Natural Resources and Civil Conflict: Evidence From a New, Georeferenced Dataset*

Michael Denly[1], Michael Findley[1]†, Joelean Hall[1],

Andrew Stravers[1], James Igoe Walsh[2]

[1] University of Texas at Austin, Innovations for Peace and Development

[2] University of North Carolina at Charlotte

May 24, 2019

*This project was supported by the Department of Defense Minerva Research Initiative and Army Research Office under grant number W911NF-13-0332 (Principal Investigator: James Walsh). We thank David Weinberg for providing the Multicolour data. Those interested in the Multicolour data may contact David at info@multicolour.com. We thank Daniela Blanco for assisting with the price conversions. For comments and advice, we thank Victor Asal, Steven Beard, Kyosuke Kikuta, Michael Gibbs, James Piazza, Jean-Claude Thill, and participants at the American Political Science Association (APSA) Conference and Midwest Political Science Association conference. We would like to thank Kyosuke Kikuta and the following graduate research fellows and research affiliates at Innovations for Peace and Development at the University of Texas at Austin for assisting with the data collection: Nicole Pownall, Annie Kilroy, Erica Colston, Vanessa Lizcano, Erin Eggleston, Iasmin Goes, Oliver Babcock, Rahim Chaudry, Daniel Chapman, Garrett Shuffield, Akshat Gautam, Abby Brown, Delainey Peterson, Evelin Caro, Jonathan Velasquez, Alex Walheim, Amanda Long, Haley McCoin, Jake Barnet, Tawheeda Wahabzada, Bianca Remmie, Anna Scanlon, Alejandra Gaytan, Vishal Duvuuru, Jennifer Johnson, Sam Gorme, Miles Hudson, Sarah Fischer, Vivianna Brown, Leah Havens, Daniela Garcia, Jennifer McGinty, Lizzette Marrero, Nathalia Rojas, Josh Hamlin, Maren Taylor, Johnny Shaw, Regan Seckel, Kiara Hays, Kolby Vidrine, Kate Adams, Anita Basavaraju, Arijit Paladhi, Arvind Ashok, Brandon Gajeton, Carlos Diaz, Destiny Alvarez, Domingo Salerno, Drew Burd, Hannah Greer, Jade Tucker, Tyler Morrow, Ji Na Gil, Kanika Varma, Karan Kanatala, Kimberly Schuster, Levi Malloy, Lila Al-Kassem, Mackenna Shull, Mariana Caldas, Marifer Guerrero, Patrick Golden, Samiya Javed, Michael Hankins, Sam Bennett, Skyler Thomas, Samantha Shoff, Beomhak Lee, Benjamin Vega, Mobin Piracha, Ashley Frey, Rama Singh Rastogi, Adityamohan Tantravahi, Jake Reynolds, Kelvin Efiya, JP Repetto, Nick Romanov, Nikola Skerl, Keeton Schenck, and Ethan Masucol.

†Corresponding author. mikefindley@utexas.edu

Abstract

Scholars have long examined the relationship between natural resources and conflict at the country level. More recently, researchers have turned to subnational analyses, using either individual countries or subnational data for a small number of resources in sub-Saharan Africa. We introduce a new sub-national dataset of 183 resources that adds many resource types, locations, countries, and local price data from Africa, the Middle East, Asia and Latin America. We examine how conflict incidence varies with the value of the collective set of resources in a given location using world prices. We then introduce new local price data, which is more relevant for conflict dynamics. Because local prices can be endogenous to conflict, we instrument local prices using U.S. and global prices. We find that subnational resource wealth is associated with higher levels of conflict using some specifications, though the results vary widely by data source and world region. Using the instrumental variables strategy lends the strongest support to this positive relationship, but only for African countries.

1. Introduction

Over the last two decades, social scientists have investigated the "resource curse"—the proposition that an an abundance of non-renewable natural resources has negative political, social, and economic consequences (for a recent and thorough review, see Ross (2015)). Much of this work has focused on links between resources and violent conflict at the country level (De Soysa, 2002; Fearon and Laitin, 2003; Collier and Hoeffler, 2004; Ross, 2004, 2012; Humphreys, 2005; Ross, 2006; Cotet and Tsui, 2013; Bazzi and Blattman, 2014; Lei and Michaels, 2014; Esteban, Morelli and Rohner, 2015; Paine, 2016; Menaldo, 2016). The role of petroleum wealth in fomenting conflict has received the most attention because it is the most valuable commodity at the global level (Ross, 2012) and because of the availability of data measuring national petroleum production and reserves at the national level. This has led some to conclude that the "resource curse" is, at the national level, really an oil curse (Ross, 2012), and studies of multiple resources have found few links between countries' overall resource wealth and conflict (Bazzi and Blattman, 2014) or that the link between resources and conflict is likely moderated by geographic and political factors (O'Brochta, 2019).

More recent research on the resource curse has taken a decidedly micro turn (Nillesen and Bulte, 2014). One reason for this shift is the recognition that while petroleum may be the most valuable and widely-traded commodity at the global level, some countries are rich in other types of resources, such as diamonds or other minerals, that may promote violent conflict. Another reason is that many conflicts are essentially local in nature, leading to violence in specific regions while the rest of the country experiences little violent contention. For these reasons, Koubi et al. (2014, 12) suggest that "the analysis of disaggregated data that are also able to capture the location and spatial aspects of resources clearly seems to be the most effective approach." That is especially relevant for understanding local conflict dynamics, the incentives for national leaders to tolerate conflict that does not threaten their control over resource revenues (Koubi et al., 2014), and how resources influence succesionist conflicts (Ross, 2012; Asal et al., 2015). Some studies draw on local data from individual countries to analyze how differences in resource endowments influence violence across localities (Aragón and Rud, 2013; Dube and Vargas, 2013; Mähler and Pierskalla, 2015; Maystadt et al., 2014; Sanchez de la Sierra, 2019). Another strand of research examines how profiting from resources by rebel groups influences conflict dynamics (Asal et al., 2015; Fearon, 2004; Conrad et al., 2019; Walsh et al., 2018).

More recent work has analyzed how resource wealth influences violence at the local level in multiple countries (Berman and Couttenier, 2016; Berman et al., 2017; Christensen, 2019), which promises to yield more general findings. However, this last approach faces significant data limitations; there is no non-proprietary dataset that provides local information about resource wealth for a wide range of resources across many countries. Balestri, Lujala, and their colleagues have developed such data for gold, diamonds, gemstones, and petroleum (Lujala, Gleditsch and Gilmore, 2005; Lujala, 2009, 2010; Balestri, 2012; Balestri and Maggioni, 2014; Balestri, 2015). These data sources are among the most widely-employed in the study of resources and conflict at the local level, in part because they are open source and included in the PRIO-GRID dataset (Tollefsen, Strand and Buhaug, 2012), but their coverage of non-renewable resources is limited.[1] Table 1 compares the dataset introduced in this paper, the Global Resources Dataset (GRD), to other open sources of data that contain spatial, subnational information on non-renewable natural resources and are frequently employed by researchers. The GRD contains information on far more resources than do the other data sources, and it provides more detail across the relevant dimensions than do other data sources. The GRD includes information for all countries in Africa and most countries in Latin America and Asia.[2] The spatial unit in the GRD is the point location (i.e. latitude and longitude) of the extraction or production site. Other data sources also provide geolo-

[1] Other authors have put forth some limited sub-national data of some key resources as well (e.g. Gervasoni, 2010; Diaz-Rioseco, 2016; Hong, 2018), but these data are not systematically available for many countries and resources.

[2] See Appendix C for a list of countries currently included in the GRD. In future work, we plan to extend the GRD to cover all countries in Africa, Asia, and Latin America. Since a key advantage of the GRD is that it allows research on subnational variation in resource endowments, we view the 103 countries included in the current version as sufficient for many analyses.

cation at this level of resolution, but they provide information on fewer resources than the GRD. Berman et al. (2017) and Harari and La Ferrara (2018) provide data at the grid cell level, which may be too highly aggregated for some research purposes. The GRD data is also time-varying, documenting both when production starts in a location and when it ends, and provides annual data on output, world prices, and local prices. These variables should be of particular value to conflict researchers, as changes in production, output, and prices might both cause and be influenced by nearby violent events.

Other researchers have made important contributions to the literature using proprietary data that measures local resource endowments across countries (Berman and Couttenier, 2016; Berman et al., 2017; Christensen, 2019). But these data sources are not widely available to researchers and still include only a small number of resources or are limited in the number of countries for which they contain data. Berman et al. (2017), for example, include fourteen minerals in sub-Saharan Africa. While the replication data for Berman et al. (2017) is available, it aggregates across multiple resources at the grid-cell level. This means that other researchers cannot use the data to identify the specific locations of resource extraction sites or disaggregate details within a grid-cell. Furthermore, many existing data sources lack information about world or local prices of the non-renewable resources they document. While world prices for many commodities are now available (Bazzi and Blattman, 2014), the absence of local prices is a potentially important omission, as the local value of a resource likely exerts a more powerful influence on conflict dynamics.

Table 1: Spatial Natural Resource Datasets

Dataset	Countries	Spatial Unit	Time-Varing	Output	World Prices	Local Prices	Resources
Global Resources Dataset	103	Point	Yes	Yes	Yes	Yes	183 resources
Berman et al. (2017)	52	Grid cell	Yes	Yes	Yes	No	14 resources
Harari and La Ferrara (2018)	46	Grid cell	Start only	No	No	No	85 resources
Balestri (2015)	110	Point	Start only	No	No	No	Gold
Lujala, Gleditsch and Gilmore (2005)	52	Point	Start only	No	No	No	Diamonds
Lujala, Röd and Thieme (2007)	107	Polygon	Start only	No	No	No	Oil and gas
Lujala (2009)	107	Point	Start only	No	No	No	Gemstones
Buhaug and Lujala (2005)	86	Polygon	Start only	No	No	No	Coca bush, opium, poppy, cannabis

The GRD provides information on a much larger number of resources, the location of their extraction and production, their output, and their local prices. The GRD includes resources that are overlooked in extant data sources, such as iron ore, stone, and phosphates, and it also includes resources that involve downstream refining and production such as petrochemicals, steel, and cement. In total, we collected information on 183 unique resources, of which we are able to collect price information for 87% of the resource-location-years.[3] To obtain these data, we coded detailed country reports from the United States Geological Survey (USGS) for countries or independent territories in Africa, the Middle East, and parts of South America and Southeast Asia.[4] Overall, the new subnationally geo-referenced dataset contains $8,534$ unique country-resource-locations and $42,597$ country-resource-location-years.

To demonstrate the value of this new data source, we examine how the collective value of resources in a given location correlates with the incidence of conflict. We pool the different resource types and use relevant multipliers to compute comparable values, such that we can understand better the overall value of non-renewable resources in a a given location. In conducting this main analysis, we find mixed results. When examining sub-Saharan African countries only and using the Armed Conflict Location and Event Dataset (ACLED) and Georeferenced Event Dataset (GED) measures for conflict, as natural resource values increase in a given location, the likelihood of conflict incidence tends to increase. This is consistent with much of the work on sub-national resources and conflict, which has focused primarily on sub-Saharan Africa (Berman et al., 2017).

We then extend our analysis by using local price data, which can differ for each country, rather than relying on world prices. Using local prices is an improvement in key ways, especially because local prices are more relevant to actors on the ground. With that said, local prices are also likely endogenous to conflict dynamics, and as such require an identification strategy. Accordingly, we instrument local prices using U.S. prices, which should not be heavily dependent on local prices, as well as with world prices, though they are less indepen-

[3] The list of resources included in the GRD can be found in the Appendix C.
[4] The list of countries included in the GRD can be found in the Appendix C.

dent given that world prices are defined through the aggregation of local prices. The use of U.S. prices is not without challenges, especially given that spillovers and interdependence are likely, and could therefore may make the instrument weak, something we discuss at length in the manuscript. In extending the analysis to use U.S. and world prices as instruments for local prices, we find strong evidence that higher natural resource values increase conflict incidence primarily in African countries, but like the base models that result does not hold in global perspective.

The paper proceeds as follows: we first discuss the GRD, including information about the resource locations as well as price information. We then carry out an investigation of the effects of natural resource values on civil conflict. As part of this, we implement an instrumental variables strategy that uses U.S. and world prices to instrument for local prices, which enables us to address concerns about endogeneity. Finally we sum up with concluding thoughts about what the use of more expansive data and analysis imply for future research on natural resources and conflict.

2. The Global Resources Dataset

This section provides an overview of the GRD.[5] The unit of analysis is individual natural resource extraction and production facilites in each year from 1994 to 2014. For each site, the dataset records the location, type of resource produced, output, and local and global prices. The dataset is based on annual narratives of the the minerals industries of most countries in the world produced by the National Minerals Information Center of the United States Geological Survey (USGS).[6] These country reports contain exhaustive entries for natural resource sites and production facilities around the world, their locations, names of the facilities, types of resources present, and the capacity of the site. We had multiple

[5] The dataset and codebook will be made publicly available. The codebook, which provides further details on our methods for collecting these data, appears in Appendix C.

[6] Available at https://www.usgs.gov/centers/nmic/international-minerals-statistics-and-information.

coders read each of these reports and extract information into a machine-readable format.[7] The USGS country reports most often simply give the name of the facility or the city/general vicinity in which it is located. These facility-years are the unit of analysis for the dataset. We took this information and used the Geonames, Google Maps, and Mindat databases to identify the most precise longitude/latitude possible. We applied a "precision code" to denote how close the latitude/longitude recorded is to the exact location of the facility. We recorded a "1" when the exact site was within the above databases itself, with accompanying satellite imagery of the facility. We recorded a "2" when the most precise we could be was the city in which/near which the site was located. The vast majority of sites were these two levels of precision. Less precise measures include a "3" or a "4," where we could be no more precise than the district or province in which the site is located, respectively. These levels of precision locate the site at the geographical center of the respective administrative division. Similarly, when we are unsure of the location of the site altogether, we recorded a "9," which places the site at the geographical center of the country. However, less than 10% of entries in the USGS data were so vague as to prevent subnational geolocation and warrant a precision code of "9".

The dataset identifies 183 unique natural resources at these extraction/production facilities, of which 87% of the resource-location-years have price data. The extraction/production sites includes both "natural" resources such as diamonds, oil, and gold, as well as the sites that refine them and those that produce downstream products such as steel or fuel. This comprehensive documentation of extraction and production facilities provides users of the data with the flexibility to address a wide range of research questions about the resource curse. For example, researchers interested in "lootable resources" can define the characteristics that make a resource lootable, and filter relevant data from the GRD. Similarly, those with an interest in particular resources that have been linked to conflict, such as petroleum or

[7] We implemented safeguards to ensure high quality data collection from the USGS country reports. First, we conducted two rounds of coding for all countries. At the end of the second round of coding, the coders randomly sampled each other's work and performed some triple-checks. A senior coder then performed spot checks throughout and adjudicated all difficult cases that were not initially clear from the documents produced by the USGS.

diamonds, or specific types of minerals, such as metals, can draw on the location, output, and price data available for these resources. The GRD also includes information regarding downstream refining and processing facilities, such as petroleum refineries and mineral processing plants. We have observed recent examples of rebel groups capturing such capital-intensive resources and profiting from their extraction and sale, so that their inclusion is a potentially important contribution to understanding how production facilities relate to conflict. Foremost among them is the Islamic State's capturing and exporting fuel from Syrian and Iraqi oil facilities, which according to some estimates earned the organization up to $1.5 million a day. Further examples are not hard to find, with the Movement for the Emancipation of the Niger Delta (MEND) group in Nigeria launching repeated attacks on oil facilities in that country. Algeria saw a similar attack from Al Qaeda in the Islamic Maghreb in 2013 on the In Amenas petroleum processing facility. During the 1990s and into the 2000s, Chechen rebels targeted oil pipelines and oil transport vehicles. Although some resources might be more easily "looted," such as gold, diamonds, or drugs, these examples show that all types of resources may generate local grievances and produce incentives to capture them.

While our data are similar in intent and precision to Lujala (2010) and Berman et al. (2017), our data are more expansive and cover a broader set of countries. First, our data include information on 183 resources, including tin, copper, cobalt, uranium, iron ore, and phosphate. Of course, not all of these resources are in every country, and some resources only show up in rare cases, but nonetheless, we include the full catalog from USGS. As we can see in the political tumult around some of these resources (phosphate in Morocco/Western Sahara and uranium in the Democratic Republic of the Congo, for instance) non-gemstones and non-hydrocarbons can play a significant role in local incentives and the political-economic structure of countries in a way that warrants analysis.

The USGS country reports also include estimates of annual output for each facility, which we include the GRD. For most of our locations, we identified annual production capacities for the years 2002–2014. For some locations, the data extend back to 1994.

While the inclusion of output marks an advance over existing open source natural resource datasets, researchers often want to estimate the value of such output, which requires price data. The GRD provides up to three prices for each natural resource. The first is the price of the resource in the United States (Matos, 2015). The second is the world price, obtained from the World Bank Economic Monitor (World Bank, 2018) and Multicolour.[8] The third is local prices obtained from United Nations Statistics Division (2018), which reports the export price of the resource. It contains prices specific not only to resources and years but also to each respective country. All price data are expressed in each resource's standard measurement unit, for which we then create multipliers so as to ensure congruence between outputs and prices.[9]

The USGS country reports also document the ownership structure of each facility, which some research has found influences the intensity of resource curse effects (Jones Luong and Weinthal, 2010). The following ownership structures are available in the data: artisanal, artisanal/military, cooperative, cooperative/industrial, industrial, industrial/government, and government. Unfortunately, sometimes ownership information is not available. Mixed categories exist for when there is more than one type of owner and neither owns a majority stake (i.e., greater than 50%). When any one of the above owns more than a 50% stake, it is classified as only one of the above categories.

The following maps depicted in Figures 1, 2, 3, and 4 illustrate the distribution of the natural resources globally, regionally, and then for Colombia and the Democratic Republic of the Congo. The maps illustrate the substantially broader coverage of our Global Resources Dataset. Recall that existing work focuses almost entirely on sub-Saharan Africa, especially recent in-depth sub-national analyses (Berman et al., 2017; Christensen, 2019). The maps also illustrates the greater variation in resource types. While the primary resources are what researchers typically imagine, such as oil or gold, we track a variety of other resources such

[8] Multicolour is a Hong Kong-based auction house that provides pricing information on many resources that are not available in other datasets. Those wishing for these data may contact its owner, David Weinberg, via email: info@multicolour.com

[9] Refer to the Codebook in Appendix C for more details.

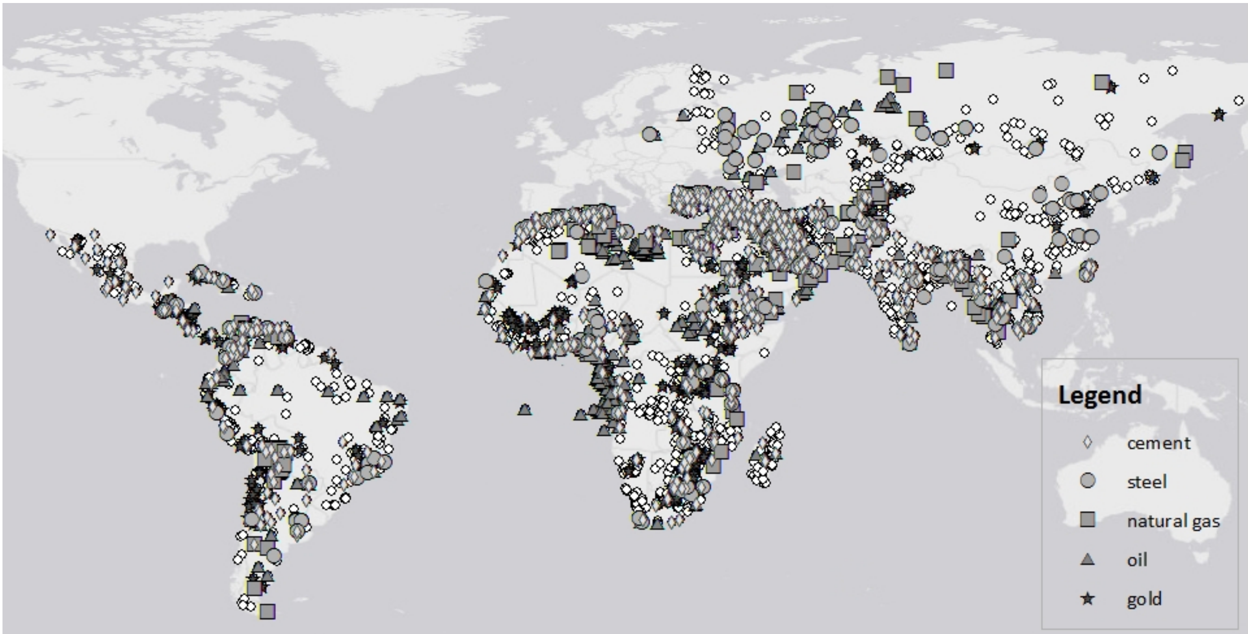

Figure 1: Natural Resource Locations Worldwide

as cement and steel, which involve downstream refining and production, but that have been tied to conflict dynamics in various ways.[10]

3. Research Design

We merged our new dataset based on spatial location in ArcGIS with the PRIO-GRID database (Tollefsen, Strand and Buhaug, 2012), which allows us to draw on that dataset's large number of covariates. The PRIO-GRID data divides the world into 0.5 degrees of longitude by 0.5 degrees of latitude squares (roughly 55 km x 55 km at the equator) to form a "grid."

We have coded all countries in sub-Saharan Africa as well as the Middle East and North Africa. Accordingly, we present the results for these two regions separately. We have coded an additional non-random set of Asian and Latin American countries. Because we do not

[10]We note again that the GRD covers all countries in sub-Saharan Africa, North Africa, and the Middle East. Although other areas may appear relatively complete, they are not yet finalized, although coding is in progress. Given the extensive coding protocols, other regions will not be complete for some time but will be made publicly available upon completion.

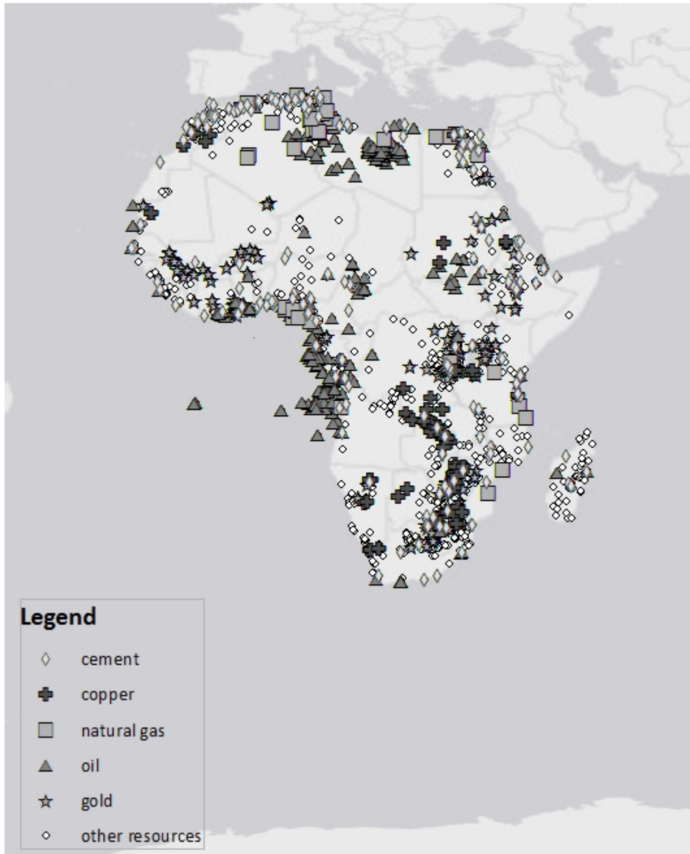

Figure 2: Natural Resource Locations Worldwide

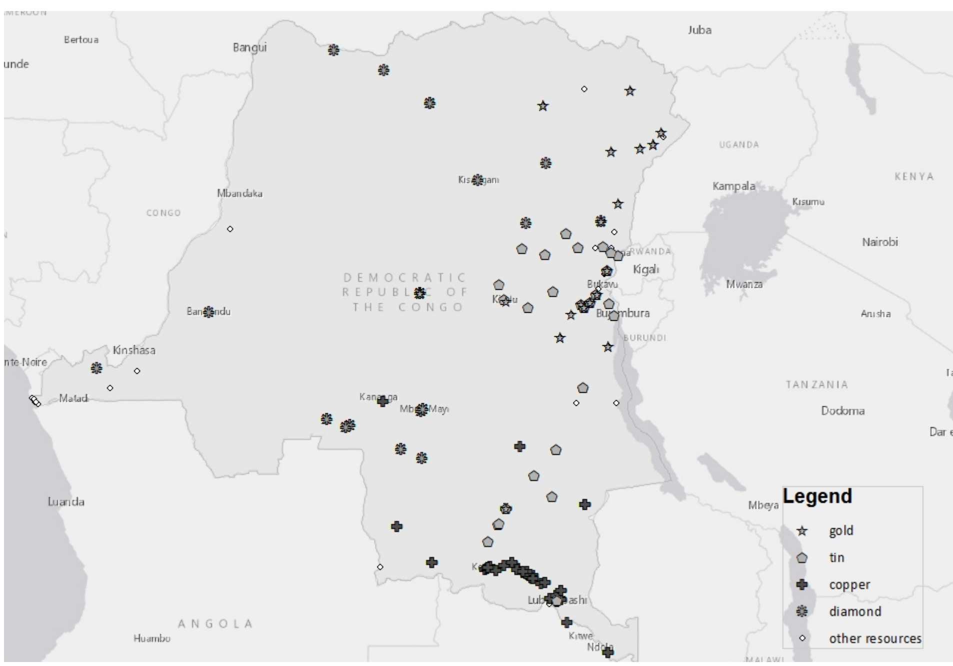

Figure 3: Natural Resource Locations in the DRC

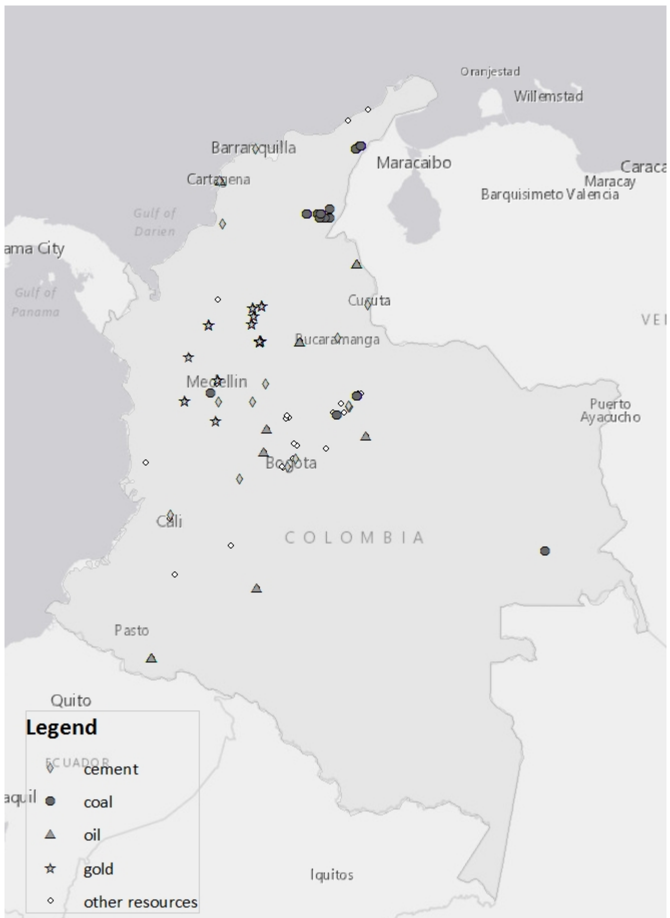

Figure 4: Natural Resource Locations in the Colombia

have a random or complete sample in these regions, we estimate a pooled model with all countries across all regions and report those following the sub-Saharan Africa and Middle Eastern country models.

3.1. Variables: Response, Explanatory, and Controls

For our dependent variables, we employ measures of conflict from the Armed Conflict Location Events Dataset (ACLED), the dependent variable employed in Berman et al. (2017), and the Uppsala Conflict Data Program's Georeferenced Event Dataset (GED). In particular, we examine conflict incidence —a dummy capturing conflict incidence in a grid cell as recorded by each of these datasets (Raleigh et al., 2010; Sundberg and Melander, 2013).

Our primary explanatory variable is the overall value of the collective set of resources in a grid cell, represented in constant 2010 USD. One advantage of the GRD over many existing datasets is that it includes both output and price information for a wide range of resources, allowing us to calculate the total value of resources produced at a location in a year. This contrasts with existing studies that rely on dichotomous measures of the existence of a resource, or that include only price but not output information (Berman et al., 2017). Measuring the total value of resources produced at a location is important because existing theory would lead one to expect that changes in these values influence incentives for conflict.[11] To do so, for a given resource we multiply the overall production amount in the year by the value of the resource in that year, and then repeat and sum for all resources in the grid cell, and then finally log that number. Following this approach allows us to capture some information about the full set of resources in a grid cell. Given the dispersion in the resource values, we logged the data. And to address some of the challenges with contemporaneous

[11] To calculate the value for resource extraction site, we compared the units for the output from USGS and the units for the prices by the World Bank, USGS, UN Comtrade and Multicolour. When the units did not match, we created a multiplier for the units to match. Then, we deflated our results using 2010 USD.

measurement, we lagged the data by one year.[12]

We supplement this measure by using local values, which are likely more theoretically relevant for most theories of resources and conflict. The local value variable is the export value of the resource in 2010 USD, based on the unit output for the resource extraction site from USGS and prices from UN Comtrade, where the resulting local values differ by country. This measure is not without challenges, most notably it likely responds to changes in conflict, while possibly also motivating conflict. We thus need to develop a causal identification strategy that minimizes the endogeneity in this measure, which we do below.

Finally, our study attempts to control for several potential confounders. These variables are at the grid-cell level. For data on ethnicity, we use the measure on excluded ethnic groups within each grid cell (Vogt et al., 2015). We take grid-cell (log) population data from HYDE (Goldewijk et al., 2017). We also control for level of development using nighttime lights data. In particular, we use the mean calibrated nighttime lights density at the grid-cell level, as measured by satellite imagery (Tollefsen, Strand and Buhaug, 2012; Elvidge et al., 2014). As is shown below, the model uses fixed effects at the grid-cell level, which explains the absence of a series of other traditional time-invariant control variables, such as distance to borders (Caselli, Morelli and Rohner, 2015) and mountainous terrain (Fearon and Laitin, 2003). Finally, we also generate spatially lagged conflict variables using the conflict data referenced above.

3.2. Estimation

Given that Berman et al. (2017) is one of the most recent and highest profile works in this area, we model the effects of natural resources on conflict similarly to provide some basis for comparison. Accordingly, we estimate our main models using a spatial heteroskedastic and autocorrelation consistent (HAC) model. Following Hsiang (2010), the spatial HAC

[12] The appropriate lag structure for the data is not immediately evident, and moving forward some theorizing is needed about the timescale on which natural resource extraction and production can be expected to translate into any conflict inducing behavior.

model takes the following form:

$$y_{kt} = \alpha + \beta_0 + \beta_p X_p + FE_k + FE_{it} + \epsilon_{kt} \tag{1}$$

where cell(k), time(t), and country(i) are all specified, FE_k are grid cell level fixed effects, and FE_{it} are additional country and year fixed effects. As should be apparent, the advantage of the spatial HAC is that it can account for multiple fixed effects. In addition, spatial HAC models estimate Conley (1999) standard errors that properly account for spatial dependence, and the Stata `.ado` routine of Hsiang (2010) allows us to specify spatial and serial correlation cutoffs. Although the spatial HAC model uses Ordinary Least Squares (OLS), and we have a binary dependent variable, our large dataset contributes to the statistical consistency of our estimates, making them (arguably) asymptotically unbiased (Berman et al. (2017) use a similar approach).

3.3. Identification through Instrumental Variables

In our primary models, discussed above and reported below in Tables 2 and 3, we lag the natural resource price variable, which can be an important though not sufficient step towards avoiding endogeneity. As a next step, in this section we introduce an instrumental variables approach that allows us to include information on the local (export) prices of natural resources.[13] To do so, we use two-stage least squares to instrument local price values with U.S. price values and world price values.

3.3.1. Criteria to Satisfy for Instrument Validity

For any instrument to be valid, it must satisfy several criteria (Angrist, Imbens and Rubin, 1996). We first discuss the first-stage, monotonicity, and the stable unit treatment

[13] Including an approach to obviate potential endogeneity between natural resources and conflict is a specific recommendation of a recent literature review from Koubi et al. (2014).

value (SUTVA) assumptions. Given that the ignorability/independence and the exclusion requirements require more explanations, we devote specific subsections to these topics.

First, a valid instrument must have a first-stage relationship: $COV(D, Z) \neq 0$. For our instrument, there must be a relationship between the endogenous variable (local values, D) and the instrument (U.S./world values, Z). In our case, log local values correlate with log World Bank values (used by Berman et al. (2017)) at 0.8, and log USGS values correlate with local values at 0.54.[14] Conventionally, instruments are thought to be strong if the F-statistic is above 12. In all of our models with control varaibles (other than for Asia), the F-statistic ranges from 29 to 225. In the Asia model, the F-statistic is 9. Accordingly, in most of the models the instrument is strong.[15]

Second, the instrument must satisfy the monotonicity assumption: $Pr(D_1 \geq D_0) = 1$ (Kern and Hainmueller, 2009). Monotonicity means that the instrument is shifting outcomes in countries in the same direction; alternatively, in the language of Imbens and Angrist (1994), there are no "defiers".[16] In this case, higher U.S./world resource values for natural resources mostly fuel civil conflict. Ross (2012) points out that there is some causal heterogeneity in the resource curse for wealthy countries such as Canada and Norway, but that is mainly not the case in Africa and the other developing countries in our sample.

Third, the instrument must satisfy the stable-unit treatment value assumption (SUTVA): $Y_i \perp\!\!\!\perp D_j \; \forall \; i \neq j$ and $Y_i = Y_{1i}D_i + Y_{0i}(1 - D_i)$. For SUTVA to hold, units must not interfere with each other, and potential outcomes must be well-defined. One could perhaps argue that mine discoveries in one grid cell could catalyze exploration and discovery of mines in neighboring grid cells. However, any spatial spillovers are prone to time lags given that discoveries and extraction in neighboring grid-cells will not happen immediately. As Menaldo (2016)

[14] The figures in the footnoted sentence reflect grid-cell level correlations. At the spatial point level, log local values correlate with World Bank values at 0.88, and log USGS values correlate with log local values at 0.86.

[15] All first-stage results available with replication files.

[16] Technically, it is possible to have an instrumental variable in which there are only "defiers" and no "compliers", but this is not the norm. For more on the compliers and defiers distinction, refer to Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996).

shows, natural resource extraction requires significant technology, capital, and investment.

3.3.2. Criterion 4: Exclusion Restriction

Fourth, the instrument needs to satify the exclusion restriction: $P(Y_{1d} = Y_{0d}|D) = 1 \in [0,1]$ (Kern and Hainmueller, 2009, 384). Our proposed instrument would violate the exclusion restriction if U.S./world values (Z) are endogenous to local conflict (Y). On that score, very prominent recent studies by Berman et al. (2017) and Christensen (2019) contend that world resource prices are exogenous to local conflict (see also Humphreys, 2010; Carter, Rausser and Smith, 2011; Rossen, 2015). According to these authors, a commodity super-cycle has been in place since roughly 1996. As many countries have become richer and more populous, world demand for minerals has spiked considerably, creating large demand-side shocks that facilitate exogeneity of resource prices to conflict.

Whether these demand-side shocks from the commodity super-cycle are so large as to offset any supply-side incentives of higher resources prices potentially fueling rebel attacks of extraction sites is difficult to test empirically. Nevertheless, in this paper we furnish (to our knowledge) the first evidence to show that natural resource companies spend significant amounts of their resources on preventing rebel attacks (see Appendix B). Rebels are generally thus not able to affect the global price at will. There are significant safeguards in place at industrial mines to avoid rebel-induced interruptions in the flow of minerals onto the world market. In turn, on a process level, local conflicts are insulated from global prices except through the mediation of local prices, so the exclusion restriction holds.

3.3.3. Criterion 5: Independence

The fifth criterion that an instrument must satisfy is the independence or ignorability assumption: $Z_i \perp\!\!\!\perp (Y_{i1}, Y_{i0}, D_{i1}, D_{i0})$. Essentially, the instrument needs to be independent of potential outcomes and the endogenous variable in its different treatment states (Morgan and Winship, 2015, 307). In this case, the independence assumption would not hold if the US/world values (Z) are a function of local conflict (Y) or the local resource values (D). We addressed the potential non-independent relationship between Y and Z in the previous section on the exclusion restriction.

Whether the relationship between Z and D suffers from Betz, Cook and Hollenbach (2018) call "spatial simultaneity" merits further discussion. For our instrument, the local resource values that we calculate from UN Comtrade prices do not constitute any form of an average or aggregate up to the US/world values that we calculate from USGS and the World Bank—and, in some cases, Multicolour (see above). In fact, none of these datasets come from the same distribution. USGS prices correspond to US resource values, which are outside our sample. Despite the literature's ubiquitous use of the world prices from the World Bank (e.g. Berman et al., 2017), the latter institution mostly draws their price data from OECD countries outside our sample (World Bank, 2018). Accordingly, our instrument does not suffer from the same concerns as the spatial averages that Betz, Cook and Hollenbach (2019) critique at length.

Betz, Cook and Hollenbach (2019) further raise the issue of spatial interdependence among outcome variables. In order to control for the possibility of spillover effects among outcome variables in neighboring units, they recommend the use of spatial two-stage least squares (S-2SLS). The latter creates a first-stage equation to predict outcome variables in neighboring cells, and it then uses the predicted values in the second-stage equation. Much of what the S-2SLS model accomplishes in practical terms is the creation of a spatial weights matrix in order to perform the two-stage equation. However, S-2SLS does not lend itself to panel data.

To address this issue, in the creation of this data set, we constructed a series of spatial weights matrices for each year of the data. After the construction of each year's spatial weights matrix, we simply appended the data from each year to produce time-series data that also contained spatially lagged variables. This simple work around allows the creation of both spatial and time-series lags, and so we included a spatially ***and*** temporally lagged dependent variable of conflict on the right-hand side of the equation.

Noticeably, the above procedure skips the first-stage of S-2SLS, but we posit this has some advantages. First, whereas S-2SLS uses predicted values from neighboring cells, we use the actual values of conflict in the neighboring cell that are both spatially and temporally lagged. This has the advantage of more realistically modeling diffusion and avoids simultaneity. Second, a predicted value from a neighboring cell relies on good model fit for an accurate prediction. Even if the prediction model is well-fit, the predicted value's relationship to the actual value should be unbiased. Thus, the use of the actual value would produce similar results to the use of predicted values. If the prediction equation is not well-fit, then the use of actual values will create results that are more accurate than biased results from a poorly fit predicted value. In some cases, the use of actual values may even be an overly conservative test for our primary independent variables, as the first-stage value may under-predict conflict, because of poor model fit. Thus, the use of actual values for temporally and spatially lagged dependent variables on the right-hand side appears an appropriate solution to the concerns about spatial interdependence.

4. Results: Natural Resource Values and Civil Conflict

We proceed by reporting the results in a series of steps. To compare with past studies, we begin by reporting the analysis for Sub-Saharan Africa when using the ACLED measure as our dependent variable. (See Table 2.) We first report the results using the local resources values without and with controls (Models 1 and 2 respectively) and then using the instrumented local price variable without and with controls (Models 3 and 4 respectively). Continuing with the ACLED conflict measure, we then expand the analysis to include the entire African continent, including North Africa also using the ACLED dependent variable (Table 4). The results of all of these analyses are strong and show that natural resources are positively associated with the incidence of conflict, a result that is consistent with past studies, especially the comprehensive Berman et al. (2017) study.

Table 2: Main Spatial HAC Model Results for ACLED Outcome on SSA (Three-Way Fixed Effects)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Natural Resource Value in Cell (Time Lag/Log)	0.0046***	0.0029***		
	(0.0004)	(0.0004)		
Resources 1st Order Spatial Lag		0.0007***		0.0002
		(0.0002)		(0.0002)
Resources 2nd Order Spatial Lag		0.0001		0.0003**
		(0.0001)		(0.0001)
Presence of Lootable Resources		0.0191*		
		(0.0106)		
Number of Excluded Ethnic Groups		-0.0105***		-0.0065**
		(0.0036)		(0.0031)
Nighttime Lights		-0.8882***		0.6385***
		(0.1623)		(0.0666)
V-Dem Democracy Index		-0.4534		
		(181.3453)		
Spatially Lagged Conflict Measure		0.0417***		0.1021***
		(0.0023)		(0.0030)
Natural Resource Value w/ Instrumented Local Price			0.0178***	0.0093***
			(0.0013)	(0.0012)
Constant			0.0652***	0.0153***
			(0.0003)	(0.0028)
Observations	195128	170493	195128	170493
R^2	0.003	0.007		

Standard errors in parentheses

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

We are interested in the broader effects of natural resources on conflict, and because we have natural resource data for the entire Middle East as well as much of Asia and Latin America, we conduct a series of analyses slowly expanding out geographically. Unfortunately the ACLED data is not available for most countries outside of Africa for a sufficiently long time period. As such, we need to shift to a different measure for armed conflict that is available more broadly, and accordingly, we use data from the UCDP's Georeferenced Event Dataset (GED). For comparability with the ACLED models, we re-estimate the Sub-Saharan Africa results using UCDP and report those results in Table 3 (compare to ACLED results in Table 2), and then extend out in successive analyses capturing Sub-Saharan Africa and North Africa (See Table 5 and compare to Table 4).

What is clear from these analyses using the UCDP measure is that the results found with the ACLED measure are no longer straightforward. In the main models, there are either null or negative relationships. And for the instrumented models, the results are inconsistent with the results negative in the models without controls, but positive for Sub-Saharan Africa using the instrumented local price model. Thus one out of eight models match the results from the ACLED models, although arguably that one model is the most critical model to match. As we have argued, the instrumented version is likely to be the most accurate specification.

Everything about the setup of these models is identical to the earlier models save for the different operationalization of conflict, suggesting that natural resources may only robustly predict certain types of conflict but not others. There are a number of key differences between ACLED and UCDP that largely reflect differences in scope, such as ACLED capturing a wider variety of violent and non-violent events with and without casualties whereas UCDP is confined to fatality-producing violent events Eck (2012), though there is often much overlapping information as well (Donnay et al., 2019). In sum, the results of these models with UCDP do not provide a robust story, though the instrumented SSA model with controls is consistent, which is an important comparison point (See Table 3).

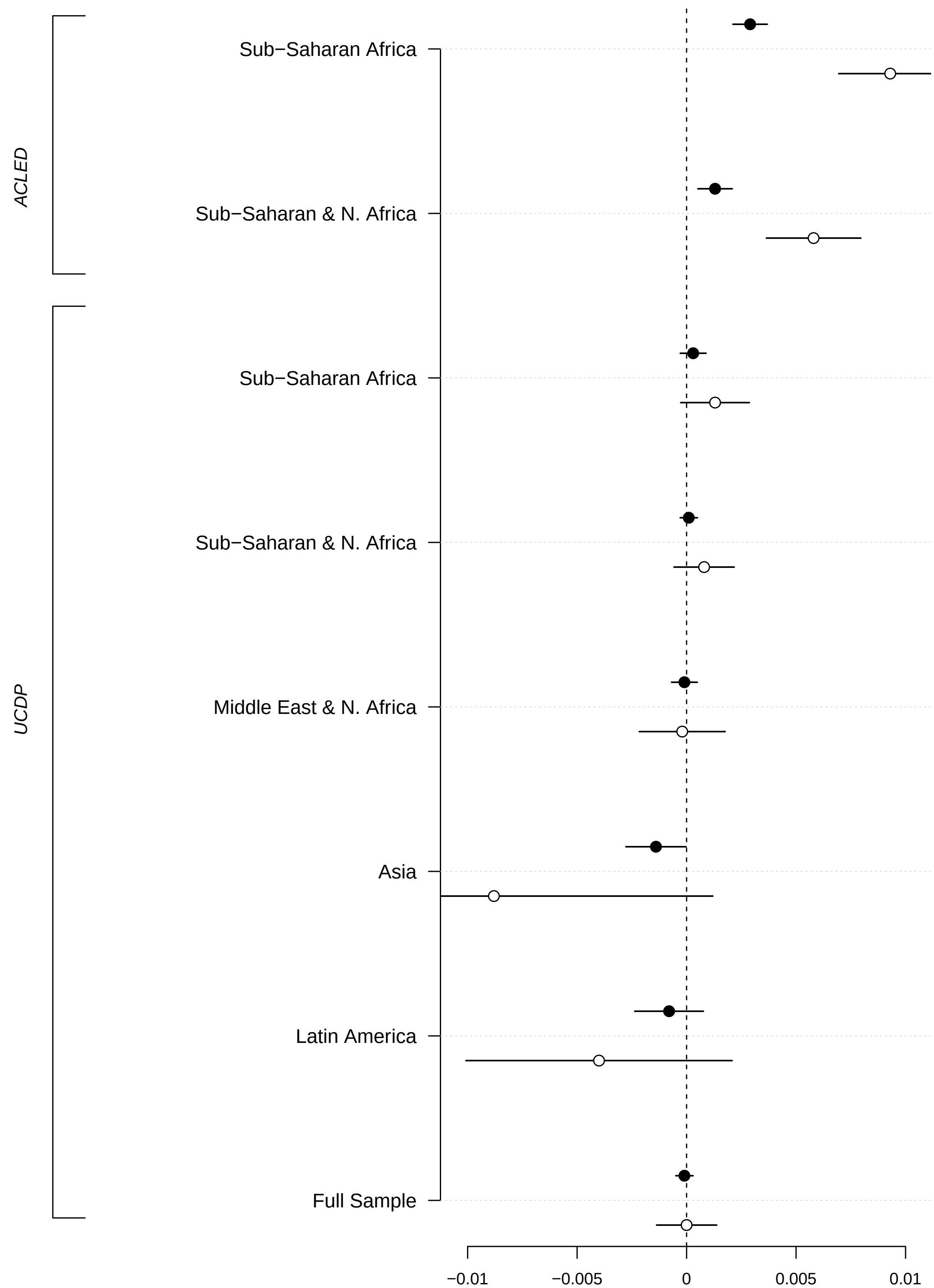

Figure 5: Results Across All Models. Solid dots represent base regression models with controls, but no instruments. The hollow dots represent the instrumental variable models.

Table 3: Main Spatial HAC Model Results for UCDP Outcome on SSA (Three-Way Fixed Effects)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Natural Resource Value in Cell (Time Lag/Log)	-0.0004	0.0003		
	(0.0003)	(0.0003)		
Resources 1st Order Spatial Lag		0.0000		-0.0000
		(0.0001)		(0.0002)
Resources 2nd Order Spatial Lag		-0.0002		-0.0002*
		(0.0001)		(0.0001)
Presence of Lootable Resources		0.0103		
		(0.0083)		
Number of Excluded Ethnic Groups		0.0144***		0.0207***
		(0.0040)		(0.0034)
Nighttime Lights		0.1461		-0.0092
		(0.2217)		(0.0510)
V-Dem Democracy Index		-4.6966		
		(777.1070)		
Mean Population Density		0.0001		0.0000
		(0.0001)		(0.0001)
Spatially Lagged Conflict Measure		0.0163***		0.0279***
		(0.0019)		(0.0020)
Natural Resource Value w/ Instrumented Local Price			-0.0022***	0.0013*
			(0.0008)	(0.0008)
Constant			0.0410***	0.0237***
			(0.0002)	(0.0043)
Observations	195128	104380	195128	104380
R^2	0.000	0.002		

Standard errors in parentheses

Once we turn to the remaining UCDP models outside of the African context, the overall story becomes even more muddled. We now consider whether the results are similar (no effect) outside of Sub-Saharan African and North Africa when using the UCDP measure. Middle East and North Africa (See Table 6), Asia (See Table 7), Latin America (See Table 8), and then all countries globally that we have coded thus far (See Table 9). For purposes of interpretation, recall that we have coded the African continent and the Middle East in their entirety, but have only coded a non-random set of countries in Asia and Latin America.

Moving to the MENA region as well as the entire sample (also Table 6), all of which rely on the UCDP measure given ACLED is not available, resources are not significantly associated with conflict incidence, or in some cases are negatively associated. The story is similar for Asia and Latin America (Tables 7 and 8). When pooling across all regions, the results remain mixed at best. In all of these additional models, resources are never positively associated with conflict, suggesting important limitations to the emerging narrative tying resources to conflict.

5. Conclusion

In this paper, we report on a new data set of 183 natural resources, geo-referenced across 103 countries. While the natural resource data could be used for many purposes, we used it here to examine its relationship to conflict. We carried out a basic set of models connecting natural resource value (using world prices) to conflict and show that in some cases, natural resources are positively correlated, though the result does not carry over to other regions and indeed changes based on whether one uses the ACLED or GED measures. We then shifted to calculating natural resource value with local price data, instrumented with U.S. and world prices, in order to address endogeneity concerns. These results indicate for the ACLED outcome, but not the GED outcome, natural resources strongly and positively predict violence. The differences across these regions and measures of conflict suggest that the relationship between local resource wealth and conflict is more complicated and heterogeneous than existing research, such as Berman et al. (2017), would suggest. Future research could profitably focus on developing more fine-grained explanations of the contextual factors that seem to lead to positive relationships between resource wealth and conflict at the local level.

While our empirical focus here has been on the links between resources and conflict incidence at the local level, the GRD could be used to address many additional research questions by scholars of conflict and of other issues. For conflict researchers, the data might be useful for understanding the intensity of conflict, the type of conflict events (i.e. battles between government and rebel forces or violence against civilians), protests (Christensen, 2019), how changes in prices influence conflict (Dube and Vargas, 2013), where rebel groups originate and establish bases and sanctuaries, human rights abuses by government and rebel forces (Weinstein, 2007), and so on. A partial list of research questions beyond the domain of armed conflict that could be investigated with the GRD includes government capacity at the local level; the incidence of corruption; public goods provision (e.g. health, environmental protection); and voting behavior.

References

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):456–458.

Aragón, Fernando M. and Juan Pablo Rud. 2013. "Natural Resources and Local Communities: Evidence from a Peruvian Gold Mine." *American Economic Journal: Economic Policy* 5(2):1–25.

Asal, Victor, Michael G. Findley, James Piazza and James Igoe Walsh. 2015. "Political Exclusion, Oil, and Ethnic Armed Conflict." *Journal of Con ict Resolution* 53(212):1689–1699.

Balestri, Sara. 2012. "Gold and Civil Conflict Intensity: Evidence from a Spatially Disaggregated Analysis." *Peace Economics, Peace Science and Public Policy* 18(3):1–17.

Balestri, Sara. 2015. "GOLDATA: The Gold Deposits Dataset Codebook, Version 1.2." *Cognitive Science and Communication Research Centre* pp. 1–30.

Balestri, Sara and Mario A. Maggioni. 2014. "Blood Diamonds, Dirty Gold and Spatial Spillovers Measuring Conflict Dynamics in West Africa." *Peace Economics, Peace Science and Public Policy* 20(4):551–564.

Bazzi, Samuel and Christopher Blattman. 2014. "Economic Shocks and Conflict: Evidence from Commodity Prices." *American Economic Journal: Macroeconomics* 6(4):1–38.

Berman, Nicolas and Mathieu Couttenier. 2016. "External Shocks, Internal Shots: The Geography of Civil Conflict." *Review of Economics and Statistics* 98(2):209–225.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner and Mathias Thoenig. 2017. "This Mine is Mine! How Minerals Fuel Conflicts in Africa." *American Economic Review* 107(6):1564–1610.

Betz, Timm, Scott J. Cook and Florian M. Hollenbach. 2018. "On the Use and Abuse of Spatial Instruments." *Political Analysis* 26:474–479.

Betz, Timm, Scott J Cook and Florian M Hollenbach. 2019. "Spatial Interdependence and Instrumental Variable Models." *Political Science Research and Methods* 00(0):1–13.

Booth, Geoffrey, Anthony Cameron, Gerry Fahey and Michael Lawlor. 2010. Anvil Mining Limited Ni 43-101 Technical Report Kinsevere Copper Project Katanga Province. Technical Report March.

Buhaug, Halvard and Päivi Lujala. 2005. "Accounting for Scale: Measuring Geography in Quantitative Studies of Civil War." *Political Geography* 24(4):399–418.

Carter, Colin A., Gordon C. Rausser and Aaron Smith. 2011. "Commodity Booms and Busts." *Annual Review of Resource Economics* 3(1):87–118.

Caselli, Francesco, Massimo Morelli and Dominic Rohner. 2015. "The Geography of Interstate Resource Wars." *Quarterly Journal of Economics* (2012):267–315.

Christensen, Darin. 2019. "Concession Stands: How Mining Investments Incite Protest in Africa." *International Organization* 73(1):65–101.

Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–595.

Conley, Timothy G. 1999. "GMM Estimation with Cross-sectional Dependence." *Journal of Econometrics* 92(1):1–45.

Conrad, Justin M., Kevin T. Greene, James Igoe Walsh and Beth Elise Whitaker. 2019. "Rebel Natural Resource Exploitation and Conflict Duration." *Journal of Con ict Resolution* 63(3):591–616.

Cotet, Anca M. and Kevin K. Tsui. 2013. "Oil and Conflict: What Does the Cross Country Evidence Really Show?" *American Economic Journal: Macroeconomics* 5(1):49–80.

Czernowalow, Martin. 2004. "Anvil to Resume Operations at Dikulushi Mine." *Creamer Media's Engineering News* .

URL: *https://www.engineeringnews.co.za/article/anvil-to-resume-operations-at-dikulushi-mine-2004-10-19/rep*$_id$: 4136

De Koning, Ruben. 2010. "Demilitarizing Mining Areas in the Democratic Republic of the Congo: the Case of Northern." *SIPRI Insights on Peace and Security* 1(January):1–20.

De Soysa, Indra. 2002. "Paradise is a Bazaar? Greed, Creed and Grievance in Civil War, 1989-1999." *Journal of Peace Research* 39(4):295–416.

Diaz-Rioseco, Diego. 2016. "Blessing and Curse: Oil and Subnational Politics in the Argentine Provinces." *Comparative Political Studies* 49(14):1930–1964.

Donnay, Karsten, Eric T. Dunford, Erin C. McGrath, David Backer and David E. Cunningham. 2019. "Integrating Conflict Event Data." *Journal of Con ict Resolution* 63(5):1337–1364.

Dube, Oeindrila and Juan F. Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." *Review of Economic Studies* 80(4):1384–1421.

Eck, Kristine. 2012. "In Data We Trust? A Comparison of UCDP GED and ACLED Conflict Events Datasets." *Cooperation and Con ict* 47(1):124–141.

Elvidge, Christopher D., Feng-Chi Hsu, Kimberly Baugh and Tilottama Ghosh. 2014. *National Trends in Satellite Observed Lighting: 1992-2012*. Boca Raton, Florida: CRC Press.

Esteban, Joan, Massimo Morelli and Dominic Rohner. 2015. "Strategic Mass Killings." *Journal of Political Economy* 123(5).

Fearon, James D. 2004. "Why Do Some Civil Wars Last So Much Longer than Others?" *Journal of Peace Research* 41(3):275–301.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(01):75–90.

Gervasoni, Carlos. 2010. "A Rentier Theory of Subnational Regimes: Fiscal Federalism, Democracy, and Authoritarianism in the Argentine Provinces." *World Politics* 62(2):302–340.

Goldewijk, Kees Klein, Arthur Beusen, Jonathan Doelman and Elke Stehfest. 2017. "Anthropogenic Land Use Estimates for the Holocene - HYDE 3.2." *Earth System Science Data* 9(2):927–953.

Harari, Mariaflavia and Eliana La Ferrara. 2018. "Conflict, Climate, and Cells: A Disaggregated Analysis." *The Review of Economics and Statistics* 100(4):594–608.

Hong, Ji Yeon. 2018. "How Natural Resources Affect Authoritarian Leaders' Provision of Public Services: Evidence from China." *Journal of Politics* 80(1):178–194.

Hsiang, Solomon M. 2010. "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America." *Proceedings of the National Academy of Sciences* 107(35):15367–15372.

Humphreys, David. 2010. "The Great Metals Boom: A Retrospective." *Resources Policy* 35(1):1–13.

Humphreys, Macartan. 2005. "Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms." *Journal of Con ict Resolution* 49(4):508–537.

Imbens, Guido W. and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2):467–475.

Institute of Developing Economies, Japan External Trade Organization. 2019. "Anvil Mining Limited - AGE (Africa Growth Enterprise).".

Jones Luong, Pauline and Erika Weinthal. 2010. *Oil is Not a Curse: Structure and Institutions in Soviet Successor States*. Cambridge: Cambridge University Press.

Kern, Holger Lutz and Jens Hainmueller. 2009. "Opium for the Masses: How Foreign Media can Stabilize Authoritarian Regimes." *Political Analysis* 17(4):377–399.

Koubi, Vally, Gabriele Spilker, Tobias Böhmelt and Thomas Bernauer. 2014. "Do Natural Resources Matter for Interstate and Intrastate Armed Conflict?" *Journal of Peace Research* 51(2):227–243.

Lei, Yu Hsiang and Guy Michaels. 2014. "Do Giant Oilfield Discoveries Fuel Internal Armed Conflicts?" *Journal of Development Economics* 110:139–157.

Lujala, Päivi. 2009. "Deadly Combat over Natural Resources: Gems, Petroleum, Drugs, and the Severity of Armed Conflict." *Journal of Con ict Resolution* 53(1):50–71.

Lujala, Päivi. 2010. "The Spoils of Nature: Armed Civil Conflict and Rebel Access to Natural Resources." *Journal of Peace Research* 47(1):15–28.

Lujala, Päivi, Jan Ketil Röd and Nadja Thieme. 2007. "Fighting over Oil: Introducing a New Dataset." *Con ict Management and Peace Science* 24(3):239–256.

Lujala, Päivi, Nils Petter Gleditsch and Elisabeth Gilmore. 2005. "A Diamond Curse? Civil War and a Lootable Resource." *Journal of Con ict Resolution* 49(4):538–562.

Mähler, Annegret and Jan H. Pierskalla. 2015. "Indigenous Identity, Natural Resources, and Contentious Politics in Bolivia: A Disaggregated Conflict Analysis, 2000- 2011." *Comparative Political Studies* 48(483):301–332.

Matos, G.R. 2015. "Historical Global Statistics for Mineral and Material Commodities (2015 version)." *U.S. Geological Survey Data Series 896* .

Maystadt, Jean François, Giacomo De Luca, Petros G. Sekeris and John Ulimwengu. 2014. "Mineral Resources and Conflicts in DRC: A Case of Ecological Fallacy?" *Oxford Economic Papers* 66(3):721–749.

Menaldo, Victor. 2016. *The Institutions Curse: Natural Resources, Politics, and Development*. Cambridge: Cambridge University Press.

Morgan, Stephen L. and Christopher Winship. 2015. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge University Press.

Nillesen, Eleonora E. and Erwin H. Bulte. 2014. "Natural Resources and Violent Conflict." *Annual Review of Resource Economics* 6:69–83.

O'Brochta, William. 2019. "A Meta-Analysis of Natural Resources and Conflict." *Research & Politics* 6(1):1–6.

Paine, Jack. 2016. "Rethinking the Conflict Resource Curse: How Oil Wealth Prevents Center-Seeking Civil Wars." *International Organization* 70(4):727–761.

Raleigh, Clionadh, Andrew Linke, Håvard Hegre and Joakim Karlsen. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset." *Journal of Peace Research* 47(5):651–660.

Rights and Accountability in Development and Association Africaines de défense des droits de L'Homme Action Contre l'Impunité pour les Droits Humains. 2005. Anvil Mining Limited and the Kilwa Incident Unanswered Questions. Technical report Rights and Accountability in Development Action Contre l'Impunité pour les Droits Humains, Association Africaines de défense des droits de L'Homme Mobutu, Democratic Republic of the Congo: .

Ross, Michael L. 2004. "What Do We Know about Natural Resources and Civil War?" *Journal of Peace Research* 41(3):337–356.

Ross, Michael L. 2006. "A Closer Look at Oil, Diamonds, and Civil War." *Annual Review of Political Science* 9:265–300.

Ross, Michael L. 2012. *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton, New Jersey: Princeton University Press.

Ross, Michael L. 2015. "What Have We Learned about the Resource Curse?" *Annual Review of Political Science* 18(1):239–259.

Rossen, Anja. 2015. "What Are Metal Prices Like? Co-movement, Price Cycles and Long-Run Trends." *Resources Policy* 45:255–276.

Sanchez de la Sierra, Raul. 2019. "On the Origin of States: Stationary Bandits and Taxation in Eastern Congo." *Journal of Political Economy* .

Sundberg, Ralph and Erik Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset." *Journal of Peace Research* 50(4):523–532.

Tollefsen, Andreas Forø, Håvard Strand and Halvard Buhaug. 2012. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research* 49(2):363–374.

United Nations Statistics Division. 2018. "Comtrade Database.".

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Con ict Resolution* 59(7):1327–1342.

Walsh, James Igoe, Justin M Conrad, Beth Elise Whitaker and Katelin M Hudak. 2018. "Funding Rebellion: The Rebel Contraband Dataset." *Journal of Peace Research* 55(5):699–707.

Weinstein, Jeremy M. 2007. *Inside Rebellion: The Politics of Insurgent Violence*. New York: Cambridge University Press.

World Bank. 2018. "Commodity Markets.".

A. Additional Results

Table 4: Main Spatial HAC Model Results for ACLED Outcome on SSA and NA (Three-Way Fixed Effects)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Natural Resource Value in Cell (Time Lag/Log)	0.0055***	0.0013***		
	(0.0003)	(0.0004)		
Resources 1st Order Spatial Lag		0.0002		0.0001
		(0.0002)		(0.0002)
Resources 2nd Order Spatial Lag		-0.0000		-0.0003**
		(0.0001)		(0.0001)
Presence of Lootable Resources		0.0009		
		(0.0113)		
Number of Excluded Ethnic Groups		0.0087*		0.0109***
		(0.0047)		(0.0039)
Nighttime Lights		-0.9805***		0.6034***
		(0.1988)		(0.0564)
V-Dem Democracy Index		4.0858		
		(754.9359)		
Mean Population Density		0.0002**		0.0002**
		(0.0001)		(0.0001)
Spatially Lagged Conflict Measure		0.0168***		0.0368***
		(0.0022)		(0.0023)
Natural Resource Value w/ Instrumented Local Price			0.0191***	0.0058***
			(0.0011)	(0.0011)
Constant			0.0556***	0.0237***
			(0.0003)	(0.0042)
Observations	250824	134345	250824	134345
R^2	0.005	0.002		

Standard errors in parentheses

Table 5: Main Spatial HAC Model Results for UCDP Outcome on SSA and NA (Three-Way Fixed Effects)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Natural Resource Value in Cell (Time Lag/Log)	-0.0005**	0.0001		
	(0.0002)	(0.0002)		
Resources 1st Order Spatial Lag		-0.0001		-0.0001
		(0.0001)		(0.0001)
Resources 2nd Order Spatial Lag		-0.0002*		-0.0002
		(0.0001)		(0.0001)
Presence of Lootable Resources		0.0122		
		(0.0075)		
Number of Excluded Ethnic Groups		0.0147***		0.0206***
		(0.0039)		(0.0034)
Nighttime Lights		-0.1453		-0.0588
		(0.1590)		(0.0418)
V-Dem Democracy Index		-3.7023		
		(777.3183)		
Mean Population Density		0.0000		0.0000
		(0.0001)		(0.0001)
Spatially Lagged Conflict Measure		0.0148***		0.0246***
		(0.0017)		(0.0018)
Natural Resource Value w/ Instrumented Local Price			-0.0020***	0.0008
			(0.0006)	(0.0007)
Constant			0.0375***	0.0254***
			(0.0002)	(0.0030)
Observations	250824	134345	250824	134345
R^2	0.000	0.001		

Standard errors in parentheses

Table 6: Main Spatial HAC Model Results for UCDP Outcome on MENA (Three-Way Fixed Effects)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Natural Resource Value in Cell (Time Lag/Log)	0.0001	-0.0001		
	(0.0003)	(0.0003)		
Resources 1st Order Spatial Lag		-0.0002		-0.0001
		(0.0002)		(0.0002)
Resources 2nd Order Spatial Lag		-0.0000		0.0005***
		(0.0001)		(0.0001)
Presence of Lootable Resources		-0.0050		
		(0.0097)		
Number of Excluded Ethnic Groups		0.1298***		-0.0327
		(0.0257)		(0.0216)
Nighttime Lights		-0.0172		0.5821***
		(0.1081)		(0.0675)
Mean Population Density		0.0001		0.0001**
		(0.0000)		(0.0001)
Spatially Lagged Conflict Measure		0.0053		0.0019
		(0.0036)		(0.0036)
Natural Resource Value w/ Instrumented Local Price			0.0008	-0.0002
			(0.0009)	(0.0010)
Constant			0.0481***	0.0116
			(0.0004)	(0.0084)
Observations	118864	61646	118864	63911
R^2	0.000	0.006		

Standard errors in parentheses

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

Table 7: Main Spatial HAC Model Results for UCDP Outcome on Asia (Three-Way Fixed Effects)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Natural Resource Value in Cell (Time Lag/Log)	-0.0010**	-0.0014**		
	(0.0004)	(0.0007)		
Resources 1st Order Spatial Lag		-0.0013**		-0.0009
		(0.0006)		(0.0007)
Resources 2nd Order Spatial Lag		-0.0000		0.0006
		(0.0005)		(0.0005)
Presence of Lootable Resources		0.0019		
		(0.0141)		
Number of Excluded Ethnic Groups		0.0737***		-0.0686***
		(0.0152)		(0.0130)
Nighttime Lights		-0.0098		-0.0138
		(0.0697)		(0.0448)
Mean Population Density		0.0000		0.0001
		(0.0001)		(0.0001)
Spatially Lagged Conflict Measure		0.0097		-0.0095
		(0.0110)		(0.0083)
Natural Resource Value w/ Instrumented Local Price			-0.0041	-0.0088*
			(0.0027)	(0.0051)
Constant			0.0351***	0.0676***
			(0.0002)	(0.0132)
Observations	150144	80678	150144	80678
R^2	0.000	0.002		

Standard errors in parentheses

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

Table 8: Main Spatial HAC Model Results for UCDP Outcome on Latin America (Three-Way Fixed Effects)

	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 3	Model 4
Natural Resource Value in Cell (Time Lag/Log)	-0.0005*	-0.0008		
	(0.0003)	(0.0008)		
Resources 1st Order Spatial Lag		-0.0013***		-0.0013***
		(0.0004)		(0.0004)
Resources 2nd Order Spatial Lag		-0.0004		-0.0001
		(0.0002)		(0.0003)
Presence of Lootable Resources		0.0072		
		(0.0130)		
Number of Excluded Ethnic Groups		0.0013		-0.0022
		(0.0012)		(0.0016)
Nighttime Lights		0.5289***		-0.0870**
		(0.1581)		(0.0354)
Mean Population Density		-0.0002*		-0.0003**
		(0.0001)		(0.0001)
Natural Resource Value w/ Instrumented Local Price			-0.0023**	-0.0040
			(0.0010)	(0.0031)
Spatially Lagged Conflict Measure				0.0000
				(.)
Constant			0.0201***	0.0433***
			(0.0002)	(0.0051)
Observations	190128	101049	190128	101166
R^2	0.000	0.002		

Standard errors in parentheses

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

Table 9: 2SLS Instrumented Local Prices Model Results for UCDP Outcome on Full Sample (Three-Way Fixed Effects)

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Natural Resource Value in Cell (Time Lag/Log)	-0.0003**	-0.0001		
	(0.0001)	(0.0002)		
Resources 1st Order Spatial Lag		-0.0003***		-0.0002
		(0.0001)		(0.0001)
Resources 2nd Order Spatial Lag		-0.0001		0.0002*
		(0.0001)		(0.0001)
Presence of Lootable Resources		0.0068		
		(0.0052)		
Number of Excluded Ethnic Groups		0.0259***		0.0005
		(0.0035)		(0.0033)
Nighttime Lights		0.0481		0.0510***
		(0.0296)		(0.0154)
V-Dem Democracy Index		-0.4654		
		(78.1578)		
Mean Population Density		0.0000		0.0000
		(0.0000)		(0.0000)
Spatially Lagged Conflict Measure		0.0147***		0.0209***
		(0.0017)		(0.0019)
Natural Resource Value w/ Instrumented Local Price			-0.0012**	-0.0000
			(0.0005)	(0.0007)
Constant			0.0236***	0.0180***
			(0.0001)	(0.0024)
Observations	988968	520796	988968	523178
R^2	0.000	0.001		

Standard errors in parentheses

B. DRC Case Study of Exclusion Restriction

The case of mining operations in the Democratic Republic of Congo (DRC) provides plentiful evidence that mining companies devote significant resources to protecting mines from outside forces. The most direct evidence comes from a mining company called Anvil Mining Limited. In 2009, it operated the Kinsevere Copper Project in Katanga Province of the DRC. During this time period, the company spent roughly $158,000 per month on direct security costs for the Kinsevere site alone (Booth et al., 2010). The Kinsevere Mine is a relatively average mining site in the DRC with an annual value of roughly $366 million per year, compared to the average location across all observations in the DRC of $365 million. It also has only a slightly higher annual output than the mean of all mines within the DRC. As such, it represents a typical mining location, and the mining company spent almost $2 million a year on direct site security for the Kinsevere site alone.

These direct costs are also only part of the broader picture of mine security costs. Mine security in the DRC is a complex issue that involves numerous government agencies, with side payments and informal agreements between the mining company and armed groups—both government and rebel. For instance, Anvil Mining was also reported to pay roughly $5,000 per month to local administrative and security officials to maintain their support in the area around the Dikulushi Mine north of Kilwa (Rights and Accountability in Development and Action Contre l'Impunité pour les Droits Humains, 2005). The same report indicates that informants claimed local administrators and sector chiefs each received roughly $420 per month. All of these payments stand in addition to the existing repatriation agreement, where the company repatriates 40% of proceeds from the mine site for use by the DRC government. For this mine site alone, that agreement amounted to the repatriation of $76 million in 2008 (Institute of Developing Economies, 2019). Other reports indicate that, while the central government agrees to provide security in return for a share of the mining profits, local officials do the same. At times, tacit agreements are formed with local commanders or even individual soldiers in return for the provision of security (De Koning, 2010).

In addition, there are tacit agreements at mine sites, which allow local authorities to use company security equipment when they need it for security purposes. In one instance, local authorities used mine security equipment to raid a local town that was supposedly harboring rebels. This shows that the mine site was heavily armed and prepared to defend against rebel groups. In fact it was even more heavily armed than government forces in the area, and so heavily armed that it was used a repository for those local security officials to conduct offensive operations against neighboring rebel groups. Furthermore, the company paid for the stationing of DRC troops and army intelligence at the mine site itself as a protective measure. It was only after the incident that the company requested additional security forces from the government in order to prevent the need for local security forces to requisition equipment from the company (Czernowalow, 2004).

All of these items indicate that mine security is taken very seriously across even medium-sized sites of average value, to prevent disruptions in the supply of raw materials to the world market. Because companies are determined to protect their resources through the direct provision of security and through explicit and implicit agreements with local officials, local prices of the resource at the mine site are unlikely to see significant shocks. Rather, what we generally see are steady operations at industrial sites that occasionally shut down for technical issues, which affects local prices but not global prices.

Furthermore, it is worth exploring the idea that global prices influence conflict on their own without the mediation of local prices. This is unlikely for a variety of reasons, but the main issue is that many minerals require a significant investment in infrastructure for them to be taken to the world market. They must enter the global market in order to be incorporated into supply chains and the process of adding value through conversion, transformation, refinement, or combining with other elements to produce finished products.

For instance, in the case of the same mine, the Dikulushi Mine in Southeast DRC, the minerals extracted are copper and silver. In order to bring these minerals to market, they must first be refined and finished. The company built pontoon ferries across 27 miles of Lake Mweru and then drive another 1,600 miles to a company processing facility in Namibia for refining. From there, the processed product would then need to be transported to an international port for loading onto ships and transport to facilities that apply further manufacturing techniques in Europe and Asia.

Rebels have very little ability to apply this process on their own, and even looted resources must be sold at local prices for them to be taken into the global market by others. Due to the technical nature of extraction and the need for significant infrastructure to transport many minerals to a point of sale, it is highly unlikely that rebels would ever be able to realize a world price rather than a local price.

Thus, because of the nature of many minerals—both their need for further value-added and the necessity of large-scale infrastructure on the ground in order to realize any value, local conflict is relatively insulated from world prices. Since companies that do the mining also expend significant time and money guarding the resource sites, local conflicts are insulated from both supply- and demand-side shocks from the global market. Therefore, the instrument meets the exclusion restriction.

C. Codebook

C.1. Overview

The purpose of this project is to geolocate all critical mineral resource extraction locations worldwide. Data are available at a subnational level on a country-by-country basis and therefore required coding along these lines. We use the cleaned data to gain inference on the larger concepts of conflict and governance.

C.2. Coding process

The unit of observation of interest for this coding procedure is the mine or resource extraction facility. The data are available for each country worldwide from the United States Geological Survey (USGS) website.

We undertook a number of safeguards to ensure high quality data. First, we undertook an initial round of coding. Next, especially since geolocations are not always clear with higher level precision codes, we undertook a second round of coding to check all of the entries for accuracy. At the end of the second round of coding, the coders randomly sampled each other's work and performed some triple-checks. In the third round of coding, coders performed an initial coding of each location-year, with another coder double-checking over each coded entry. Senior coders also performed spot checks throughout and adjudicated all difficult cases that were not initially clear from the PDF documents produced by the United States Geological Survey (USGS). After the second and third rounds of coding, we further examined instances in which the same location was given different latitudes and longitudes for different location-years. Accordingly, an expert coder then re-checked those locations and assigned a final latitude and longitude to them ex post.

C.3. Variables

This section outlines the variables in the dataset.

C.3.1. year

This information is taken from United States Geological Survey (USGS). Years range from 1994–2014. Data availability varies by country.

C.3.2. resource

This information is taken from United States Geological Survey (USGS). Details on the individuals resources covered in this dataset are found in Table 13. In total, there 183 different resources in the dataset.

C.3.3. country

The country information comes from the United States Geological Survey (USGS). The dataset includes the following countries: Afghanistan, Algeria, Angola, Argentina, Bahrain, Bangladesh, Belize, Benin, Bhutan, Bolivia, Botswana, Brazil, Burkina Faso, Burundi, Cape Verde, Cambodia, Cameroon, Chad, Chile, China, Colombia, Costa Rica, Cote D'Ivoire (Ivory Coast), Cuba, Democratic Republic of Congo, Djibouti, Dominican Republic, Ecuador, Egypt, Equatorial Guinea, Eritrea, Ethiopia, French Guiana, Gabon, Ghana, Guatemala, Guinea, Guyana, Honduras, India, Iran, Iraq, Israel, Jordan, Kazahkstan, Kenya, Kuwait, Kyrgyzstan, Lebanon, Lesotho, Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mexico, Morocco, Mozambique, Myanmar (Burma), Namibia, Nepal, Nicaragua, Niger, Nigeria, Oman, Panama, Paraguay, Pakistan, Peru, Qatar, Republic of Congo, Reunion, Russia, Rwanda, Saudi Arabia, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, South Sudan, Sudan, Sri Lanka, Swaziland, Syria, Taiwan, Tajikistan, Tanzania, Thailand, Togo, Tunisia, Turkey, Uganda, United Arab Emirates, Venezuela, Vietnam, Western Sahara, Yemen, Zaire, Zambia, and Zimbabwe.

C.3.4. region_wb

This variable corresponds to World Bank region of the mine location or resource extraction site. There are five regions in the dataset: (Subsaharan) Africa; Middle East and North Africa; Latin America and Caribbean; South Asia; and East Asia and Pacific.

C.3.5. continent

This variable corresponds to the continent of the mine location or resource extraction site. The dataset contains observations from Asia; the Americas (South and Central America); and Africa.

C.3.6. standard_measure

This information is taken from United States Geological Survey (USGS). Data are recorded using the following units: 42-gallon barrels, 42-gallon barrels per day, billion cubic meters, carats, cubic meters, kilograms, metric tons, metric tons per day, million 42-gallon barrels, million bricks, million cubic meters, million cubic meters per day, million metric tons, square meters, thousand 41-gallon barrels, thousand 41-gallon barrels, thousand 42-gallon barrels per day, thousand 42-gallon barrels per day, thousand bricks, thousand carats, thousand cubic meters, thousand metric tons, and thousand square meters.

C.3.7. comtrade_unit

This information is taken from UN Comtrade. It describes the unit measure for the respective UN Comtrade prices. Prices are expressed in carats, cubic meters, kilograms, and liters.

C.3.8. wb_unit

This information is taken from the World Bank's Global Economic Monitor. The variable describes the unit corresponding to the world price of the respective mineral or resource. Prices are expressed in 42-gallon barrels, metric tons, troy ounces, and mmbtu.

C.3.9. usgs_unit

This information is taken from the United States Geological Survey (USGS). The variable describes the unit corresponding to the US prices of the respective mineral or resource. Prices are expressed in metric tons.

C.3.10. multicolour_unit

This is information is taken from Multicolour. The variable describes the unit corresponding to the world price of the respective mineral or resource. All Multicolour prices are given in carats. For more inquiries on Multicolour prices, please contact David Weinberg at Multicolour: info@multicolour.com.

C.3.11. APIforoil

Table 10: API Gravity to Density Conversions

API Gravity Measure	**Corresponding Density (kg/m^3)**
20	933.993
25	904.152
30	876.161
35	849.850
40	825.073
45	800.8

This information refers to the American Petroleum Institute (API) gravity measure for oil/petroleum or products thereof. It is the industry standard for expressing density, as compared to the density of water. Higher API gravities entail lower densities, which in turn return higher prices on commodity spot markets. When oil has a lower API gravity/higher density, yielding a heavier 42-gallon oil barrel/drum, it requires additional processing steps to make the oil usable.

Table 10 provides the densities in kg/m^3 corresponding to the API gravity measures for a sample of API gravities used in this dataset. The data availability for API gravity based on oil field assays is limited. Thus, when we were unable to find the API gravity each oil field, we approximated the API gravities by country based on information here, here, here, here, here, other websites, and:

Awadh, Salih Muhammed, and HebaSadoon Al-Mimar. 2013. "Statistical Analysis of the Relations between API, Specific Gravity, and Sulfur Content in the Universal Crude Oil." *International Journal of Science and Research* 4(5): 1279-1284.

C.3.12. SGforoil

This variable pertains to the specific gravity of oil/petroleum and products thereof. The specific gravity can be calculated as follows:

$$Specific\ Gravity = 141.5/(131.5 + API\ Gravity)$$

C.3.13. density

This information refers to the density of variables for which output data is expressed in terms of mass but price data is given in volume or heat content —or vice-versa. Table 11 provides the relevant densities (kg/m^3) used in this dataset. Note that densities are only relevant when converting between mass, volume, or heat content units.

C.3.14. heat_content

This variable describes the heat content of certain resources in MMBtu/bbl. Refer to Table 12 for the resource for which it was necessary to have heat content information due to conversions between mass, volume, and heat content units. Heat contents by resource can be found on the website of the Society for Petroleum Engineers.

Table 11: Density by Resource

Resource	Corresponding Density (kg/m^3)
clay (bricks)	1900
gasoline	719.7
granite	2075
helium	147
limestone	2360
liquified petroleum gas	550
liquified natural gas	450
marble	2700
natural gas	0.8
oil	see Table 10
salt	1025
stone	2515

Table 12: Heat Content by Resource

Resource	Heat Content (MMBtu/bbl)
liquified natural gas	3.735
natural gas	3.735
oil/petroleum	5.8
petrochemicals	5.976
petroleum products	5.976

C.3.15. specific_surface_area

This variable corresponds to the specific surface area of stone, sandstone, granite, and marble in meters^2/grams. This variable is necessary for these minerals because USGS annual allocation capacity figures are expressed in square meters. We obtained data from the following resources:

- Keppert, Martin, Jaromir Zumar, Monika Cachova, Dana Konakova, Petr Svora, Zbysek Pavlik, Eva Vejmelkova, and Robert Cerny. 2016. "Water Vapor Diffusion and Adsorption of Sandstones." *Advances in Materials Science and Engineering* (2016). DOI:10.1155/2016/8039748
- Ticknor, Kenneth V., and Preet P.S. Saluja. 1990. "Determination of Surface Areas of Mineral Powders By Adsorption Capacity" *Clays and Clay Minerals* (38)4: 437-441.

C.3.16. locationname

This information is taken from United States Geological Survey (USGS). The location information describes the closest available city, town, or point of interest to the mine or resource extraction site.

C.3.17. minetype

This information comes from United States Geological Survey (USGS). The following different types of mines are available in the data: artisanal, artisanal/military, cooperative, cooperative/industrial, industrial, industrial/government, and government. When ownership information is not available, it has been listed as "n/a". The mixed categories with more than one type of owner are for instances in which there is more than one owner and neither owns a majority stake (i.e. greater than 50%). When any one of the above owns more than a 50% stake, it is classified as only one of the above categories.

C.3.18. admin1

This information is taken from GeoNames (www.geonames.org) on the basis of the location name from USGS. This information corresponds to the administrative level 1 precision code. Generally, it corresponds to a province/department/state.

C.3.19. admin2

This information is taken from GeoNames (www.geonames.org) or Google Maps on the basis of the location name from USGS. This information corresponds to the administrative level 2 precision code. Generally, it corresponds to a district/municipality.

C.3.20. latitude

This information is taken from GeoNames (www.geonames.org) or Google Maps on the basis of the location name from USGS. In instances where there are multiple location names that match the USGS description, the coder arbitrates between the locations given clues on the USGS document, such as province information given by USGS. Further, geonames provides aerial shots of the location, which can be used to pinpoint a probable mine location.

C.3.21. longtitude

This information is taken from GeoNames (www.geonames.org) or Google Maps on the basis of the location name from USGS. In instances where there are multiple location names that match the USGS description, the coder arbitrates between the locations given clues on the USGS document, such as province information given by USGS. Further, geonames provides aerial shots of the location, which can be used to pinpoint a probable mine location.

C.3.22. precisioncode

This information is derived from GeoNames (www.geonames.org) or Google Maps on the basis of the location name from USGS. We use the following precision codes:

- 1: Mine/production facility itself
- 2: Nearby city
- 3: District level
- 4: Province
- 9: Unsure if location is correct

C.3.23. comtrade_price_mult

This variables corresponds to the UN Comtrade export price of the resource, expressed in its standard measure output unit (see above). Thus, prices are available for specific resources and years but also each respective country. All prices are deflated to represent their 2010 United States dollar value. To access the deflators, refer to the World Bank's World Development Indicators.

C.3.24. wb_price_mult

This variables corresponds to the World Bank price for the resource, expressed in its standard measure unit (see above). All prices, which are world prices, are deflated to represent their 2010 United States dollar value. To access the deflators, refer to the World Bank's World Development Indicators.

C.3.25. usgs_price_mult

This variables corresponds to the USGS for the resource, expressed in its standard measure unit (see above). All prices, which are world prices, are deflated to represent their 2010 United States dollar value. To access the deflators, refer to the World Bank's World Development Indicators.

Kindly also note the following:

1. We merge antimony and antimony ore into one antimony price variable. There are few antimony ore observations in our dataset, and pure antimony is a very rare in occurrence. So, it is logical to use one price for antimony.
2. We merge boron and boron refined concentrates into one boron price. There are few boron observations in the dataset.

C.3.26. multicolour_price_mult

This variable corresponds to the Multicolour price for the resource, expressed in its standard measure unit. All prices, which are world prices, are deflated to represent their 2010 United States dollar value. To access the deflators, refer to the World Bank's World Development Indicators. For all information regarding Multicolour, please contact David Weinberg: info@multicolour.com

Kindly also note the following:

1. We merge bi-color tourmaline with chrome tourmaline into one tourmaline price. Often, it is possible to find tourmalines of different colors in the same mines.
2. We merge color change sapphire, fancy sapphire, and sapphire into one sapphire price. It is possible to find sapphires of different colors in the same mine.
3. We merge grossular garnet, tsavorite, color change garnet, and garnet into one garnet price. Garnets of different colors can be found in the same mine.
4. We merge chrysocolla quartz, rose quartz, rutilated quartz, and quartz into one quartz price.

C.3.27. multiplier_comtrade

This variable corresponds to the multiplier used for the conversion of the UN Comtrade price unit conversion into the standard measure unit.

C.3.28. multiplier_wb

This variable corresponds to the multiplier used for the conversion of the World Bank price unit conversion into the standard measure unit.

C.3.29. multiplier_usgs

This variable corresponds to the multiplier used for the conversion of the United States Geological Service (USGS) price unit conversion into the standard measure unit.

C.3.30. multiplier_multicolour

This variable corresponds to the multiplier used for the conversion of the USGS or World Bank price unit conversion into the standard measure unit.

C.3.31. annualallocationcapacity

This information is taken from United States Geological Survey (USGS). It measures yearly output of the mine or resource extraction site in the standard measure unit.

C.3.32. exp_annual_value_location1

This variable accounts for annual value of the location in 2010 United States Dollars (USD). This measure of the annual value of the location prioritizes UN Comtrade export prices first. Then, it incorporates prices from the World Bank, followed by those of the USGS. The variable excludes prices from Multicolour.

A few reasons underpin our rationale provide one set of prices without Multicolour values. First, not each resource-year in the Multicolour dataset has a high number of observations. Second, Multicolour sales tend to be a on a very small scale, with typical prices being at the gram or carat level. Accordingly, small fluctuations in the Multicolour prices per carat, which is normal given factors such as gem quality size, clarity, and color, can make a significant difference in the price. By contrast, the prices for most minerals from UN Comtrade, USGS, the World Bank tend to be aggregated at the kilogram, metric ton, or thousand metric ton levels, making them less prone changes from small fluctuations.

C.3.33. exp_annual_value_location2

This variable accounts for annual value of the location in 2010 United States Dollars (USD). This measure of the annual value of the location prioritizes UN Comtrade export prices first. Then, it incorporates world prices from World Bank, USGS, and Multicolour (in that order). The variable is calculated by multiplying the annual allocation capacity of the mine/resource extraction site by `export_price_first_mult2`.

C.3.34. wd_annual_value_location1

This variable accounts for annual value of the location in 2010 United States Dollars (USD). This measure of the annual value of the location prioritizes world prices from World Bank. Then, it incorporates US prices from USGS, followed by country-specific export prices from UN Comtrade. The variable excludes prices from Multicolour. The variable is calculated by multiplying the annual allocation capacity of the mine/resource extraction site by `world_price_first_mult1`.

A few reasons underpin our rationale provide one set of prices without Multicolour values. First, not each resource-year in the Multicolour dataset has a high number of observations. Second, Multicolour sales tend to be a on a very small scale, with typical prices being at the gram or carat level. Accordingly, small fluctuations in the Multicolour prices per gram or carat, which is normal given factors such as gem quality size, clarity, and color, can make a significant difference in the price. By contrast, the prices for most minerals from UN Comtrade, USGS, the World Bank tend to be aggregated at the kilogram, metric ton, or thousand metric ton levels, making them less prone to changes from small fluctuations.

C.3.35. wd_annual_value_location2

This variable accounts for annual value of the location in 2010 United States Dollars (USD). This measure of the annual value of the location prioritizes world prices from World

Bank and US prices from USGS. Then, it incorporates export prices from UN Comtrade. The variables excludes prices from Multicolour. The variable is calculated by multiplying the annual allocation capacity of the mine/resource extraction site by `world_price_first_mult2`.

C.3.36. local_value

This variable corresponds to the annual value of the location using only export prices from UN comtrade.

C.3.37. wb_value

This variable corresponds to the annual value of the location using only world prices from the World Bank's Global Economic Monitor Commodities Pink Sheet.

C.3.38. usgs_value

This variable corresponds to the annual value of the location using only US prices from the United States Geological Survey (USGS).

C.3.39. world_value_nomc

This variable corresponds to the the annual value of the location using world prices from the World Bank and US prices from USGS, excluding world prices from Multicolour. We include USGS prices alongside World Bank ones since, based our data,`wb_value` and `usgs_value` correlate at 0.91. That is even before logging the data, too. After logging `wb_value` and `usgs_value`, they correlate at 0.9982, which further strengthens our rationale for grouping World Bank and USGS prices.

C.3.40. world_value_withmc

This variable corresponds to the the annual value of the location using world prices from the World Bank and US prices from USGS, including world prices from Multicolour. We include USGS prices alongside World Bank ones since, based our data, `wb_value` and `usgs_value` correlate at 0.91. That is even before logging the data, too. After logging `wb_value` and `usgs_value`, they correlate at 0.9982, which further cements our rationale for grouping World Bank and USGS prices.

C.3.41. lootable

This is a dummy variable indicating, based on our research, that the resource is potentially lootable. To be lootable, a resource must have high value and low barriers to entry/extraction. We say "potentially" lootable because certain types of resources can be found in different extraction sites, and some of these extraction sites make it easier to extract than others. For example, gold may be mined through placer techniques, which can be done by most anyone. By the same token, gold can also be mined through the use of expensive dredging or digging machinery. Even though not everyone has access to the expensive machinery, the fact that almost anyone can mine gold through placer techniques makes the resource "lootable" for the purposes of this dataset.

C.4. Resource Price Data Availability

Table 13 provides the availability of prices used in this dataset by resource. In cases when there are prices from more than one source by variable, refer to Section C.3 for how we calculate the respective prices.

Table 13: Source of Resource Prices

Resource	UN Comtrade	USGS	World Bank	Multicolour
aggregates				
alumina		X		
aluminum	X	X	X	
aluminum floride	X			
amazonite				
amethyst				X
ametrine				X
ammonia	X			
ammonium nitrate				
andalusite	X		X	
anhydite				
antimony	X	X		
apatite			X	
aquamarine				X
arsenic	X			
arsenic trioxide				
asbestos	X	X		
asphalt	X			
attapulgite				
barite	X	X		
basalt	X			
bauxite	X	X		
bentonite	X	X		
beryl				X

Continued on next page

Table 13 : Source of Resource Prices – *continued*

Resource	**UN Comtrade**	**USGS**	**World Bank**	**Multicolour**
beryl and emerald				
bismuth	X	X		
black carbon	X			
boron	X	X		
cadmium	X	X		
calcite				X
carbon dioxide	X			
cast iron				
caustic soda				
celestine		X		
cement		X		
chromite	X			
chromium	X			
citrine				X
clay	X	X		
clinker				
coal	X		X	
coal tar				
cobalt	X	X		
cobalt-copper				
coke				
coking coal				
columbium-tantalite				
columbium and niobium		X		

Continued on next page

Table 13 : Source of Resource Prices – *continued*

Resource	UN Comtrade	USGS	World Bank	Multicolour
copper	X	X	X	
copper-cobalt				
diamond	X	X		
diatomite		X		
diesel				
dolomite	X			
emerald	X			X
feldspar	X	X		
ferroalloys	X			
ferrochromium	X			
ferromanganese	X			
ferromolybdenum	X			
ferrosilicon	X			
ferrotungsten				
ferrovanadium	X			
fertilizer		X		
fluorite				
fluorspar	X	X		
gallium		X		
garnet		X		X
gasoline				
gemstones		X		
germanium	X	X		
glass				

Continued on next page

Table 13 : Source of Resource Prices – *continued*

Resource	**UN Comtrade**	**USGS**	**World Bank**	**Multicolour**
gold	X	X	X	
gold and precious stones				
granite	X			
granite and marble				
graphite		X		
gravel	X			
gypsum	X	X		
helium		X		
hydroxide				
ilemite				
indium		X		
iron	X			
iron and steel		X		
iron oxides		X		
iron pyrites	X			
jet fuels	X			
kaolin	X			
kerosene	X			
kyanite	X			
labradorite				X
lapis				X
lead	X	X	X	
lead-zinc				
lime	X	X		

Continued on next page

Table 13 : Source of Resource Prices – *continued*

Resource	**UN Comtrade**	**USGS**	**World Bank**	**Multicolour**
lime and limestone				
liquified natural gas	X		X	
liquified petroleum gas	X			
liquified petroleum gas and natural gas liquids				
lithium		X		
magnesite	X			
magnesium	X	X		
manganese	X	X		
marble	X			
matte				
mercury	X	X		
methanol				
mica	X	X		
molybdenum	X	X		
morganite				X
naphtha	X			
natural gas	X			
nickel	X	X	X	
nickel-copper-cobalt				
niobium	X			
niobium and tantalum	X			
nitrogen	X	X		
nitrogen ammonia		X		

Continued on next page

Table 13 : Source of Resource Prices – *continued*

Resource	**UN Comtrade**	**USGS**	**World Bank**	**Multicolour**
oil	X		X	
oil and gas				
onyx				
opal				X
palladium	X			
peat	X	X		
perlite	X	X		
petrochemicals	X			
petroleum products			X	
phosphate	X	X	X	
phosphoric acid	X			
platinum	X		X	
polished gemstones	X			
potash				
pozzolan				
pozzolana				
pyrophyllite		X		
quartz	X			
rare earths		X		
rhodium	X			
ruby	X			X
ruthenium	X			
salt	X			
sand				

Continued on next page

Table 13 : Source of Resource Prices – *continued*

Resource	UN Comtrade	USGS	World Bank	Multicolour
sand and gravel		X		
sandstone				
sapphire	X			X
selenium	X	X		
sepiolite				
silica	X	X		
silicomanganese	X	X		
silicon	X	X		
silver	X	X	X	
soapstone				
soda ash	X	X		
sodium silicate				
sodium tripolyphosphate				
steel		X		
stone	X	X		
sulfur		X		
sulfuric acid	X	X		
tantalite				
tantalum	X	X		
tanzanite				X
tellurium		X		
tin		X	X	
titanium		X		
travertine	X			

Continued on next page

Table 13 : Source of Resource Prices – *continued*

Resource	**UN Comtrade**	**USGS**	**World Bank**	**Multicolour**
triple superphosphate				
tuff				
tungsten	X	X		
turquoise				X
uranium	X			
urea			X	
vanadium	X			
vanadium pentoxide	X			
vermiculite	X			
wolframite				
wollastonite				
zeolite				
zinc	X	X	X	
zircon	X		X	
zirconium	X	X		