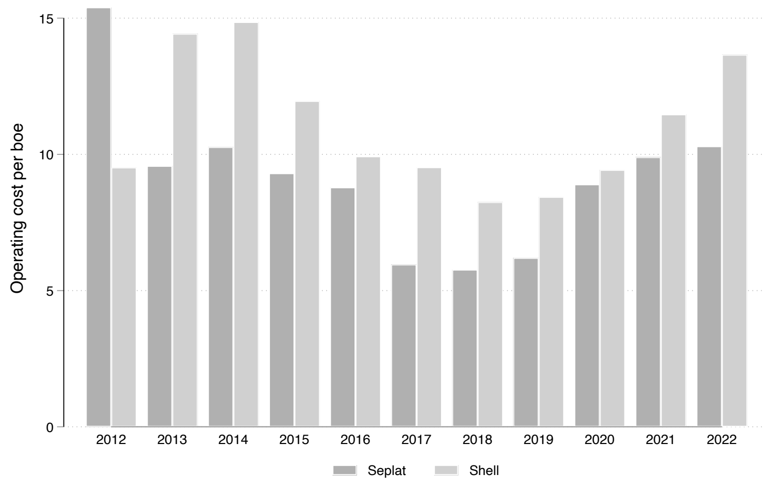
Figure A21 plots annual average operating costs for two firms – Shell, a multinational, and Seplat, a local firm – using data from annual company reports. Comparing average operating costs suggests that local firms are no less efficient than multinationals; in fact, Seplat has lower average operating costs than Shell over the period. However, this comparison has several caveats: Shell costs include the entire Africa portfolio, the data don't adjust for the asset mix (onshore / offshore composition), as the only Nigerian firm listing internationally, Seplat is likely to be positively selected, and operating expenditures likely include security spending, and so do not purely measure technical efficiency. As such, it would be misleading to use these cost differentials as an input in the welfare model. Still, the similarity in the magnitudes does not suggest a particularly large cost advantage for multinationals.

H3. *Illicit market effects*

Crime externalities and enforcement costs: An important social cost of the black market is violence, which contributes to premature death among the young men involved in gang activity. The is captured in $\frac{\partial v_f}{\partial f}$, the fatalities estimates taken from Table A12 column (1), and e_v, the cost of these deaths. I convert the effect on fatalities into life-years by assuming the average black market participant is 25 years old. I then value these life-years e_v using the same VSL estimate as above, 3×2016 GDP per capita. However, increased law enforcement effort to combat oil theft requires mobilizing real resources, captured by $\frac{\partial \varrho_f}{\partial f} \alpha\eta$, where ϱ is the enforcement level and $\alpha\eta$ is the per-unit cost of policing. While we have data on enforcement quantities, we lack data on policing costs and so are unable to estimate this cost.

Illicit market effects: Finally, the reduction of oil theft recoups several inefficiencies associated with black market theft, sale, and processing of crude oil. To estimate the costs of black market activity, I use summary statistics on illicit extraction costs, price discounts, and refining value added from Stakeholder Democracy Network (SDN), a local Niger Delta NGO

Figure A21. Operating costs for two firms



Note: Figure shows average operating expenditures per barrel of oil equivalent (boe) in current USD for two firms, the multinational Shell and the local Seplat Nigeria. Figures come from company annual financial reports, and cover the entirety of Shell's Africa portfolio.

with experience conducting field surveys of actors in the illegal sector value chain. These estimates come from the data underlying SDN (2019*b*), which was provided to the author by SDN. Each per-unit cost of black market activity is multiplied by $\frac{\partial n_f}{\partial f}$, the change in the size of the black market. I measure this as the number of barrels transferred from the black to the official market as a result of indigenization, estimated in Table A11 row 4.

First, black market extraction costs duplicate extraction costs already paid by the firm: these are $\frac{\partial n_f}{\partial f}c$, the change in theft multiplied by the per-unit theft cost. I measure c using the cost of theft per barrel, which is just USD 0.75 according to SDN (2019*b*); I then aggregate across all treated field-years. Second, black market crude oil is sold at a substantial discount to the global oil price, entailing welfare losses, given by $\frac{\partial n_f}{\partial f}q(\tilde{p} - p^*) < 0$, where $\tilde{p}$ is the black market price for crude oil. On average, participants in the SDN (2019*b*) survey between 2012-2016 sold crude oil for just USD 12.21/barrel, a nearly 85% discount on the average global crude price over the sample period. In Scenario 1, I multiply this per-barrel discount by the change barrels transferred from the black market and aggregate across all treated-field years to obtain the total surplus regained by indigenization, while in Scenario 3 these are treated as transfers and set to zero. Third, $\frac{\partial n_f}{\partial f}p^*\kappa$ are spillage losses associated with theft. However, as in Appendix C.C4, I exclude these estimates because their magnitude is negligible relative to the scale of other losses.

Finally, despite these inefficiencies, the illicit market also creates surplus. In particular, value-addition in the illicit sector, $\frac{\partial n_f}{\partial f}VA_n$, comes from illicit refining and may be sizable. This is particularly significant in the Nigerian context, where the domestic refining sector is severely constrained and most crude oil is exported directly; for example, in 2015, official refineries processed just 1.3% of total Nigerian oil production. To measure per-barrel value added VA_n in the illicit refining sector, I use estimates from a survey of refining camps between 2012-2016 (SDN, 2019*b*), which shows that profits are on average USD 14.76 per barrel of crude processed. I aggregate these gains across all treated field-years; these gains are of course lost as a result of indigenization.