



**HAL**  
open science

## Mineral resources and the salience of ethnic identities

Nicolas Berman, Mathieu Couttenier, Victoire Girard

► **To cite this version:**

Nicolas Berman, Mathieu Couttenier, Victoire Girard. Mineral resources and the salience of ethnic identities. *The Economic Journal*, 2023, 133 (653), pp.1705-1737. 10.1093/ej/uead018 . hal-04135257

**HAL Id: hal-04135257**

**<https://amu.hal.science/hal-04135257v1>**

Submitted on 23 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MINERAL RESOURCES AND THE SALIENCE OF ETHNIC IDENTITIES\*

*Nicolas Berman, Mathieu Couttenier and Victoire Girard*

This paper shows how ethnic identities may become more salient due to natural resources extraction. We combine individual data on the strength of ethnic—relative to national—identities with geo-localised information on the contours of ethnic homelands, and on the timing and location of mineral resources exploitation in 25 African countries, from 2005 to 2015. Our strategy takes advantage of several dimensions of exposure to resources exploitation: time, spatial proximity and ethnic proximity. We find that the strength of an ethnic group identity increases when mineral resource exploitation in that group’s historical homeland intensifies. We argue that this result is at least partly rooted in feelings of relative deprivation associated with the exploitation of the resources. We show that such exploitation has limited positive economic spillovers, especially for members of the indigenous ethnic group; and that the link between mineral resources and the salience of ethnic identities is reinforced among members of powerless ethnic groups and groups with strong baseline identity feelings or living in poorer areas, or areas with a history of conflict. Put together, these findings suggest a new dimension of the natural resource curse: the fragmentation of identities, between ethnic groups and nations.

Social identities are a first-order determinant of a large group of social, economic, institutional and political outcomes (Alesina *et al.*, 1999; Alesina and La Ferrara, 2005). In Africa, the high levels of social fragmentation resulting from identification to ethnic groups—rather than nations—has long been acknowledged as a primary cause of the ‘growth tragedy’: identity fragmentation appears to dampen trust, cooperation, public goods provision and government performance (Easterly and Levine, 1997; Habyarimana *et al.*, 2007; 2009). However, we are still

\* Corresponding author: Victoire Girard, Nova School of Business and Economics, R. da Holanda 1, 2775-405 Carcavelos. Email: [victoire.girard@novasbe.pt](mailto:victoire.girard@novasbe.pt)

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The authors were granted an exemption to publish parts of their data because access to these data is restricted. However, the authors provided the Journal with temporary access to the data, which enabled the Journal to run their codes. The codes for the parts subject to exemption are also available on the Journal repository. The restricted access data and these codes were also checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: <https://zenodo.org/record/7576268>.

The project leading to this publication has received funding from the French government under the ‘France 2030’ investment plan managed by the French National Research Agency (ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University—A\*MIDEX. Mathieu Couttenier acknowledges financial support from the IDEXLYON from the University of Lyon (French National Research Agency, ‘Programme Investissements d’Avenir’ ANR-16-IDEX-0005) and the Swiss National Foundation (grant Ambizione-PZ00P1 161451). This project benefited from the financial support of the chair Energy and Prosperity. Victoire Girard acknowledges funding by the Swiss National Foundation (Scientific Exchange grant IZSEZ0 179672), the Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722) and POR Norte.

a long way from understanding to which extent these groups' identities are fixed, and, if not, what makes them evolve (Anderson, 2006; Akerlof and Kranton, 2010).

This paper studies a potential driver of individual identification to ethnic groups in sub-Saharan Africa (SSA): the exploitation of mineral resources. Mineral resources exploitation offers a unique setting to investigate the mechanisms underlying variations in ethnic identities for three main reasons. First, minerals are point-source commodities, which are attached to specific locations. In SSA, these locations belong to historical, pre-colonial ethnic homelands, entities which still matter today and have been shown to influence current economic and political outcomes (Michalopoulos and Papaioannou, 2020) despite colonisation and subsequent nation-building policies. The exploitation of such resources is therefore intrinsically connected to the ethnicity question. Second, the perceived ownership of such resources is a recurrent subject of disputes between individuals who may claim the rights to ownership of natural resources—and therefore their potential rents (Collier, 2017): nationals from the country hosting the resource and those who feel ownership of the land above the resource (the locals or the members of the resource-rich ethnic group). These disputes may escalate up to the emergence of secessionist parties or to secession attempts, as in Aceh, Biafra and Katanga (Aspinall, 2007; Gehring and Schneider, 2020). Finally, given the debate on who should benefit from resources rents, minerals' extraction is potentially associated with positive economic spillovers as well as with various grievances.<sup>1</sup> By studying how mining affects a variety of individual outcomes—wealth, grievances, or actual and perceived ethnic inequality—we can discuss *how* mineral resources affect ethnic identification, and therefore provide results which may inform us more broadly about the drivers of ethnic—or local—identification beyond the resources case.

To study the link between ethnic identities and natural resources, we combine three main sources of data covering 25 SSA countries between 2005 and 2015. First, we exploit geolocalised individual data on the salience of ethnic (*versus* national) identities from four survey waves of the Afrobarometer. Second, we match the information on the ethnic group of each Afrobarometer respondent with a list of historical ethnic groups, which gives us the geographical boundaries of the historical ethnic homelands. Last, we make use of time-varying information on the location and exploitation of mineral resources to compute the yearly number of active large-scale mines in the ethnic homeland of each Afrobarometer respondent—this number of mines is our baseline proxy of resources exploitation in a homeland. Equipped with these data, we estimate how respondents from different ethnic backgrounds respond to changes in the relative intensity of natural resources exploitation in their historical ethnic homelands.

Our baseline identification strategy exploits three dimensions of variation in exposure to mineral exploitation: time, geographic proximity and identity proximity. Consider two individuals living at a given point in time, but belonging to different ethnic groups. Based on our combined data on the contours of each ethnic homeland and on time variations in mineral exploitation, we identify how changes in relative mining activity, over time and across the ethnic groups of these two individuals, affect their identification to their ethnic groups. The structure of the data allows us to partial out any local time-varying shocks affecting both individuals, and to account for any time-invariant differences in the level of identification to a given ethnicity in a given country, such as a group size or its historical political dominance.

We first document that ethnic identification becomes stronger when mineral resources exploitation intensifies in the historical homeland of the respondent's ethnic group. Interestingly,

<sup>1</sup> For a literature review on the effects of natural resources on local development, see for instance, Van der Ploeg (2011), Aragón *et al.* (2015), Cust and Poelhekke (2015) and Venables (2016).

the effect of resources is significant, both when exploitation takes place in the respondent's country of residence and when it takes place abroad; this resonates with the idea that ethnic borders continue to matter in a continent where national borders largely ignore them (the 'scramble for Africa'; e.g., Michalopoulos and Papaioannou, 2020). We also find that the effect of mineral resources on ethnic identification is persistent (and stronger two to three years after the exploitation starts) and tends to be magnified, during electoral periods. The baseline estimates are robust to a large battery of checks, including various definitions of ethnic homelands, or using alternative measures of mineral resources exploitation, namely changes in the value of mining production generated by variations in the world prices of minerals. We also reproduce our main analysis using ethnic homeland-specific variations in weather; we do not find any effect of such weather shocks, which suggests that our main results are driven by specificities of mineral resources extraction, rather than by general income fluctuations. Finally, we find that the effect of mining on the salience of ethnic identities holds regardless of whether the individuals live in their historical homeland. Overall, this first part of the paper shows that mineral exploitation causes more fragmented identities, by increasing the divergence in the importance of the ethnic or national identities reported by the members of the different groups present in each country.

The second part of the paper focuses on the channel underlying these findings. At the core of our discussion are the ethnic group-specific feelings of 'relative deprivation' and grievances. New mining activities may raise aspirations related to economic development, in particular among members of the ethnic groups hosting the resources (feeling a sense of entitlement over the rent from resources extracted in their group's homeland; Collier, 2017). However, under incomplete information, these aspirations may be unrelated to the reality of resources' extraction, resulting in feelings of deprivation. We build on specific case studies and recent literature linking economic development to aspirations and identities (notably Aspinall, 2007; Olzak, 2011; Cheeseman and Larmer, 2015; Binzel and Carvalho, 2017; Must, 2018). These deprivation feelings may heighten group identification, either in a decentralised manner (individuals valuing a social identity as a way to cope with unfulfilled economic aspirations, as in Binzel and Carvalho, 2017), or following an opportunistic manipulation by 'ethnic political entrepreneurs' (who use ethnic grievances to divert attention from their own rent extraction, or gain political leverage, as in Aspinall, 2007; Cheeseman and Larmer, 2015). The main competing explanation would be that natural resources make ethnic identities more salient, by increasing the economic or political 'pay-offs' associated with belonging to a particular group. This channel would be observed if minerals had strong positive local economic effects that are felt disproportionately by historical local groups, or if they increased the political power and representation of these groups or political patronage along ethnic lines. The actual pay-offs of belonging to a given group would increase which would make identities more salient. Combined with intra-ethnic spillovers (due to migration, trade, remittances or solidarity), both the 'pay-off' and the 'deprivation' channel could also explain why individuals residing outside their homeland display stronger identity feelings. However, the pay-off interpretation has a number of implications which differ from the deprivation channel which we can test empirically.

We provide a set of empirical exercises which globally support an interpretation rooted in the impact of natural resources on feelings of deprivation and ethnic grievances. We first show that in our sample, the effect of mineral resources on household wealth is limited, while the reports of economic deprivation and pessimism about economic conditions increase significantly with mineral exploitation. Interestingly, these effects are magnified for co-ethnics living outside their homeland, possibly because the discrepancy between the expectations and the reality of the

economic spillovers from resources exploitation is magnified in this subgroup. Using information on mining companies' characteristics, we also find that domestic mines tend to magnify self-reports of deprivation more than non-domestic mines, although the effect on individual wealth is similar for both types of mines ownership. Second, we investigate how the local effects of mines vary with distance to mines, and across ethnic groups. We find that mines have limited local economic effects and, crucially for the deprivation channel, that there is no additional economic benefit for individuals belonging to the group whose historical homeland hosts mining activities. However, individuals belonging to that group report more deprivation when mining activity intensifies. Taken together, these results support the deprivation rather than the pay-off channel.

Finally, we argue that, if indeed caused by deprivation feelings, the link between natural resources extraction and ethnic identities should be magnified in areas where feelings of ethnic inequality are pre-existing, e.g., relatively poor regions, politically excluded ethnic groups, or in areas with a history of violent conflicts (as appears in the case studies of Aspinall, 2007; Must, 2018). We bring these predictions to the data, making use of information contained in the Ethnic Power Relations (EPR) dataset on political exclusion, and of data on conflict and regional poverty. We show that the effect of mineral resources on ethnic identification is magnified, both for groups without any ethnic political representation and for powerless groups. We also find that a recent history of conflict in the ethnic homeland strengthens the mining-identity relationship.<sup>2</sup> Finally, we document that this relationship is magnified for members of relatively poor ethnic groups, as compared to the country average. These results corroborate the importance of the grievance channel in the mines-identity relationship.

Our paper contributes to three main research strands. The first is the recent research emphasising the endogenous nature of social identities. Since the seminal paper of Easterly and Levine (1997), numerous work has documented the economic consequences of fragmented identities (as reviewed in Alesina and La Ferrara, 2005). However, the strength of these identities feelings might not be fixed (Anderson, 2006; Akerlof and Kranton, 2010). We contribute to the recent line of research which shows that the existence of group identities, and the distance between them, are malleable.<sup>3</sup> Any individual has many social identities—based on nationality, gender, class, ethnicity, etc. In theory, which of these identities is the strongest can depend on how close one feels to other members of that identity relative to members of other identities, on the social salience of that identity, or on the pay-offs one may expect from that identity (Shayo, 2009). Empirically, strong group identities have thus been associated with either major events affecting the distance between individuals and the salience of certain identity lines—such as an ethnic conflict or a national sport victory (Rohner *et al.*, 2013; Depetris-Chauvin *et al.*, 2020)—or with the existence of positive group pay-offs—such as an ethnic group being

<sup>2</sup> This result uses information on the *pre-existing* intensity of violent conflicts in the homeland, that is, conflicts that took place before the beginning of our study. The more general relationship between minerals, identities and conflicts is beyond the scope of this paper. Whether the exacerbation of ethnic identity triggers conflict, or the opposite, or whether these two variables are co-determined (both may be means of reaching the same objective for the leaders) remains an open question.

<sup>3</sup> On identity endogenisation, see, for example, Klor and Shayo (2010), Rohner *et al.* (2013), Robinson (2014), Bisin *et al.* (2016), Ahlerup *et al.* (2017), Green (2020) and Dehdari and Gehring (2022). On the importance of distances across identities on development processes, the most common measures of identities fragmentation, namely the polarisation index and the fractionalisation index, are weighted by distance (Esteban *et al.*, 2012). The fractionalisation index assumes by default a maximum value for distance (at 1). Given the difficulty to measure intergroup distance, the same has often been done in studies using the polarisation index. In the absence of any intergroup distance, these measures would take a value of zero, no matter how fragmented the society is.

historically richer or ethnic competition in political campaigns (Eifert *et al.*, 2010; Pengl *et al.*, 2022). Our paper contributes to this literature by highlighting how individuals react to a shock which is ethnic group specific and time varying. This setting allows us to draw attention on the importance of pay-off perceptions and grievances highlighted in case studies (Aspinall, 2007; Cheeseman and Larmer, 2015; Must, 2018), and to document how specific ethnic group characteristics are likely to reinforce the cycle between identity fragmentation, deprivation, and exclusion.

Our results also call for giving more attention to social identities and grievances in the large literature on the effects of natural resources' extraction (surveyed by Van der Ploeg, 2011; Cust and Poelhekke, 2015; Venables, 2016). A body of work has documented the political economy side of the so-called natural resource curse, and outlined the risks of conflicts about rent sharing across identity groups or regions.<sup>4</sup> Using data on mines in Africa similar to ours, Berman *et al.* (2017) find that increases in the world prices of minerals are associated to more violence locally, and that variations in the value of mines within specific ethnic homelands foster the activity of rebel groups related to these ethnic groups. These findings call for a better understanding of what motivates individuals living in resource-rich environments to engage in violence or support fighting groups. By documenting a link between natural resources, identity and deprivation, our work can be viewed as an additional step in that direction. Recent works argue that, even before exploitation starts, the existence of resources triggers an optimism unrelated to their economic reality, among both specialised growth forecasters and households (Cust and Mihalyi, 2017; Cust and Mensah, 2020). Here, we show that after resources exploitation starts, and even if resources may have positive economic spillovers (Aragón and Rud, 2013),<sup>5</sup> group-specific grievances and identity fragmentation arise. These results suggest a new political economy channel through which resources exploitation may hinder development.

Finally, our work also indirectly speaks to the literature on political borders in Africa, which shows that, partly because of the arbitrary way in which national borders were designed by colonial powers (the 'scramble for Africa'), pre-colonial ethnic institutions persist and have still observable effects today (see the survey by Michalopoulos and Papaioannou, 2020). We find that what occurs inside historical ethnic homelands matter, even for individuals living outside their historical homelands. We also find that natural resources in one's homeland have a significant effect on ethnic identification, when the resources are located in a part of the homeland which is outside the respondent's country of residence. Put differently, our results provide new evidence supporting the prevalence of ethnic borders over administrative ones.<sup>6</sup>

The rest of the paper is organised as follows. Section 1 presents the data and our baseline empirical strategy. Section 2 contains the results on how mineral resources activity shape ethnic identification. In Section 3, we discuss theory and evidence supporting the deprivation channel. Section 4 concludes.

<sup>4</sup> See, for instance, Monteiro and Ferraz (2010), Caselli and Michaels (2013), Morelli and Rohner (2015), Loayza and Rigolini (2016), Berman *et al.* (2017), Collier (2017), Fenske and Zurimendi (2017), Vogt (2017), Mamo and Bhattacharyya (2018), Christensen (2019), Lessmann and Steinkraus (2019) and Girard *et al.* (2021).

<sup>5</sup> The relative deprivation we document, grounded in a discrepancy between expectations and reality, will be exacerbated if the forecasts of the resources economic spillovers are over optimistic.

<sup>6</sup> Our results align well with the literature stressing how ethnic networks and linkages (Fafchamps, 2000; 2003; Aker *et al.*, 2014), ethnic leaders (Asiwaju, 1985; Aspinall, 2007; Michalopoulos and Papaioannou, 2020), or sheer ethnic solidarity (Habyarimana *et al.*, 2007; 2009; Lowes *et al.*, 2015), all craft behaviours, and matters potentially as much or more than individuals' physical location.

## 1. Data and Empirical Strategy

To assess the impact of mineral resources on ethnic identification, we combine data on (i) the location and incidence of resources' production; (ii) the historical borders of ethnic groups' homelands; and (iii) individuals' level of identification to their ethnic group. The Online Appendix contains more details about the dataset construction.

### 1.1. Ethnic Identification

We exploit rounds three to six of the Afrobarometer surveys.<sup>7</sup> The data are repeated cross-sections and cover a total sample of more than 100,000 respondents in 25 sub-Saharan African countries over the period 2005–2015.<sup>8</sup> For each round, the Afrobarometer reports the exact centroid coordinate of respondents' town, village or neighbourhood of residence (BenYishay *et al.*, 2017), and provides detailed individual characteristics such as age, gender, education level, employment status, as well as ethnic group identification.

Our main outcome variable is the strength of individual feelings towards their ethnic group identity. It comes from the following question: *'Let us suppose that you had to choose between being a [National] and being a [R's Ethnic Group]. Which of the following best expresses your feelings?'* Our baseline variable of 'ethnic identification' takes the inverse of the values suggested in the questionnaire, hence the highest values denote the strongest identity feeling: *5 = I feel only (R's ethnic group), 4 = I feel more (R's ethnic group) than national, 3 = I feel equally national and (R's ethnic group), 2 = I feel more national than (R's ethnic group), 1 = I feel only national.* Alternatively, we use a binary variable taking the value 1 if the respondent's answer is *I feel only* or *I feel more* or *I feel equally*. We also consider a more restrictive binary measure, which equals 1 only if the respondent feels more or only ethnic. In total, 53% of the respondents declare to feel at least equally strongly or more strongly about their ethnic identity than about their national identity.

The Afrobarometer surveys contain several other variables that we exploit to document the channels of transmission. In particular, we make use of various proxies for wealth and economic deprivation, as well information about the respondents' perception of their economic conditions. Finally, we also exploit a question where respondents are asked whether they believe that their ethnic group is treated unfairly.<sup>9</sup> We describe these variables in Section 3 and present their summary statistics in Online Appendix, Section A3.

<sup>7</sup> The question on ethnic identity that we exploit was asked consistently in these rounds, while earlier rounds gave respondents a list of several alternative identities.

<sup>8</sup> The Afrobarometer coverage is not exhaustive and calls for caution in generalising our results. However, we are confident that our results are informative about a significant share of the continent population as in 2015, 70% of the population of SSA was living in one of the countries included in our sample (population numbers source: the United Nations World Population Prospects 2019). The countries covered are: Benin, Botswana, Burkina Faso, Cameroon, Cote d'Ivoire, Ghana, Guinea, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mozambique, Namibia, Niger, Nigeria, Senegal, Sierra Leone, South Africa, Tanzania, Togo, Uganda, Zambia and Zimbabwe (see Online Appendix, Section A1, for the year of survey of each country). Each survey round in each country has a nationally representative sample of either 1,200 or 2,400 respondents. Data are available at <http://www.afrobarometer.org>. The Online Appendix reports the year of survey for each country.

<sup>9</sup> Another relevant variable relates to interpersonal trust, particularly to trust on co-ethnics versus other ethnic groups. Unfortunately, this variable only in one of the four Afrobarometer rounds that we consider, which makes it unusable in our context.

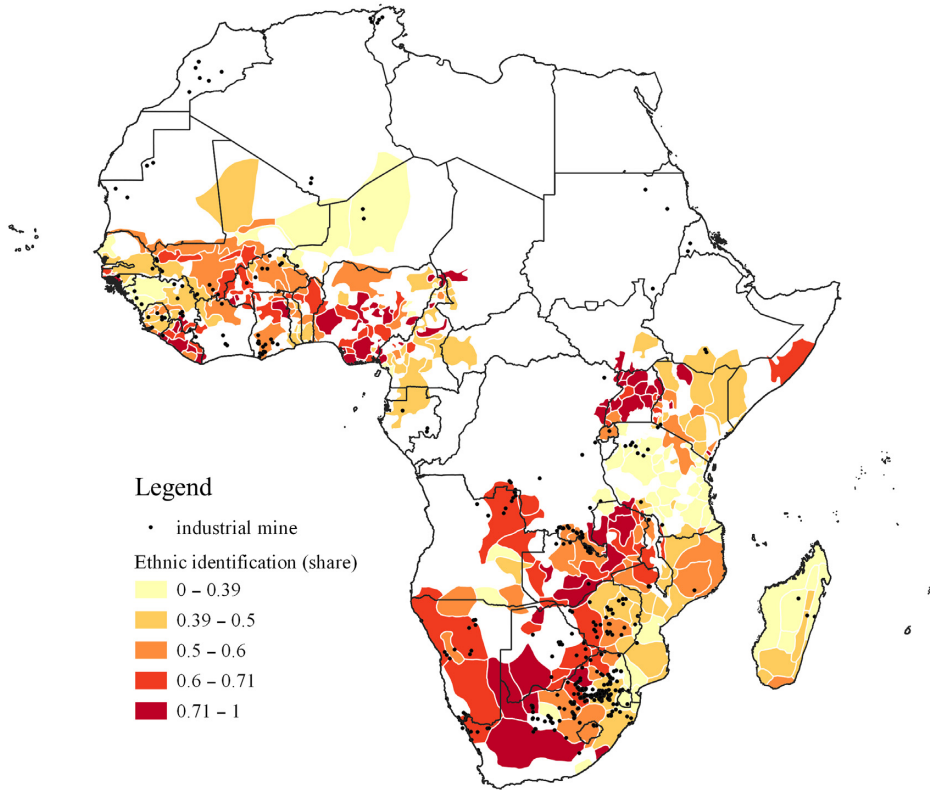


Fig. 1. *Ethnic Homelands, Mining and Ethnic Identification.*

*Notes:* Authors' computations from the Afrobarometer surveys, Murdock and S&P data. See main text for data sources. The map shows the location of each industrial mine that has been active between the years 2005 and 2015, and the baseline matching of Afrobarometer and Murdock ethnic groups. For each ethnic group, the map shows the share of members of that group who value their ethnic identity equally or more than their national identity. Each share takes in account the answers of all the members of each ethnic group, independently of the homeland in which these members live.

## 1.2. *Boundaries of Ethnic Homelands*

We combine the Afrobarometer surveys with information on the boundaries of respondents' ethnic homelands to assess individuals' reactions to shocks occurring in these homelands. Our baseline estimations exploit the digital maps of historical ethnic homelands of Nunn (2008), based on the compilation by Murdock (1959) of the ethnographic information available in the late nineteenth century. We match the self-reported ethnic group of respondents to a Murdock homeland for 91% of the respondents.<sup>10</sup> This results in a total of 321 ethnic homelands located in 42 countries.<sup>11</sup> These homelands are represented in Figure 1, which also displays the average level of ethnic identification by ethnic homeland, computed from Afrobarometer data. More

<sup>10</sup> We attribute to the respondent an ethnic group based on the following question: 'What is your ethnic community, cultural group or tribe?' Only 0.5% of the respondents do not answer the question and 1% of them answer a national identity only.

<sup>11</sup> This number is larger than the 25 countries included in the Afrobarometer, as the homelands of the Afrobarometer's respondents often span over countries that are not necessarily in the Afrobarometer.



precisely, the map shows, for each ethnic group, the share of members of that group who value their ethnic identity equally or more than their national identity: there is a substantial variation both across and within countries in the salience ethnic identification.

In our baseline estimations, we match each ethnic group of the Afrobarometer surveys with a single, main group from Murdock (1959). This matching procedure is the most direct, and also used by Nunn (2008) and Nunn and Wantchekon (2011). It is, however, restrictive: Afrobarometer and Murdock groups might not share a unique correspondence. An Afrobarometer group, for instance, might appear as several subgroups in Murdock.<sup>12</sup> This explains why multiple blank areas appear in Figure 1, even in countries covered by the Afrobarometer. To take into account this limitation, we also match Afrobarometer and the Murdock Atlas homelands using an alternative matching rule, allowing for multiple correspondence between the two datasets. We make use of an aggregated definition of groups present in the Murdock Atlas, namely the ‘cultural’ groups. The Murdock Atlas attributes each of its 843 ethnic homelands to 104 distinct cultural groups.<sup>13</sup> In this approach using cultural groups, a single ethnic group of the Afrobarometer surveys can be matched with several Murdock homelands, ensuring a more exhaustive coverage of the Afrobarometer countries. Last, we resort to an alternative matching procedure based on the recently released LEDA: Linking Ethnic Data from Africa (Müller-Crepon *et al.*, 2022). While our baseline matching resorts to historical information on groups evolution and ancestry (like Nunn, 2008), the LEDA relies on groups proximity after linking these groups to the Ethnologue language tree. The outcome of both the cultural group matching and the LEDA matching is shown in Figures A.1 and A.2 in the Online Appendix, Section A2, where the matching procedures are described in more details.

Though the Murdock data is widely used by the literature (e.g., Gennaioli and Rainer, 2007; Nunn, 2008; Michalopoulos and Papaionnaou, 2013; Alsan, 2015), the fact that homelands are a snapshot from historical maps comes both as an advantage and as an inconvenience. On the negative side, settlement patterns may be outdated. Also, as homelands maps were built based on existing anthropological work, the level of accuracy and resolution varies for different regions in the continent (coastal areas are typically more accurately described than remote inland areas, see e.g., Michalopoulos and Papaioannou, 2020, for a discussion). On the positive side, this historical snapshot alleviates concerns of ethnic group location being endogenous to the evolution of ethnic identities salience that we aim at explaining.

As an alternative strategy, we also match ethnic groups of the Afrobarometer to homelands from the Ethnic Power Relations (EPR) dataset. The EPR dataset being based on both contemporaneous and time-varying information, the borders of the EPR ethnic groups are more likely to be affected by contemporaneous shocks than the Murdock data. To mitigate this concern, we compute EPR-related information based on the EPR dataset *before* the start of our analysis, in the year 2004. We also note that the focus of the EPR data on politically relevant groups leads to omitting many ethnic groups—and sometimes entire countries—from the EPR sample.<sup>14</sup> However, boundaries of the EPR data are *de facto* more accurate. Moreover, we make use of the richness of the EPR

<sup>12</sup> Indeed, although the Afrobarometer questionnaire on ethnic groups is supposed to use the same classification as Murdock (1959), the overlap is far from perfect. The most extreme case may be in the Ivory Coast, where the Afrobarometer only records five different ethnic groups, while the Murdock Atlas contains 36 distinct groups.

<sup>13</sup> The ‘Lobi’ cultural group territory will, for example, consist of the homelands of four ethnic groups, namely that of the Birifon, Dorosie, Kulango and Lobi.

<sup>14</sup> Vogt *et al.* (2015, p. 1329) note that ‘An ethnic group is considered politically relevant if at least one political organization has claimed to represent its interests at the national level or if its members are subjected to state-led political discrimination’.

Table 1. *Summary Statistics for the Main Variables.*

	Mean	SD	1st Quartile	Median	3rd Quartile
Ethnic identity	2.35	1.19	1	3	3
Ethnic identity (dummy) <sup>a</sup>	0.53	0.5	0	1	1
Ethnic identity (dummy, alt.) <sup>a</sup>	0.13	0.34	0	0	0
No. mines in homeland	2.06	5.79	0	0	1
No. mines in homeland, LEDA homelands <sup>b</sup>	1.48	5.15	0	0	1
No. mines in homeland, cultural	18.18	35.42	0	4	14
No. mines in homeland, EPR homelands <sup>b</sup>	6.25	21.91	0	0	1
No. mines in homeland, in country	1.36	4.93	0	0	0
No. mines in homeland, abroad	0.71	3.24	0	0	0
Residence in homeland (dummy)	0.46	0.5	0	0	1
Residence in rural area (dummy)	0.63	0.48	0	1	1
Female (dummy)	0.5	0.5	0	0	1
Primary education or more (dummy) <sup>c</sup>	0.61	0.49	0	1	1
Active (employed or looking) (dummy) <sup>c</sup>	0.64	0.48	0	1	1
Age	36.69	14.49	25	33	45
Number of respondents = 115,105					

*Notes:* Authors' computations from the Afrobarometer surveys, Murdock and S&P data. See main text for data sources. <sup>a</sup>Ethnic identity (dummy) equals 1 if the respondent's answer is *I feel only* or *I feel more* or *I feel equally* ethnic. Ethnic identity (dummy, alt.) equals 1 only if the respondent feels more or only ethnic. <sup>b</sup>The descriptive statistics for the LEDA matching come from the sample of 116,117 households who belong to groups identified in the LEDA dataset, the descriptive statistics for the EPR come from the sample of 85,342 households who belong to groups identified in the EPR dataset. <sup>c</sup>In our estimations, we control for all nine education categories and all four employment categories provided by the Afrobarometer surveys.

records on the political status of each group. These data allow us to estimate how our results vary across ethnic groups with different degrees of access to political power.

### 1.3. *Mineral Resources*

Data on natural resources come from S&P Global - SNL Metals and Mining. The dataset includes large-scale mines (industrial mines), usually owned and operated by multinationals or national firms. It covers 33 minerals, as well as information on the location of the mine, whether it is active, the volume of production, and the year in which production started.<sup>15</sup> We are therefore able to compute the number of active mines in the ethnic homeland  $e$  of individual  $i$  at time  $t$ , which we use as our main explanatory variable. The location of mines is shown in Figure 1. Mining activity is clustered: on average, each ethnic homeland contains two mines, homelands with at least a mine contain on average six mines.

### 1.4. *Descriptive Statistics*

Table 1 displays descriptive statistics on our final sample made of 115,105 respondents (the sample of our baseline estimates; Table 2, column 1). Most respondents live outside their ethnic homeland and in rural areas (54% and 63%, respectively). Respondents split evenly across genders, 61% of them have completed a primary education qualification (or more), and 64% are active (either employed or looking for a job).

<sup>15</sup> The following minerals are covered: U3O8, bauxite, chromite, chromium, coal, cobalt, copper, diamond, ferro chrome, gold, graphite, heavy mineral sand, ilmenite, iron, lanthanide, lead, lithium, manganese, molybdenum, nickel, niobium, phosphate, platinum, potash, rutile, silver, tantalum, tin, titanium, tungsten, vanadium, zinc and zircon. It omits artisanal and small-scale mines.

The respondents are divided across 296 ethnic groups in the final sample, with 20.9% of these ethnic groups hosting at least one mineral activity in their homeland over the 2005–2015 period. As resource-rich ethnic groups are on average larger than resource-poor ethnic groups, 34% of the respondents report coming from a group whose homeland hosts minerals.<sup>16</sup> The average number of mines substantially varies when we move from the most restrictive to the broader homeland definition: the average number of mines in our final sample is 2.1 with our baseline measure, and goes up to 18.2 when we use Murdock’s more aggregated ‘cultural groups’ definition. Note that, as the mine variable is right-skewed, we check the robustness to the presence of outliers in our sensitivity analysis.

### 1.5. Identification Strategy

The purpose of our empirical strategy is to estimate how natural resources extraction—in the form of minerals—taking place in the historical homeland of an ethnic group affects the strength of identity feelings declared by the members of that ethnic group. Our source of identification relies on the comparison of two individuals living in the same region, in a given point of time, but belonging to two different ethnic groups.<sup>17</sup> Combining our data on the contours of each individual’s ethnic homeland and the time-varying information of mining activity location, we identify how changes in relative mining activity across each individual’s homelands, over time, affect their relative ethnic identification:

Formally, for an individual  $i$  belonging to ethnic group  $e$ , living in region  $r$  of a country  $c$  at time  $t$  (the specific month of each year), we estimate the following specification:

$$\text{Ethnic identification}_{i,e,r,t} = \alpha \times (\text{No. mines})_{i,e,t} + \mathbf{c}'_i \beta + \mathbf{FE}_{e,c} + \mathbf{FE}_{r,t} + \varepsilon_{i,e,r,t}. \quad (1)$$

The dependent variable represents the strength of ethnic identification of individual  $i$  (as a categorical or as a dummy variable), and our main explanatory variable is the number of active mineral resources exploitation in the ethnic homeland  $e$  of individual  $i$  at time  $t$  (**No. mines** <sub>$i,e,t$</sub> ). We test the robustness of our results to using various definitions of ethnic homelands. We consider two different matching procedures between the Afrobarometer’s ethnic groups and Murdock’s homelands, as explained in Section 1.2. We also show results using the alternative contours of homelands boundaries from EPR.

The vector  $\mathbf{c}_i$  includes a large set of respondent characteristics: gender, age and its square, a set of dummies controlling for education levels and employment status, and residence in a rural area. We also control for a dummy coded one if individual  $i$  is living in their ethnic homeland at the time of the survey, and zero otherwise. We include ethnic homeland/country fixed effects ( $\mathbf{FE}_{e,c}$ ), which capture any time-invariant feature of each ethnic group in each country, as well as inherent differences in the level of identification to a given ethnic group in each country (and other historical aspects like its historical political dominance). We also include region of residence  $\times$  year  $\times$  month fixed effects ( $\mathbf{FE}_{c,r,t}$ ), where a region is defined at the first administrative level

<sup>16</sup> The complete sample—just after matching the self-reported ethnic group of respondents and Murdock homelands—covers 321 groups. In this complete sample, 34.5% of the respondents report coming from an ethnic group rich in minerals.

<sup>17</sup> This approach is fundamentally distinct to the paper maybe closest to us (Ahlerup *et al.*, 2017). Ahlerup *et al.* (2017) document how local income, proxied by night-time light and instrumented by mineral production, is associated with more identification to the nation. Our work differs in terms of research question (taking resources as a group-specific shock rather than considering their average effect), identification strategy (allowing resources to matter through other channels than a wealth effect), and sample (extended in time and space, allowing us to control for any time-invariant ethnic group characteristics, such as pre-existing differences in the level of identification across groups).

(Admin-1).<sup>18</sup> This allows us to account for local shocks such as climate variations or crop prices. Any local economic spillover of natural resources exploitation that would be specific to a location, without affecting differentially the members of the different ethnic groups, is also captured by  $\mathbf{FE}_{c,r,t}$ . Given the inclusion of these fixed effects,  $\alpha$  identifies the effect of mining activity in individual  $i$ 's homeland on that individual's ethnic identification, relative to individuals of a different ethnic group living in the same region at a given point in time. The inclusion of two sets of high dimensional fixed effects leads us to use a linear probability model to estimate (1), to avoid the incidental parameters problem. Finally, SEs are clustered at the ethnic group level, though we show that our baseline results are extremely similar when allowing for spatial correlation in the error term.

### 1.6. Identification Issues

Saturating the model with a large array of fixed effects limits omitted variables concerns and ease of interpretation, but the remaining variance of our variable of interest  $\mathbf{No. mines}_{i,e,t}$  may become small, and may be affected by specific cases. We consider this issue in Section A5 of the Online Appendix, where we show how the residual variation in the number of mines is affected by our various dimensions of fixed effects. As shown in the Online Appendix, Table A.3, the raw variable has an SD of 5.79. When purged from the contribution of our baseline set of individual controls  $\mathbf{c}_i$ , the SD is barely affected (5.72). The variation drops by much more when ethnic homeland or ethnic homeland  $\times$  country fixed effects are included: the SD goes down to around 1. This is expected, as these fixed effects capture all the variations in the homelands in which the number of mines does not change over our sample period—which is about 80% of the homeland  $\times$  countries. Interestingly, the SD remains high (4.17) when the region of residence  $\times$  time fixed effects are included. This is because individuals of different ethnic groups are present at this level, making it more likely that a different number of mines is exploited in their homeland. When all fixed effects are included, the SD falls at 0.77. Hence, thinking about the effect of 'one additional mine'—the metric we use when commenting our results—is not very far from a (conditional) SD increase in the number of mines. Note that we discuss in the robustness, in Section 2.3, how sensitive our results are to using more or less restrictive sets of fixed effects.

Finally, a potential endogeneity concern is that observed changes in mining activity over time—opening and closing of mines—could be driven by ethnic groups' behaviours that correlate with the strength of ethnic identities. For instance, the political power of the ethnic group might drive both ethnic identification and changes in natural resource production. We will show that many of our results are difficult to reconcile with this view: in fact, the strongest effect is found for the least politically relevant groups. The number of active mines is also easy to observe for citizens. Still, in our robustness exercises we use an alternative measure of mining activity, which combines information on the volume of production at the beginning of the period with (exogenous) yearly variations in world prices of minerals.

<sup>18</sup> The Afrobarometer survey being a repeated cross-section, the exact village or town neighbourhood where the survey takes place moves across waves. Making use of the exact centroid coordinate of respondents' town, village or neighbourhood of residence, to define the region of residence would result in more than 12,000 local  $\times$  year  $\times$  month fixed effects (median of five respondents by level of fixed effects). We thus instead use the region of residence defined at the regional (Admin-1) level. In our sensitivity analysis, we show that our baseline results are robust to using a more disaggregated definition of a region, or a definition of region that does not rely on political boundaries.

Table 2. *Natural Resources Extraction and Ethnic Identification.*

Dep. Var.	(1)	(2)	(3)	(4)	(5)
	Salience of ethnic group identification				
Version	Categorical	Dummy	Dummy (alt.)	Categorical	Categorical
Sample	Full	Full	Full	Full	Outside homeland
No. mines: ethnic homeland	0.018 <sup>a</sup> (0.003)	0.006 <sup>a</sup> (0.002)	0.004 <sup>a</sup> (0.001)		0.016 <sup>a</sup> (0.003)
No. mines: ethnic homeland, in country				0.017 <sup>a</sup> (0.005)	
No. mines: ethnic homeland, abroad				0.018 <sup>a</sup> (0.005)	
Residence in homeland	0.044 <sup>a</sup> (0.012)	0.014 <sup>a</sup> (0.005)	0.012 <sup>a</sup> (0.003)	0.044 <sup>a</sup> (0.012)	
Fixed effects	Country × Ethnic group, Region × Time				
Individual controls	Yes				
Observations	115,105	115,105	115,105	115,105	61,790
R <sup>2</sup>	0.185	0.191	0.110	0.185	0.180
Sample mean/SD dep. var.	2.35/1.19	0.53/0.50	0.13/0.34	2.35/1.19	2.31/1.18

Notes: OLS estimations. Significant <sup>a</sup>at 1%. SEs are clustered at the ethnic group level. In column 5, we restrict the sample to individual living outside their ethnic homeland. The dependent variable is the level of ethnic identification with respect to state identification. In columns 1, 4 and 5, the variable ranges from categories 1 to 5. In column 2 we use as dependent variable a dummy taking the value 1 if the individual identifies at least as much to the ethnic group than the country. In column 3 we use an alternative definition of the binary dependent variable, taking the value 1 if the individual identifies strictly more to the ethnic group than the country. Controls include age and its square, gender, rural/urban dummy, education categories and employment status. There are 1,824 Admin-1 region × year × month fixed effects and 401 country × ethnic group fixed effects. Sample mean/SD dep. var. are respectively the sample mean and SD of the dependent variable in the sample of the corresponding column.

## 2. Mineral Resources Activity and Ethnic Identification

### 2.1. Baseline Estimates

Table 2 displays the baseline estimates. All columns include country × ethnic group fixed effects and region of residence × year fixed effects, where a region is defined at the Admin-1 level and time corresponds to month-year. We also control for a set of respondents' characteristics—gender, education, whether the respondent is in a rural or urban area and whether the respondent is located in their historical homeland. Table A.2 in the Online Appendix reports the complete set of coefficients for these control variables. In line with the literature (Robinson, 2014), female, uneducated and rural respondents tend to identify more to their ethnic group.

We start by estimating the impact of the number of mines in the respondent's ethnic homeland on the categorical measure of ethnic identification (column 1). The measure ranges from 1 to 5, with higher levels denoting stronger identification of the respondent to their ethnic group. The coefficient on the number of mines is positive and significant: ethnic identities become more salient as the number of mines in the ethnic homeland increases. Quantitatively, an additional mine raises the level of ethnic identification by 0.018 (column 1), i.e., around 1.5% of the (unconditional) SD of ethnic identification. Though it appears limited at first glance, the order of magnitude of the estimated effect of mining is comparable to that of other documented determinants of ethnic identities (Robinson, 2014). For instance, female respondents, inactive individuals, those living in a rural area, or residing in the historical homeland of the ethnic group respectively exhibit 0.051, 0.050, 0.032 and 0.044 higher levels of ethnic feelings (see

Online Appendix, Table A.2, column 1), i.e., between 2.7% and 4.3% of the SD of the ethnic identification variable.<sup>19,20</sup> The effect is qualitatively similar when using the binary measure of identification, that takes the value one for all respondents who report feeling ‘only ethnic’, ‘more ethnic than national’ or ‘equally ethnic and national’ (column 2), or alternatively only the first two categories (column 3).

It is well known that the historical homelands of many ethnic African ethnic groups have been partitioned by the national borders drawn by colonial powers: in our sample 167 of the 296 ethnic homelands span over more than one country. In Table 2, column 4, we allow the effect of natural resources production to differ, depending on whether it takes place in a part of the respondent’s homeland which lies inside or outside the respondent’s country. We do not detect any significant difference. Note that these results should not be seen as a comparison within ethnic groups: less than 3% of the respondents in our sample have an ethnic homeland partitioned across national borders and hosting natural resources on both sides of the border. Hence, the coefficients shown in Table 2, column 4, are identified across rather than within groups, which prevents us from comparing their relative size. The fact that mines outside the respondent’s country are found to significantly affect ethnic identification is, however, informative, as it echoes the literature on political borders in Africa, according to which ethnic partitions are still relevant entities today. This result also aligns with the general story we uncover in the mechanisms section—that natural resources foster ethnic identification by triggering feelings of exclusion and economic deprivation for relatively powerless groups: indeed, those feelings may arise regardless of whether the homeland of the individual is located within or outside the country. These feelings may travel through political borders as ethnic networks and linkages, ethnic leaders or sheer ethnic solidarity craft behaviours, and may at times be more relevant than (often artificial) national borders.<sup>21</sup>

Our estimations, so far, do not make any distinction between mines operating within the region of residence of the individuals or outside that region. These two cases are, however, quite different, if only because local mines may directly impact local economic conditions. In Table 2,

<sup>19</sup> We do not know, when an individual resides outside of their historical homeland, whether it is due to the evolution in time of the boundaries of the homeland, or to a decision to move away taken by the individual or their ancestors. If an individual decision, next to usual motives like joining an urban centre or a job opportunity, identity feelings might drive self selection into migration. If this were the case, we would expect individuals with least attachment to their ethnic identity to be the ones most likely to move away from their ethnic group homeland. This could contribute to explain the positive relationship that we observe between residence in one’s group historical homeland and the strength of ethnic identity. In this scenario, the coefficient we estimate for residence in the homeland would be upward biased. The fact that the magnitude of the effect of active mines is about half of that of residence in the homeland, would be an even stronger signal of the importance of our channel on ethnic group identification.

<sup>20</sup> Quantitatively, our baseline effect is also comparable to the one found by Depetris-Chauvin *et al.* (2020), who show that ethnic identification decreases in the days following sport victories of the national team. In their specification closest to our baseline (Online Appendix, Table A.11, col. 2), a national team’s victory decreases the level of ethnic identification by around 3.9% of the SD.

<sup>21</sup> Co-ethnics living on both sides of a country’s border may share strong economic linkages, given the importance of ethnic networks in the credit market, for trust, trade or migration (Fafchamps, 2000; 2003; Aker *et al.*, 2014). Co-ethnics may also refer to common ethnic leaders and institutions. As an illustration, Asiwaju (1985) cites a Ketu king claiming during colonial times ‘we regard the boundary [between Benin-Dahomey and Nigeria] separating the English and the French, not the Yoruba’. These ethnic institutions and leaders have endured the colonial and post-independence era and are, still today, the relevant unit of projection for many individuals (Michalopoulos and Papaioannou, 2020). Furthermore, some leaders may strategically manipulate ethnic identities (Aspinall, 2007, acting as ‘ethnic political entrepreneurs’). Finally, co-ethnics living in different countries might share their deprivation feelings through the pure empathy or solidarity effect documented in the experimental literature, which is independent from any personal or economic ties (Habyarimana *et al.*, 2007; 2009; Lowes *et al.*, 2015).

column 5, we restrict the sample to respondents who live outside of their ethnic homeland. The magnitude of our coefficient of interest barely changes.<sup>22</sup>

Our interpretation of these estimates is that ethnic identity is likely to become increasingly salient for individuals belonging to ‘resource-rich’ ethnic groups—whose homelands host natural resources—as opposed to national identity becoming more salient for resource-poor individuals. We focus the interpretation on resource-rich individuals for two main reasons. Firstly, when minerals are located abroad, minerals should not impact identification of the individuals who are co-nationals of the members of the resource-rich groups, but without having any resources themselves. The significant effect of resources abroad which appears in Table 2, column 4, is thus likely driven by an increase in ethnic identification of resource-rich individuals. Second, though we do not fully rule out that members of resource-poor groups may react to the presence of minerals within their country,<sup>23</sup> we expect members of resource-rich groups to react to resources exploitation at least as much as the members of other groups.

Overall, our estimates correspond to an increase in identities fragmentation, as we document an increase in identities distance between the resource-rich group members and other individuals. To be more precise, identity fragmentation is in theory a combination of the existence of identity boundaries between two or more groups, and the pair-wise distance between the members of these different identities (where various definitions of the identity groups are possible, such as ethnicity, religion or nation; Esteban and Ray, 2011). We here document a widening of the distance between the identities of the resource-rich and the resource-poor groups, a concerning finding in terms of comparative development.<sup>24</sup>

## 2.2. Persistence

### 2.2.1. Timing of the effects

The results of Table 2 show how individuals react to variations to the stock of active mines in their homeland, but are silent on the moment when the mining activity affects individual ethnic identification. Indeed, given that the rounds of the Afrobarometer occur several years apart, the changes in mineral exploitation could have occurred at the time of the survey or several years before. In other words, our baseline estimates could reflect either a short-lived ‘news shock’ effect, or a more persistent one. To test when the impact of mining activity on ethnic identification starts taking place, we turn to a specification considering the time since the change in the number of mines.

<sup>22</sup> We also allow the effect of natural resources to vary with individual characteristics such as gender, education or type of residence; the estimates do not display significant differences along these dimensions.

<sup>23</sup> The reaction of members of resource-poor groups could magnify our estimates if, as these individuals cannot claim the natural resources wealth through ethnic linkages, they use their national identity to justify that resources rents must be shared across all citizens. Alternatively, the reaction of resource-poor individuals might attenuate our estimates if members of all groups witness an increase in their ethnic identity feelings when exploitation intensifies, and the only reason we still identify a divergence is because the increase is strongest for members of the resource-rich groups. A particular case that could lead to such a reaction is the one in which two groups contest an ethnic border with a mine; we could then expect that their members feel more ethnic given this contest. Such a contest would play against finding a significant effect as it would lead to attenuation bias.

<sup>24</sup> To be more precise, our results may be documenting different situations of concern. First, we may be documenting an increase in distance between two groups in the country. For example, an increase in the distance between the resource-rich group—whose ethnic identity becomes more salient—and all other groups—whose identities remain unchanged or more leaning towards their nation. These dynamics will lead to a more polarised society, a potentially problematic situation, in particular for the resource-rich group members if the rest of the nation turns against them (e.g., Morelli and Rohner, 2015). Alternatively, we may be documenting an increase in the distance between all ethnic groups in the country, for example, between the Mossi, the Lobi, the Fula, etc. leading to a more fractionalised society.

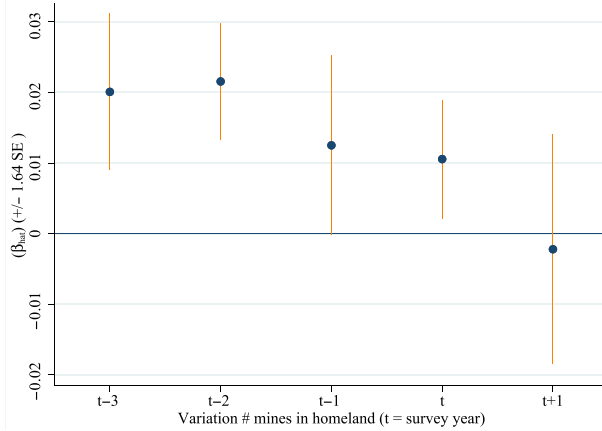


Fig. 2. *Timing of the Effect of Natural Resource Extraction on Ethnic Identification.*

*Notes:* This figure reports the results of an estimation akin to equation (1), except that instead of including the number of active mines at the time of survey, we include variables representing variations in the number of active mines in the year of the survey as well as in the previous and subsequent years. We estimate the following equation:

**Ethnic identification**,  $i,e,r,t = \sum_{l=-1}^3 \alpha^{t-l} \times \Delta \text{mines}_{i,e,(t-l,t-l-1)} + \mathbf{c}'_i \beta + \mathbf{FE}_{e,c} + \mathbf{FE}_{c,r,t} + \varepsilon_{i,e,r,t}$  where  $\Delta \text{mines}_{i,e,(t-l,t-l-1)}$  is the yearly change in the number of active mines in the homeland  $e$  of individual  $i$  interviewed during year  $t$ , considering changes that took place up to three years before the year of the interview and a year after. The figure depicts the coefficients ( $\hat{\alpha}$ ) and 90% confidence intervals.

Figure 2 shows the impact of changes in mining activity taking place from three years before the survey up to a year after the survey. The estimation strategy is the same as in Table 2, column 1, but we replace the variable representing the stock of the number of active mines at the time of the survey with separate variables representing the contemporaneous, past and future changes in the yearly number of active mines.<sup>25</sup> The figure shows estimated coefficients as well as 90% confidence intervals. We find that the effect of changes in mining activity on identification to one's ethnic groups is quite persistent: in fact, the effect gets reinforced over time.

Before the start of mining operations, a phase of one or two years is required for investment (Benshaul-Tolonen, 2019). Locations where a mine started producing the year of the survey could already witness mine-related activity (and most likely media coverage) the year before the start of the production. Yet, as shown in Figure 2 we detect no significant effect of changes in mining activity occurring a year following the survey. Hence, production appears to matter *per se*, not the prospects of production (Arezki *et al.*, 2017).

### 2.2.2. Electoral periods

We pursue an alternative strategy to study persistence: we match our data with information on the timing of presidential elections and estimate whether the proximity to elections affects our

<sup>25</sup> This approach would be equivalent to using leads and lags of the stock if we had yearly data with a large variance in the explanatory variable. However, the Afrobarometer provides a repeated cross-section with gaps of three or four years between surveys rather than a yearly panel. Moreover, about 50% of the groups with changes in the number of mines witness only one change over the survey rounds.



Table 3. *Natural Resources Extraction and Ethnic Identification During Electoral Periods.*

Dep. Var.	(1)	(2)	(3)
	Ethnic group identification		
Version	Categorical		
No. mines: ethnic homeland, in country	0.079 <sup>b</sup> (0.035)		0.227 <sup>a</sup> (0.079)
No. mines: ethnic homeland × Dist. elec.	−0.009 <sup>c</sup> (0.005)		
No. mines: ethnic homeland × Elec. (0–6months)		0.043 <sup>b</sup> (0.018)	
No. mines: ethnic homeland × Elec. (6–12months)		0.022 (0.016)	
No. mines: ethnic homeland × Elec. (12+ months)		0.025 <sup>c</sup> (0.014)	
No. mines: ethnic homeland × Dist. elec. (after)			−0.014 <sup>b</sup> (0.006)
No. mines: ethnic homeland × Dist. elec. (before)			−0.016 <sup>a</sup> (0.006)
No. mines: ethnic homeland, abroad	0.018 <sup>a</sup> (0.005)	0.019 <sup>a</sup> (0.005)	0.018 <sup>a</sup> (0.005)
Fixed effects	Country × Ethnic group, Region × Time		
Individual controls	Yes		
Observations	109,122	109,122	109,122
R <sup>2</sup>	0.187	0.187	0.187
Sample mean/SD dep. var.	2.36/1.19		

Notes: OLS estimations. Significant <sup>a</sup>at 1%; <sup>b</sup>at 5% and <sup>c</sup>at 10%. SEs are clustered at the ethnic group level. The dependent variable is the level of ethnic identification with respect to state identification in all columns, ranging from categories 1 to 5. Controls include age and its square, gender, rural/urban dummy, education categories and employment status. There are 1,767 Admin-1 region × year × month fixed effects and 387 country × ethnic group fixed effects. Sample mean/SD dep. var. are respectively the sample mean and SD of the dependent variable in the corresponding column.

estimates.<sup>26</sup> The underlying idea is that, if the effect is stronger during electoral periods, natural resources might have persistent impacts on ethnic identities through voting patterns, even if their impact is otherwise temporary. In the spirit of Eifert *et al.* (2010), we interact our main variable with the (log-)distance to the closest election (Table 3, column 1). We find that the impact of changes in the number of mines on ethnic identity is larger in proximity to elections. Alternatively, we interact our main variable with dummies equal to one if an election takes place in the 6 months, 6 to 12 months and more than 12 months around the survey. The effect is stronger in the six months around elections and significantly different from the other periods (column 2). Last, the effect of the distance to the election does not depend on whether the closest election took place before or after the interview (column 3).<sup>27</sup>

<sup>26</sup> To construct our measure of proximity to elections, we make use of the recently released Database of Political Institutions 2020 (DPI 2020, available at <https://www.iadb.org/en/research-and-data/dpi2020>). Similar to Eifert *et al.* (2010), we focus on presidential elections, which dates correspond to the dates of executive elections in the DPI 2020, complemented with the date of legislative elections for the three countries from our sample where the assembly elects the president (namely Botswana, Lesotho and South Africa).

<sup>27</sup> Online Appendix, Section A6, shows similar results when we change the definition of the dependent variable for our baseline dummy of identification to ethnic group identities.

### 2.3. Sensitivity Analysis and Additional Results

In this subsection we test the robustness of the baseline estimates of Table 2 to a variety of sensitivity checks.

#### 2.3.1. Alternative level of fixed effects

Our baseline model includes a restrictive set of fixed effects, at the country  $\times$  ethnic group as well as at the region of residence  $\times$  year  $\times$  month level. A natural concern is that the variation in the number of mines left after including all these fixed effects is limited, which complicates the generalisation of the results. Online Appendix, Section A5, investigates this issue. We already discussed previously Table A.3 in the Online Appendix, which shows that the residual variation in the number of mines remains reasonable (a SD of 0.77) after purging the variable from all our control variables and fixed effects. Online Appendix, Figure A.3, shows how the coefficient of interest (Table 2, column 1) varies when including more or less restrictive sets of fixed effects. In all estimations, we include ethnic group  $\times$  country fixed effects to ensure that the coefficient is identified across ethnic groups (hence country fixed effects are de facto included). Two conclusions emerge. First, it is important to control for time-varying local shocks. Controlling only for year dummies leads to insignificant results and a coefficient close to zero. It is likely that both natural resource exploitation and ethnic feelings have a strong time-varying country-specific component.<sup>28</sup> Second, as long as we control for time-varying local shocks, the definitions of ‘time’ (year or year-month) and ‘local’ (country, Admin-1 region, Admin-2 region, or apolitical subdivisions, i.e., cells of  $1 \times 1$  or  $0.5 \times 0.5$  degree of latitude and longitude) do not affect much the precision of the estimates, despite the drastic changes in the number of fixed effects (from 93 when including only country  $\times$  year fixed effects to 1,817 with country  $\times$  region  $\times$  year  $\times$  month and 5,966 with cells of  $0.5 \times 0.5$  fixed effects). This result is consistent with the absence of strong local effects of mining as discussed in Section 3.2, and is reassuring: considering a wider set of variations in the number of mines produces results similar to our baseline, both qualitatively and quantitatively.

#### 2.3.2. Endogeneity of mining activity

As mentioned earlier, variations in minerals’ extraction in a specific homeland could be a function of the ethnic group’s political power, which in turns might correlate with ethnic identification. This could bias our estimates upwards. We consider this concern as quite unlikely given the findings discussed in Section 3 : our estimates appear to be mostly driven by politically powerless groups. Still, to further ensure that our results do not reflect such potential endogeneity bias, we use an alternative measure of mineral resources (Online Appendix, Section A7). For each mineral produced in the ethnic homeland of a respondent, we compute the average pre-sample production (2000–2004, or 2004) evaluated at the yearly world prices of the mineral. Then, we sum across minerals to get the production value in the ethnic homeland  $e$  at year  $t$ :  $\mathbf{Y}_{e,t} = \sum_{k=1}^N \mathbf{Q}_{e,k,t_0} \times \mathbf{P}_{k,t}$ , where  $\mathbf{Q}_{e,k,t_0}$  is the production of mineral  $k$  in the ethnic homeland  $e$  during the reference period  $t_0$  (here the average pre-sample production), and  $\mathbf{P}_{k,t}$  is the real world price of the mineral  $k$  during year  $t$ . The data on (real) world prices come from the World Bank Commodity Dataset. The inclusion of ethnic group  $\times$  country fixed effects absorbs differences in the levels of production

<sup>28</sup> Though smaller in magnitude and less precisely estimated, we do find significant effects when controlling for year  $\times$  month fixed effects only. Yet, year  $\times$  month fixed effects are likely to capture part of country  $\times$  year shocks, as the month of survey is partly country-specific.

at the beginning of the period, hence the identification relies on yearly variations in world commodity prices which are arguably exogenous to local ethnic identification, especially given that the countries in our sample are typically small producers at the world level.<sup>29</sup> A concern with this measure is that individuals have arguably less information about changes in the world value of minerals than about the activity status of a mine.<sup>30</sup> Overall, our results, however, are globally robust to this alternative measure of mining activity (Online Appendix, Table A.5). We find positive coefficients, with  $p$ -values around 0.11–0.16,<sup>31</sup> depending on the columns and definitions. Quantitatively, the effects are close to our baseline estimates: for instance, an SD increase in production in Table A.5, column 1, triggers an increase in ethnic identification equivalent to 1.3% of the SD of the dependent variable (compared to 1.5% for an additional mine in Table 2, column 1).

### 2.3.3. *Weather shocks in ethnic homelands*

Are our results specific to mining or do they reflect a more general effect of income shocks? In the Online Appendix, Table A.6, we add to our baseline estimations measures of homeland-specific weather shocks. We include, sequentially or jointly, measures of aggregate rainfall and average temperature (columns 1 to 3), and measures of anomalies for rainfall and temperature (columns 4 to 6).<sup>32</sup> The estimates are statistically insignificant. These results are useful for two reasons. First, as a robustness, these tests further ensure that our baseline results are not caused by omitted factors that might correlate with mining activity. Second, and more importantly, the fact that variations in climatic conditions do not appear to affect the salience of ethnic identification suggests that our baseline results are indeed driven by specificities of minerals' extraction, rather than economic shocks in general.<sup>33</sup> We come back to the question of economic shocks in the next section.

### 2.3.4. *Placebo*

To further rule out the possibility of false positive in the main results, we run a placebo test as follows: for each year in the sample, we randomly assign the number of mines across ethnic homelands, and re-estimate the specification of Table 2, column 1, with this random variable. We repeat this Monte Carlo procedure in 1,000 draws. Figure A.4 in the Online Appendix plots the sampling distribution of the obtained coefficient, compared to our baseline coefficient drawn as a red vertical axis (0.018). Reassuringly, the Monte Carlo coefficients are distributed far from the baseline estimates.

<sup>29</sup> See Berman *et al.* (2017) for a discussion on this question using similar data.

<sup>30</sup> Even today, local populations may be unaware of some important economic shocks. More than half of the respondents in Armand *et al.* (2020) have limited knowledge about the recent major discoveries of gas offshore of their region. A possible solution could have been to instrument the number of mines with variations in world prices; but this strategy cannot be implemented, because we do not observe the set of minerals that could have been produced in places before the mines opened.

<sup>31</sup> The coefficients in columns 1 and 3 become significant at the 1% level, when we correct for spatial correlation in the error term instead of our baseline specification (clustering SEs at the level of the ethnic homeland).

<sup>32</sup> Temperature and rainfall information are aggregated at the Murdock homeland level from the Climatic Research Unit-gridded time series that provide climate dataset at a resolution of  $0.5 \times 0.5$  latitude and longitude (Harris *et al.*, 2020). Data from <https://crudata.uea.ac.uk/cru/data/hrgr/>. Introducing the square of these weather variables does not change the conclusion of these estimations.

<sup>33</sup> Weather shocks are significant economic shocks (Barrios *et al.*, 2010; Dell *et al.*, 2012). In our sample, and in line with existing literature, we find that rainfall shocks in the historical homeland of a respondent improves the material conditions of that respondent (measured through the wealth index, Online Appendix, Table A.7).

### 2.3.5. *Sensitivity to specific groups, countries, time periods and local effects*

In the Online Appendix, we show that our results are not driven by specific subsets of observations. We first drop countries and survey waves one by one (Figure A.5 and Table A.8, respectively). The results are remarkably stable. Figure A.6 displays the estimated coefficients for each ethnic group separately: 61% of the coefficients are positive (70% if we only consider the set of coefficients statistically significant at 10%). We come back to ethnic group heterogeneity in Section 3.3. Second, we investigate the robustness of our baseline results to the exclusion of outliers (Table A.9).<sup>34</sup> Last, given the spatial dimension of the data, we allow the error term to be spatially correlated and auto-correlated. The SEs of our variable of interest remains stable (Table A.10); for instance, the SE of Table 2, column 1, remains in the interval 0.004–0.006 (compared to 0.003 in the baseline), depending on the spatial radius chosen (up to 1,000 km).<sup>35</sup>

### 2.3.6. *Alternative matching of ethnic groups' homelands*

As our results rely on the match between the historical ethnic homelands recorded in the Murdock Atlas and today Afrobarometer groups, we replicate our baseline estimates for the three alternative definitions of ethnic homelands discussed in Section 1.2. First, the broader cultural groups recorded in the Murdock Atlas, which ensure an exhaustive geographical coverage. Second, the LEDA matching from Afrobarometer ethnic groups to Murdock ethnic homelands, which relies on the Ethnologue language tree. Third, we turn to a matching of the Afrobarometer ethnic groups to the contemporaneous ethnic homelands as recorded in the EPR. With few exceptions, our results are qualitatively unchanged, and quantitatively similar for the LEDA matching, but coefficients are smaller for two alternative definitions (Online Appendix, Tables A.11 and A.12). This is because the size of the homelands—and therefore the variation in the number of mines contained in them—varies across definitions. Using our baseline specification with the categorical measure of ethnic identification (Online Appendix, Table A.11), we find that an increase of one additional mine raises ethnic identification by 0.005–0.007 when using the cultural group definition; 0.017–0.018 when using LEDA; and 0.007–0.009 when using EPR. But considering an SD increase in the number of mines (after purging the variable from all fixed effects and controls) leads to similar results across columns: an increase in ethnic identification by 0.011 (cultural groups), 0.013 (LEDA) or 0.009 (EPR) to be compared with 0.013 in our baseline.

### 2.3.7. *Strength of an identity versus changing identity*

Finally, note that all our estimations rest on the ethnic groups that the respondents declare. This declaration, i.e., the ethnic identity of the respondent, could itself react to natural resources. In particular, if natural resources benefit certain groups, it might create some incentives for individuals to change their ethnic identities. Although ethnic identities are usually considered to

<sup>34</sup> In columns 1 and 2, we consider the log and the inverse hyperbolic sine transformation of the number of mines in the ethnic homeland of the respondent. From columns 3 to 5 we exclude observations that are 3, 2 and 1 SD away from the residual mean. Then, we exclude observations with high leverage, i.e., when the individual leverage is superior to  $2k = N$  (with  $k$  being the number of predictors and  $N$  the number of observations, column 6). Finally, we exclude observations that shift the estimate at least to one SE (dfbeta; column 7), and observations that shift the estimate at least to  $4/N$  (Cook's distance; column 8).

<sup>35</sup> We use the Stata routine developed by Colella *et al.* (2020). More precisely, we apply a spatial HAC correction to our SEs, allowing for both cross-sectional spatial correlation and location-specific serial correlation (Conley, 1999). Imposing no constraint on the temporal decay for the Newey-West/Bartlett kernel that weights serial correlation across time periods, the horizon at which serial correlation is assumed to vanish can be infinite (i.e., 100,000 years). In the spatial dimension, we estimate the SEs using alternatively thresholds at 250, 500, 750 or 1,000 km.

be inherited at birth and quite stable, the literature has documented that ‘passing’ may respond to economic incentives (Cornwell *et al.*, 2017; Dahis *et al.*, 2019), and that even ethnic identities can be prone to noisy signalling and manipulation in a lab setting (Harris *et al.*, 2018). We check whether, additionally to the salience of the group identity, mineral resources affect the endogenous choice of ethnic identity, in the sense of passing from one group to another. We perform two different tests (Online Appendix, Table A.13). First, we estimate how mining activity in the ethnic homeland of a group affects the population share of that group at the country level. For each ethnic group  $\times$  survey year, we compute the share of respondents that self-report belonging to that group. We regress this variable on the number of mines in the historical homeland of the group, as in our baseline estimations, also including ethnic group  $\times$  country and country  $\times$  year fixed effects. The estimates are close to zero and statistically insignificant across all four different definitions of ethnic groups (columns 1 to 4). Second, we estimate whether resources exploitation in an ethnic homeland affects the probability for the respondent to choose that ethnic group among all groups present in the country. We include the same set of fixed effects as in our baseline estimates. Column 5 shows a small negative estimate which is significant at the 5% level, but the number of observations is very large (about 1.8 million). Hence, overall, the results suggest that mining activity affects the strength of ethnic identities, but do not significantly causes changes from one ethnic identity to another. These results must, however, be interpreted with caution, as our data are not the most well-suited to estimate the determinants of changing identity.

### 3. How Mining Affects Ethnic Identity

We find that an increase in mineral resources exploitation strengthens feelings of ethnic identity for the group whose historical homeland hosts the resource. In this section, we provide a set of empirical exercises which globally support an interpretation of this finding rooted in the impact of natural resources on ethnic groups’ grievances. At the core of our argument are the ethnic group-specific feelings of ‘relative deprivation’. Inequality, real or perceived, generates discontent which increases the salience of identities (Olzak, 2011). This inequality may, for example, arise as the economic or political benefits of natural resource extraction are not disproportionately perceived to fall on local indigenous groups, or whenever natural resources are seen to cause a degradation of these groups’ status or welfare. Building on existing work and case studies presented below, our empirical approach considers the cycle between identity fragmentation, deprivation and exclusion, which may arise either in a decentralised manner or as a result from manipulation by ‘ethnic political entrepreneurs’ (Aspinall, 2007).

#### 3.1. *Conceptual Background and Examples*

##### 3.1.1. *Relative deprivation*

At the individual level, feelings of relative deprivation—that is, the discrepancy between aspirations and reality—can be grounded either in inaccurate aspirations or in a detrimental economic reality. The start of industrial mining activities may raise aspirations related to economic development and resource rents (Collier, 2017), but, under incomplete information, these aspirations may be unrelated to the reality of resources’ extraction. Cust and Mihalyi (2017), for instance, show how natural resources discoveries trigger expectations unrelated to their economic reality, including among actors specialised in macroeconomic projections. At the individual level,

Christensen (2019) argues that the link between mining activity and riots is rooted in incomplete information and (mis)perception about mining projects' profitability. This can be translated to the question of ethnic identities. Binzel and Carvalho (2017) show how identities can evolve in response to individual feelings of relative deprivation. Taking the example of the Egyptian revolution, they argue that economic development raises expectations, but does not always fulfil these. A religious identity then allows individuals to alter their reference point and to cope with these unfulfilled aspirations. They conclude that economic development may thus paradoxically make societies more prone to religious revivals. In our context, as mines' openings are taking place within the historical homeland of ethnic groups, we expect the discrepancy between aspirations and reality to be strongest for the members of these groups.<sup>36</sup> The aspirations and deprivation feelings of all the members of the resource-rich ethnic groups may increase due to economic linkages, which would be amplified in case of a local multiplier effect (Moretti, 2010), or through an ethnic multiplier effect (being embedded in credit trade or migration networks with their co-ethnics, or through sheer in-group bias and solidarity; Fafchamps, 2000; 2003; Habyarimana *et al.*, 2009; Aker *et al.*, 2014; Iwanowsky, 2018). The Afrobarometer surveys include questions related to deprivation experiences and the perception of living conditions, which we use to test the grievances channel directly.

### 3.1.2. *The role of leaders*

Ethnic leaders, who range from traditional ethnic leaders—still a key local institution in sub-Saharan Africa (Michalopoulos and Papaioannou, 2020)—to national political representatives (Aspinall, 2007), may create or magnify individual feelings of relative deprivation along ethnic lines. Such leaders may have an interest in doing so to divert attention from their own rent extraction, or to use natural resource exploitation and ethnic discrimination in an opportunistic way in their political agenda, e.g., to gain popular support (Cheeseman and Larmer, 2015). The movement of Michael Sata in Zambia illustrates how a political leader exploits economic grievances related to mineral exploitation, and articulates them along the lines of ethnic identity. Michael Sata practised 'ethnopolitism' (Cheeseman and Larmer, 2015), building on the grievances of the Bemba identity group in the Zambian Copper belt.<sup>37</sup> We lack sufficient data to directly test this interpretation; however, we note that, although it is quite indirect, our finding that the relation between resources and identity is magnified during electoral periods, is consistent with 'ethnic political entrepreneurs' creating or instrumenting ethnic-based grievances in their campaigns.

### 3.1.3. *Natural resources, conflict, and ethnic identities*

Further anecdotal evidence suggests that leaders may be particularly likely to successfully capitalise over individual ethnic identity feelings in areas where feelings of ethnic inequality are pre-existing, such as in poor regions hosting resources, politically excluded ethnic groups or in areas with a history of violent conflicts (Aspinall, 2007; Must, 2018). According to Aspinall (2007), who studies the Free Aceh movement in Indonesia, 'rather than any intrinsic qualities

<sup>36</sup> Expectations should increase most among people who feel closest to the resource, here the members of the resource-rich ethnic group. In a context where work identity is key, in Kazakhstan, Girard *et al.* (2021) show that expectations increase most among workers from the oil sector (compared to workers from other sectors) during an oil boom.

<sup>37</sup> In the 2000s, 'One consequence of the minerals boom was rising copper prices, which enabled Zambia's newly privatized mining companies to make vast profits, increasing public discontent with falling wages and living standards. [Michael Sata's] rhetoric chimed with an enduring perception that wealth produced by workers of Bemba origin was being illegitimately consumed by an incumbent ruling elite, but also with an ethnically diverse urban constituency in [the capital]' (Cheeseman and Larmer, 2015). See Section A14 in the Online Appendix for more details on the country's background.

of natural resource extraction, the key factor was the presence of an appropriate identity-based collective action frame'. Aspinall argues that claims about the unjust exploitation of natural resources made by ethnic political entrepreneurs resonated powerfully in the population, because they reinforced the long-standing discourse of deprivation which infused Acehnese identity by the 1970s.<sup>38</sup> The qualitative work of Must (2018) suggests that similar forces were at play in Tanzania after large gas discoveries.<sup>39</sup> This literature suggests that leaders' capitalisation over ethnic identity may go as far as triggering ethnic conflicts, suggesting that identity exacerbation might be a missing link in the natural resources–conflict literature. The timing of the conflict–ethnic identity relationship is beyond the scope of this paper—one may follow the other, and both are likely to reinforce each other. However, this literature also suggests that pre-existing grievances and identity feelings, born from political exclusion, a history of conflict or long-standing poverty, may exacerbate the effect of natural resources on ethnic identity, in particular because these can be manipulated by ethnic political entrepreneurs. We take this prediction to the data at the end of this section, making use of information contained in the EPR dataset on political exclusion, and of data on local conflict and poverty. Section A14 in the Online Appendix contains a more detailed description of the cases of Zambia, Indonesia and Tanzania.

#### 3.1.4. *Competing channels and testable predictions*

The main competing explanation of our findings is that mineral resources make ethnic identities more salient because they increase the pay-offs associated with being part of a particular group. This could be the case if minerals have strong positive local economic effects (Aragón and Rud, 2013; Cust and Poelhekke, 2015; Benschaul-Tolonen, 2019; Mamo *et al.*, 2019) that are felt disproportionately by historical local groups; or if they increase the political power and representation of these groups (Mamo and Bhattacharyya, 2018), or political patronage along ethnic lines (De Luca *et al.*, 2018; Dickens, 2018). The actual pay-offs of belonging to a given group would increase, which would make identities more salient. Combined with intra-ethnic spillovers due to migration, trade, remittances or solidarity, this mechanism also explains why individuals residing outside their homeland display stronger identity feelings. This 'pay-off' interpretation generates predictions which differ from the 'deprivation' channel, and can be tested by looking at how minerals' extraction affect individual wealth and whether the wealth of indigenous groups respond differently to mining.

### 3.2. *Mining, Wealth, and Feelings of Economic Deprivation*

#### 3.2.1. *Baseline*

We first consider a number of Afrobarometer measures to study how individual wealth and *perceptions* of economic conditions react to mineral exploitation. To distinguish local effects from ethnic spillovers, the coefficients are estimated, either on our baseline full sample, or on the sample of individuals residing outside their ethnic group's historical homeland. We proxy the objective material wealth of the household computing a 'family of outcomes' index as the

<sup>38</sup> Indonesia comprises different resource-rich provinces (Aceh, Riau, East Kalimantan), which all experienced virtually identical processes of natural resource exploitation. However, the context of historical violence and institutionalisation of ethnicity made Aceh distinct. As a result, only Aceh saw the development of ethnic tensions and intense separatism.

<sup>39</sup> Despite large regional marginalisation, Tanzania remained peaceful for decades and an example of national integration (Green, 2013). Must (2018) insists on the role of leaders, perceived deprivation and identities, concluding that 'natural resource mismanagement and subsequent leadership framing increased the salience of a regional identity and exacerbated felt group grievances in southern Tanzania'.

mean of the standardised variables using questions on whether the respondent personally owns a radio, a TV, and/or a vehicle (Kling *et al.*, 2007).<sup>40</sup> Under the assumption that mineral resources' exploitation affects more (positively) the members of the ethnic group where mines are located, we would expect respondents to report owning more of those goods. To measure economic deprivation, we build two different indicators. For the first, we compute a 'family of outcomes' index based on five different questions, in which individuals report whether their household has lacked over the past year essential welfare-related items: food, cash income, clean water, medicine or fuel to cook.<sup>41</sup> The second indicator uses information on individuals' perceptions of changes in their living conditions. Individuals are asked about how their living conditions today compare to their living conditions 12 months before.<sup>42</sup> The last two variables have both an objective component—the individuals' actual wealth—and a subjective one, which reflect the individuals' perception of well-being compared to some reference point (see e.g., Ravallion, 2014). If resources exploitation do not alter the respondents' perceptions, we would expect all three variables to have a consistent relation to resources exploitation. Results will, however, diverge if resources exploitation affect respondents' perceptions and heighten deprivation feelings independently of reality—as only the first variable is a purely objective measure of wealth.

The results are provided in Table 4. Note that, contrary to the existing literature, these regressions do not estimate the overall effect of mining on wealth or welfare, given the presence of region  $\times$  time fixed effects. Rather, they tell us whether the effects of mining are disproportionately felt by individuals belonging to the ethnic groups whose homeland hosts the mines—we come back to this issue in the next subsection. Several interesting findings arise. First, the effect of mineral resources on individual wealth is limited. The coefficient is positive, with a  $p$ -value of 0.096 in column 1, and both the magnitude and significance of the coefficient decrease in the sample of individuals residing outside their homeland ( $p$ -value = 0.21, column 2). Quantitatively, an additional mine increases wealth by around 0.5%–0.8% of the SD of the wealth index, which, to put things in perspective, represents between 3% and 4% of the effect of gender, and between 2% and 3% of the effect of being inactive (see Table A.14 in the Online Appendix for the estimates of the control variables' coefficients). Second, results are more stable when using self-reports of economic deprivations and economic conditions (columns 3 to 6): in all columns individuals report worse outcomes, though the statistical significance varies in the case of perceptions of economic conditions. An additional mine implies a change equivalent to 0.6–1.1% of the SD of the deprivation variable, i.e., between 30% and 72% of the effect of gender, and between 5% and 10% of the effect of being inactive. Third, though individuals living outside their homeland do not report significant changes in wealth (column 2), they report more deprivation (columns 4 and 6). Taken together, these results suggest that, for co-ethnics who do not live in their historical homeland, mines opening translates into larger deprivation feelings, possibly because those individuals, despite expecting benefits, do not enjoy the same

<sup>40</sup> The precise question is the following: '*Which of these things do you personally own: [Radio] / [TV] / [Vehicle]?*'. Variables are standardised by subtracting their mean and dividing by their SD. Though the resulting variable is only a proxy of actual wealth, it correlates well, and as expected, with weather shocks (Section A8 in the Online Appendix).

<sup>41</sup> The precise question is the following: '*Over the past year, how often, if ever, have you or anyone in your family gone without: [Food] / [Cash income] / [Enough clean water for home use] / [Medicines or medical treatment] / [Enough fuel to cook your food]?*'. Answers are the following: 0 = Never, 1 = Just once or twice, 2 = Several times, 3 = Many times, 4 = Always.

<sup>42</sup> The original question is the following: '*Looking back, how do you rate the following compared to 12 months ago: Your living conditions?*'. Answers are the following: 1 = Much worse, 2 = Worse, 3 = Same, 4 = Better, 5 = Much better.



Table 4. *Natural Resources Extraction, Wealth, and Deprivation.*

Dep. Var. Sample	(1)	(2)	(3)	(4)	(5)	(6)
	Wealth index		Deprivation index		Economic conditions	
	Full	Outside homeland	Full	Outside homeland	Full	Outside homeland
No. mines: ethnic homeland	0.006 <sup>c</sup> (0.004)	0.004 (0.003)	0.004 <sup>b</sup> (0.002)	0.008 <sup>a</sup> (0.003)	-0.004 (0.004)	-0.009 <sup>c</sup> (0.004)
Residence in homeland	-0.028 <sup>a</sup> (0.006)		0.020 <sup>b</sup> (0.009)		0.003 (0.011)	
Fixed effects	Country × Ethnic group, Region × Time					
Individual controls	Yes					
Observations	117,773	63,813	117,786	63,819	80,665	43,436
R <sup>2</sup>	0.326	0.337	0.263	0.283	0.154	0.154
Sample mean/SD dep. var.	-0.018/0.71	0.018/0.72	-0.004/0.72	-0.029/0.72	2.95/1.05	2.99/1.035

Notes: OLS estimations. Significant <sup>a</sup>at 1%; <sup>b</sup>at 5% and <sup>c</sup>at 10%. SEs are clustered at the ethnic group level. In columns 2, 4 and 6, we restrict the sample to individuals living outside their ethnic homeland. The dependent variable in columns 1 and 2 is an index of wealth based on ownership of radio, television and vehicles. In columns 3 and 4, it is the deprivation index, based on whether the respondent declares having gone without enough of the following items over the last year: food, income, water, medicine, fuel. Finally, in columns 5 and 6, the dependent variable is a categorical variable measuring the respondent's perception of their own living conditions compared to 12 months before. Even numbered columns restrict the sample to individual living outside their ethnic homeland. Controls include age and its square, gender, rural/urban dummy, education categories and employment status. Sample mean/SD dep. var. are respectively the sample mean and SD of the dependent variable in the corresponding column.

direct benefit from the natural resources in comparison to individuals who live in their historical homeland.

### 3.2.2. *Mines' characteristics*

At this stage, these interpretations are slightly speculative, because our specification may be too rough to correctly identify the local economic effects of mining: the effects of mines may depend on technological characteristics (Pelzl and Poelhekke, 2021); it may also be very localised (Cust and Poelhekke, 2015), and vary non-linearly with the distance to the mines (De Haas and Poelhekke, 2019). Failing to consider these elements might lead our estimates to be noisy. To go further, we start by studying the role of mines' characteristics. Next, we consider a different specification which allows to measure more precisely the local effects of mines, depending on the distance between the respondents and the resources exploitation.

Our data contains information on two types of characteristics that might affect the impact of mines on ethnic identities and economic spillovers. The first is the ownership structure of mines. Mines operated by large multinational companies might behave differently from large domestic or state-owned mines (e.g., Berman *et al.*, 2017). Even if they do not, the very fact that a mine is domestic might affect how individuals perceive this mine and expect benefits from it. Second, our data also allows to proxy production technology and its labour/capital intensity. The recent literature has documented a Dutch disease type of effect of mineral production which depends on the technology of mining production (Pelzl and Poelhekke, 2021). Capital intensive mining tends to benefit more the non-mining sectors because they cause no upward pressure on wages, contrary to labour-intensive mining which tends to be associated with a reduction in employment. Online Appendix, Section A16, studies how these characteristics affect the link between mines and our various outcomes. We find a slightly stronger effect of domestic mines on the salience of ethnic identity. In addition, and consistent with Pelzl and Poelhekke (2021), relatively capital

intensive mines may be associated with more improvements in the wealth index than relatively labour-intensive mines ( $p$ -value of labour-intensive mines at 0.106). In both cases, however, the coefficients are noisy and, overall, we cannot conclude to statistically significant differences across mines' types in most specifications. We interpret these findings as suggestive that the existence of local economic effects—and hence the pay-offs associated with belonging to the local ethnic group—do not appear to be prime determinants of ethnic identification.

### 3.2.3. *Mines' local effects*

We go further by estimating the local effects of mines on wealth and deprivation feelings through a different econometric strategy. In the spirit of De Haas and Poelhekke (2019), our identification strategy exploits variations in the distance between the respondents and the mines. We first count the number of active mines located within a certain radius around each individual's residence, namely, 0–50 km and 51–150 km.<sup>43</sup> We consider separately mines which, within this radius, are located in the ethnic homeland of the individual and those which are not. We then estimate the impact of this set of mining activity variables on our proxies for wealth and economic deprivation, controlling for fixed effects at the country  $\times$  ethnic group level, and at the country  $\times$  time level. We use these aggregated fixed effects because our aim here is to estimate the average local effects of mines across all individuals (rather than to identify whether this effect varies across groups). As this approach prevents us to include the local time-varying fixed effects and makes endogeneity concerns stronger than in our baseline estimates, we also report results using mineral price variations to identify changes in mineral rents in Online Appendix, Table A.16.

Table 5 displays the estimated coefficients for the full sample of individuals (columns 1, 3 and 5) and for the subsample of individuals residing in the historical homeland of their ethnic group (columns 2, 4 and 6). The results confirm our previous interpretation: local wealth effects are limited (columns 1 and 2). Though most coefficients are positive, statistical significance is weak. More importantly, we fail to find any additional effect of resource extraction when the mine is within the homeland of the individual; jointly, the wealth effects appear significantly smaller in this case (column 2). Strikingly, while overall deprivation feelings seem to slightly decrease, and perceptions of economic conditions appear to improve in proximity to mines (columns 3 to 6), it is the opposite for mines located in the historical homeland of the individual. Put differently, mines increase deprivation feelings only for individuals belonging to resource-rich ethnic groups, those who host the mines in their historical homeland. These results are largely confirmed—and statistically reinforced—when using mineral price variations to identify changes in mineral rents (Online Appendix, Table A.16). Overall, these findings are consistent with the idea that resources windfalls do not appear to be felt by the ethnic group whose historical homeland hosts the resource. Rather, resource-rich individuals may feel economically deprived, in particular individuals living in their historical homeland—or leaders in these homelands may use mines to foster such feelings—and these perceptions might be shared by or transmitted to co-ethnics living outside these homelands.

<sup>43</sup> The cut-off at 150 kilometres is similar to the one used by De Haas and Poelhekke (2019). We use a minimum radius of 50 kilometres around the mines instead of the 20 kilometres used by De Haas and Poelhekke (2019). Four percent of the sample in Table 4, column 1, lives within 50 kilometres of a mine which is in their homeland, and 1% lives within 20 kilometres of such a mine. If we split by residence in the homeland to estimate the coefficients in column 2, 1.5% of respondents lives within 50 kilometres of a mine that is not in their homeland, and 0.1% of the sample lives within 20 kilometres of such a mine (87 individuals). We thus opt for cut-offs at 50 and 150 kilometres.

Table 5. *Natural Resources Extraction: Local Effects.*

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)
	Wealth index		Deprivation index		Economic conditions	
Sample	Full	Residents	Full	Residents	Full	Residents
No. mines (0–50km)	0.002	0.038 <sup>b</sup>	–0.002	–0.027	–0.000	0.016
excl. homeland	(0.002)	(0.017)	(0.002)	(0.017)	(0.004)	(0.039)
No. mines (50–150km)	0.000	0.004	0.001	–0.005	0.002 <sup>c</sup>	0.011 <sup>b</sup>
excl. homeland	(0.000)	(0.005)	(0.001)	(0.003)	(0.001)	(0.005)
No. mines (0–50km)	0.001	–0.006	0.007	0.015	–0.001	–0.003
homeland	(0.006)	(0.006)	(0.009)	(0.012)	(0.007)	(0.011)
No. mines (50–150km)	0.002	0.007	0.004 <sup>a</sup>	0.007	–0.002	–0.010
homeland	(0.002)	(0.006)	(0.001)	(0.006)	(0.002)	(0.011)
Residence in homeland	–0.027 <sup>a</sup>		0.026 <sup>b</sup>		0.001	
	(0.006)		(0.013)		(0.015)	
Diff. in hom./	0.001	–0.041 <sup>c</sup>	0.012	0.053 <sup>a</sup>	–0.004	–0.040
outside hom.	(0.008)	(0.022)	(0.010)	(0.020)	(0.008)	(0.029)
Fixed effects	Country × Ethnic group, Country × Time					
Individual controls	Yes					
Observations	117,930	53,915	117,943	53,922	80,725	37,169
R <sup>2</sup>	0.295	0.295	0.198	0.202	0.103	0.124
Sample mean/SD dep. var.	–0.017/0.71	–0.060/0.7	–0.004 / 0.72	0.026/0.71	2.96 / 1.04	2.9/1.06

Notes: OLS estimations. Significant <sup>a</sup> at 1%, <sup>b</sup> at 5% and <sup>c</sup> at 10%. SEs are clustered at the ethnic group level. In columns 2, 4 and 6, we restrict the sample to individual living inside their ethnic homeland. The dependent variable in columns 1 and 2 is an index of wealth based on ownership of radio, television and vehicles. In columns 3 and 4, it is the perceived deprivation index, based on whether the respondent declares having gone without enough of the following items over the last year: food, income, water, medicine, fuel. In columns 5 and 6, the dependent variable is a categorical variable measuring the respondent’s perception of their own living conditions compared to 12 months before. Controls include age and its square, gender, rural/urban dummy, education categories and employment status. Sample mean/SD dep. var. are respectively the sample mean and SD of the dependent variable in the corresponding column.

### 3.3. Political Exclusion and Ethnic Group Characteristics

As discussed in Section 3.1, though the mechanisms discussed above can in principle apply to many contexts, we expect them to be particularly relevant when feelings of ethnic inequality are pre-existing. This would be the case for politically excluded ethnic groups, for groups with a strong pre-existing sense of ethnic identity, and for those originating from (relatively) poor areas or areas with a history of violent conflict (Aspinall, 2007; Vogt, 2017; Must, 2018). Differences across ethnic groups along these dimensions may contribute to explain why we find a substantial amount of heterogeneity in the link between mining and ethnic identification (Figure A.6). In this section, we take these predictions to the data.

#### 3.3.1. Political exclusion

We start by investigating the role of political exclusion (Table 6). We make use of specific information about the political power of ethnic groups from the Ethnic Power Relations Dataset (EPR, Wucherpfennig *et al.*, 2011; Vogt *et al.*, 2015), which records contemporaneous ethnic homeland boundaries as well as time-varying political power. The aim of the EPR data is to focus on ethnic groups who have a clear political relevance in their countries, leaving many groups outside its sample.<sup>44</sup> As a preliminary check, we thus replicate the results of the baseline

<sup>44</sup> It may even be entire countries which do not appear in the EPR data. Vogt *et al.* (2015, p. 1329), note that ‘An ethnic group is considered politically relevant if at least one political organization has claimed to represent its interests at

Table 6. *Natural Resources Extraction and Political Exclusion.*

Dep. Var.	(1) Ethnic identification	(2) Group treated unfairly	(3) Wealth index	(4) Deprivation index	(5) Economic conditions
No. mines: ethnic homeland × powerful	0.009 (0.008)	0.022 <sup>a</sup> (0.006)	-0.003 (0.002)	0.000 (0.002)	-0.006 (0.006)
No. mines: ethnic homeland × powerless	0.186 <sup>a</sup> (0.021)	0.333 <sup>a</sup> (0.064)	-0.014 (0.015)	0.033 (0.022)	-0.060 <sup>b</sup> (0.028)
Residence in homeland	0.058 <sup>a</sup> (0.018)	0.064 <sup>a</sup> (0.015)	-0.010 (0.007)	0.020 <sup>b</sup> (0.008)	-0.001 (0.013)
Fixed effects	Country × Ethnic group, Region × Time				
Individual controls	Yes				
Observations	90,053	86,026	91,107	91,116	62,088
R <sup>2</sup>	0.178	0.235	0.345	0.282	0.150
Sample mean/SD dep. var.	2.35/1.2	2.36/1.2	0.04/0.74	-0.05/0.73	2.99/1.065
Test differences ( <i>p</i> -value)	0.000	0.000	0.443	0.135	0.041

Notes: OLS estimations. Significant <sup>a</sup> at 1%; <sup>b</sup> at 5% and <sup>c</sup> at 10%. SEs are clustered at the ethnic group level. The dependent variable in column 1 is the level of ethnic identification with respect to state identification in all columns, ranging from categories 1 to 5. In column 2, the dependent variable tells if the respondents considers that their ethnic group has been treated unfairly by the government, ranging from categories 0 (never) to 3 (always). The dependent variable in column 3 is an index of wealth based on ownership of radio, television and vehicles. In column 4, it is the perceived deprivation index, based on whether the respondent declares having gone without enough of the following items over the last year: food, income, water, medicine, fuel. Finally, in column 5, the dependent variable is a categorical variable measuring the respondent's perception of their own living conditions compared to 12 months before. Ethnic homeland and ethnic groups are based on data from EPR. Powerful groups which fall into one of the following EPR categories in the EPR dataset at the beginning of the period: dominant, junior and senior partners. Powerless are groups will fall in the following categories: discriminated, irrelevant, powerless. We separately control for mines abroad. Controls include age and its square, gender, rural/urban dummy, education categories and employment status. Test differences (*p*-value) tells the *p*-value of the test that the effect of the number of mines in the homeland of a powerful group is equal to the effect of the number of mines in the homeland of a powerless group. Sample mean/SD dep. var. are respectively the sample mean and SD of the dependent variable in the corresponding column.

Table 2, after restricting the sample to groups which *do not* appear in the EPR sample. The magnitude of the estimates of mineral resources extraction in an ethnic homeland is systematically higher (more than three times) in the restricted sample than in the baseline sample (Table A.17 in the Online Appendix). In other words, the effect of mineral resources on ethnic identification appears to be drastically magnified in the sample of groups without any ethnic political representation.

We then exploit the information present in the EPR dataset. We interact the mining variable with dummies capturing the degree of political power of the group, allowing heterogeneous effects by group political power (compared to the aggregate effect presented in the Online Appendix, Table A.11). While all groups of the EPR sample are politically relevant, they differ in their ability to influence political decisions. We split the groups into two mutually exclusive categories: powerful groups, which have some access to power, and powerless ones, which do not.<sup>45</sup> Because the EPR status may be endogenous to variations in mineral resource production, we use the political status

the national level or if its members are subjected to state-led political discrimination). Group identities may be salient in everyday lives, like in wedding or business networks, while still not having a formal political representation. In Burkina Faso, for example, the share of interethnic marriages among married couples in 2003 was only 11.5%, while this rate would have been 67.5% if marriage matches were orthogonal to ethnic group identities (Crespin-Boucaud, 2020, online appendix, table 2). Yet, Burkina Faso is a country where the EPR dataset does not record a single ethnic group.

<sup>45</sup> We include the following EPR categories in the groups with access to power: dominant, junior partner and senior partner. A total of 84% of the Afrobarometer respondents belong to a group with access to power (70 groups), while 16% of the respondents belong to a powerless group (24 groups). The powerless groups contain the following EPR categories: discriminated, irrelevant, powerless. The other categories (e.g., monopoly, self-exclusion) are never present in

of groups before the start of our study period, in 2004. We find that hosting mineral resources production has no significant effect on the level of ethnic identification of groups with access to power (Table 6, column 1). However, the effect of resources is large and significant for powerless groups (Table 6, column 1)—an order of magnitude higher than in our baseline estimates (Table 2, column 1).<sup>46</sup>

In column 2, we assess more directly the impact of mining activity on the sharing of power across ethnic groups, as perceived by local households. We make use of an Afrobarometer question in which individuals are asked whether they believe that their ethnic group is treated unfairly by the government.<sup>47</sup> We find a positive and significant effect of mining activity on perceptions of unequal ethnic group treatment, which is again quantitatively much stronger for individuals belonging to powerless groups. We interpret this result as a strong signal that mineral exploitation increases identity-based feelings of deprivation, be these feelings linked to pure economic motives, or to a drop in the trust that individuals have in their national government.

Finally, in Table 6, columns 3–5, we show that the effects found in the previous section on economic wealth and deprivation indexes are reinforced in the case of powerless groups. Though neither powerless nor powerful groups' wealth appear to improve with mining activity (column 3), powerless groups exhibit stronger increases in deprivation (column 4;  $p$ -value = 0.13) and more pessimism about economic conditions (column 5). The extraction of mineral resources in the homeland of politically powerful ethnic groups has a smaller and always statistically insignificant effect on the various outcomes.

### 3.3.2. *Conflict and poverty*

Finally, we consider the exacerbating role of poverty, conflict and pre-existing ethnic feelings (Table 7). We construct three ethnic group-specific measures. The first investigates the consequence of pre-existing conflicts.<sup>48</sup> We use information from the Armed Conflict Location Events Data (ACLED; Raleigh *et al.*, 2010) on the location, date and types of conflict events, to build a control for the level of conflicts taking place in each homeland and year. For each ethnic homeland, we compute the number of battle events and other conflict events recorded in ACLED over the period 2000–2004 (before our sample starts). Second, we proxy development level using nighttime luminosity data, which have been widely used by recent literature as a proxy for local income

our sample. The EPR categories definition are as follows. '1. The group rules alone: monopoly or dominant. In contrast to monopoly power, the status of dominant indicates "token" representation of other ethnic groups in the executive. 2. The group shares power: senior partner or junior partner, depending on the group's absolute influence in the executive (i.e., irrespective of group size). 3. The group is excluded: powerless, discriminated, or self-exclusion. While powerless means that the group is simply not represented (or does not have influence) in the executive, discrimination indicates an active, intentional, and targeted discrimination by the state against group members in the domain of public politics' (Vogt *et al.*, 2015, p. 1331).

<sup>46</sup> Note that we split the effect for the number of mines in the ethnic homeland in the country of the respondent, because our sample mines located abroad are systematically in the homelands of groups who have access to political power. This lack of variation prevents us from identifying the heterogeneous effect of mines abroad by access to power. However, it is interesting to note that the coefficients have closer magnitude for resources abroad and resources in the homeland of a powerful group, than for resources in the homeland of a powerless group.

<sup>47</sup> The question is stated as: '*How often is [Respondent's Ethnic Group] treated unfairly by the government*'. Answers are the following: 0 = Never, 1 = Sometimes, 2 = Often, 3 = Always.

<sup>48</sup> The timing of the relation between conflict and identity is subject to debate, as conflicts may exacerbate existing identities (Atkin *et al.*, 2021), but existing identities can also be used to mobilise conflict participants (Esteban *et al.*, 2012). We acknowledge that both contemporaneous conflicts and contemporaneous identities may be jointly determined, potentially one and only one thing, and prevent us from distinguishing the timing or the relevance of each force within our baseline results. Rather, what we do here is to focus on how pre-existing conflicts and other pre-existing conditions related to relative deprivation interact with mineral exploitations.

Table 7. *Ethnic Group-Specific Factors.*

Dep. Var. Sample	(1)	(2)	(3)	(4)	(5)	(6)
	Salience of ethnic identification					
	Full			Outside homeland	Full	Outside homeland
No. mines: ethnic homeland	0.018 <sup>a</sup> (0.003)	0.029 <sup>a</sup> (0.005)	0.027 <sup>a</sup> (0.005)	0.017 <sup>a</sup> (0.006)	-0.008 (0.013)	-0.014 (0.012)
× no. battles <sub>t0</sub>	0.016 <sup>a</sup> (0.006)		0.012 <sup>c</sup> (0.006)	0.018 <sup>a</sup> (0.006)		
× no. other conflicts <sub>t0</sub>	-0.000 <sup>a</sup> (0.000)		-0.000 (0.000)	-0.000 (0.000)		
× night-time lights <sub>t0</sub>		-0.009 <sup>a</sup> (0.003)	-0.008 <sup>c</sup> (0.004)	-0.002 (0.005)		
× strong ethnic identification <sub>t0</sub>					0.043 <sup>a</sup> (0.016)	0.039 <sup>b</sup> (0.016)
Residence in homeland	0.045 <sup>a</sup> (0.012)	0.045 <sup>a</sup> (0.012)	0.045 <sup>a</sup> (0.012)		0.041 <sup>a</sup> (0.014)	
Fixed effects	Country × Ethnic group, Region × Time					
Individual controls	Yes					
Observations	115,105	115,105	115,105	61,790	76,103	39,697
R <sup>2</sup>	0.185	0.185	0.185	0.180	0.181	0.178
Sample mean/SD dep. var.	2.35/1.19			2.35/1.18	2.28/1.18	2.24/1.17

Notes: OLS estimations. Significant <sup>a</sup>at 1% and <sup>c</sup>at 10%. SEs are clustered at the ethnic group level. In columns 4 and 6, we restrict the sample to individual living outside their ethnic homeland. The dependent variable is the level of ethnic identification with respect to state identification in all columns, ranging from categories 1 to 5. Controls include age and its square, gender, rural/urban dummy, education categories and employment status. We interact the number of active mines in the homeland with variables accounting for the pre-existing level of battles, other conflicts, nightlights and identification. No. battles<sub>t0</sub> (respectively no. other conflict<sub>t0</sub>) is the number of battles (resp. other conflict events), in the ethnic homeland of the individual, over the period 2000–2004 (ACLED). Night-time lights<sub>t0</sub> is the average level of night-time lights over 2000–2004 in the homeland relative to the country. Strong ethnic identification<sub>t0</sub> is a dummy variable taking value one for groups where the average level of ethnic identification in the first Afrobarometer wave is above the sample median (this first wave is excluded from the sample in columns 5 and 6, we interact this group-level dummy with the number of active mines). Sample mean/SD dep. var. are respectively the sample mean and SD of the dependent variable in the corresponding column.

(Donaldson and Storeygard, 2016). The data come from the National Oceanic and Atmospheric Administration and was curated by Li *et al.* (2020). For each homeland, we compute the average total luminosity over the period 2000–2004, relative to the country’s average total luminosity. Finally, we measure overall pre-existing ethnic feelings by computing, for each homeland, the average ethnic identity in the first wave of the Afrobarometer during which the ethnic group is surveyed, and define homelands with strong pre-existing feelings of ethnic identity as homelands where this average is above the median.

The results point out that groups most likely to have pre-existing feelings of ethnic inequality are those who react the most to mineral resources exploitation in their historical ethnic homeland (Table 7). We first show that violent conflict (battles) history in the homeland magnifies our coefficient of interest, while other types of conflict events have little or no effect, be it in the full sample or focusing on households living outside of their historical homeland (columns 1, 3 and 4). However, a higher pre-period income attenuates our baseline coefficient, i.e., respondents whose ethnic homelands are poorer are the one who react the most, in terms of ethnic identification, when natural resources’ exploitation intensifies in their homeland (columns 2 and 3), although this effect loses statistical significance for individuals residing outside their homeland (column 4). Finally, pre-existing ethnic group specific levels of ethnic identification matter. Groups with

stronger pre-existing identification feelings (measured in the first survey wave) identify more to their ethnic groups when mineral exploitation takes place in their homeland, while respondents from groups with low pre-existing levels of identification do not react to the presence of natural resources. These various results are consistent with the channel of relative deprivation and unequal treatment.

#### **4. Conclusion**

While we know the risks of fragmented identities, we still know little about how one identity becomes so central that it shapes our actions and lives. Here, we document why natural resources may be an important driver of identities fragmentation in sub-Saharan Africa. We exploit geo-localised individual data on the strength of ethnic versus national identification in 25 countries of sub-Saharan Africa from 2005 to 2015, matched to the respondents' ethnic homelands and mineral resources exploitation in these homelands.

Our results suggest that mineral resources foster identity fragmentation. We document a significant increase in the strength of ethnic identification, relative to national identification, as a response to the exploitation of mineral resources in the ethnic group's historical homeland. The effect is persistent and tends to be magnified during elections periods, which implies that changes in resource exploitation might have a long-lasting impact on ethnic identification through voting patterns. We argue that the mechanism underlying this finding is one of economic deprivation, ethnic grievances and political exclusion. We fail at identifying substantial local economic spillovers from mining activity in our sample, and any significant additional economic benefit for individuals belonging to the group whose historical homeland hosts mining activities. However, the latter individuals report more economic deprivation and pessimism about economic conditions. We find that our results are mainly driven by individuals belonging to politically excluded groups. We show that pre-existing feelings of ethnic inequality, through political exclusion, conflict or poverty, matter. An open question is whether such grievances arise in a decentralised manner or result from manipulation by 'ethnic political entrepreneurs'—an interesting extension would be to study leaders' speeches to answer this question.

Overall, this paper documents a potentially understudied consequence of natural resources extraction and, at the same time, sheds new light on the sources of identities fragmentation, an issue which has potential implications for intergroup conflict as well as economic performance. Our results suggest that perceptions of economic or political deprivation can play a substantial role in driving identity constructions, regardless of actual economic achievements.

*Aix-Marseille University, CNRS, AMSE, CEPR, France*

*University of Lyon, Ecole Normale Supérieure de Lyon, Center for Economic Research on Governance, Inequality and Conflict and CEPR, France*

*NOVAFRICA, Nova SBE, Portugal, and LEO, University of Orléans, France*

Additional Supporting Information may be found in the online version of this article:

**Online Appendix  
Replication Package**

## References

- Ahlerup, P., Baskaran, T. and Bigsten, A. (2017). 'Regional development and national identity in sub-Saharan Africa', *Journal of Comparative Economics*, vol. 45(3), pp. 622–43.
- Aker, J.C., Klein, M.W., O'Connell, S.A. and Yang, M. (2014). 'Borders, ethnicity and trade', *Journal of Development Economics*, vol. 107(C), pp. 1–16.
- Akerlof, G.A. and Kranton, R.E. (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, Princeton, NJ: Princeton University Press.
- Alesina, A., Baqir, R. and Easterly, W. (1999). 'Public goods and ethnic divisions', *The Quarterly Journal of Economics*, vol. 114(4), pp. 1243–84.
- Alesina, A. and La Ferrara, E. (2005). 'Ethnic diversity and economic performance', *Journal of Economic Literature*, vol. 43(3), pp. 762–800.
- Alsan, M. (2015). 'The effect of the tsetse fly on African development', *American Economic Review*, vol. 105(1), pp. 382–410.
- Anderson, B. (2006). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, London: Verso.
- Aragón, F.M., Chuhan-Pole, P. and Land, B.C. (2015). 'The local economic impacts of resource abundance: What have we learned?', Working paper, World Bank.
- Aragón, F.M. and Rud, J.P. (2013). 'Natural resources and local communities: Evidence from a Peruvian gold mine', *American Economic Journal: Economic Policy*, vol. 5(2), pp. 1–25.
- Arezki, R., Ramey, V.A. and Sheng, L. (2017). 'News shocks in open economies: Evidence from giant oil discoveries', *The Quarterly Journal of Economics*, vol. 132(1), pp. 103–55.
- Armand, A., Coutts, A., Vicente, P.C. and Vilela, I. (2020). 'Does information break the political resource curse? Experimental evidence from Mozambique', *American Economic Review*, vol. 110(11), pp. 3431–53.
- Asiwaju, A. (1985). 'The conceptual framework', in (A. Asiwaju, ed.), *Partitioned Africans*, pp. 1–18, London: C. Hurst and Co.
- Aspinall, E. (2007). 'The construction of grievance: Natural resources and identity in a separatist conflict', *Journal of Conflict Resolution*, vol. 51(6), pp. 950–72.
- Atkin, D., Colson-Sihra, E. and Shayo, M. (2021). 'How do we choose our identity? A revealed preference approach using food consumption', *Journal of Political Economy*, vol. 129(4), pp. 1193–251.
- Barrios, S., Bertinelli, L. and Strobl, E. (2010). 'Trends in rainfall and economic growth in Africa: A neglected cause of the African growth tragedy', *The Review of Economics and Statistics*, vol. 92(2), pp. 350–66.
- Benshaul-Tolonen, A. (2019). 'Local industrial shocks and infant mortality', *ECONOMIC JOURNAL*, vol. 129(620), pp. 1561–92.
- BenYishay, A., Rotberg, R., Wells, J., Lv, Z., Goodman, S., Kovacevic, L. and Runfola, D. (2017). 'Geocoding Afro-barometer rounds 1–6: Methodology and data quality', Working paper, AidData at The College of William and Mary.
- Berman, N., Couttenier, M., Rohner, D. and Thoenig, M. (2017). 'This mine is mine! how minerals fuel conflicts in Africa', *American Economic Review*, vol. 107(6), pp. 1564–610.
- Binzel, C. and Carvalho, J.P. (2017). 'Education, social mobility and religious movements: The Islamic revival in Egypt', *ECONOMIC JOURNAL*, vol. 127(607), pp. 2553–80.
- Bisin, A., Patacchini, E., Verdier, T. and Zenou, Y. (2016). 'Bend it like Beckham: Ethnic identity and integration', *European Economic Review*, vol. 90(C), pp. 146–64.
- Caselli, F. and Michaels, G. (2013). 'Do oil windfalls improve living standards? Evidence from Brazil', *American Economic Journal: Applied Economics*, vol. 5(1), pp. 208–38.
- Cheeseman, N. and Larmer, M. (2015). 'Ethnopolitism in Africa: Opposition mobilization in diverse and unequal societies', *Democratization*, vol. 22(1), pp. 22–50.
- Christensen, D. (2019). 'Concession stands: How mining investments incite protest in Africa', *International Organization*, vol. 73(1), pp. 65–101.
- Colella, F., Lalive, R., Sakalli, S. and Thoenig, M. (2020). 'ACREG: Stata module to perform arbitrary correlation regression', Discussion paper 12584, IZA.
- Collier, P. (2017). 'The institutional and psychological foundations of natural resource policies', *The Journal of Development Studies*, vol. 53(2), pp. 217–28.
- Conley, T. (1999). 'GMM estimation with cross sectional dependence', *Journal of Econometrics*, vol. 92, pp. 1–45.
- Cornwell, C., Rivera, J. and Schmutte, I. (2017). 'Wage discrimination when identity is subjective: Evidence from changes in employer-reported race', *Journal of Human Resources*, vol. 52, pp. 719–55.
- Crespin-Boucaud, J. (2020). 'Interethnic and interfaith marriages in sub-Saharan Africa', *World Development*, vol. 125, article ID 104668.
- Cust, J. and Mensah, J.T. (2020). 'Natural resource discoveries, citizen expectations and household decisions', Working Paper 9372, World Bank.
- Cust, J. and Mihalyi, D. (2017). 'Evidence for a presource curse? Oil discoveries, elevated expectations, and growth disappointments', Working Paper WPS8140, World Bank.
- Cust, J. and Poelhekke, S. (2015). 'The local economic impacts of natural resource extraction', *Annual Review of Resource Economics*, vol. 7(1), pp. 251–68.



- Dahis, R., Nix, E. and Qian, N. (2019). 'Choosing racial identity in the United States, 1880–1940', Working Paper 26465, NBER.
- De Haas, R. and Poelhekke, S. (2019). 'Mining matters: Natural resource extraction and firm-level constraints', *Journal of International Economics*, vol. 117, pp. 109–24.
- Dehdari, S.H. and Gehring, K. (2022). 'The origins of common identity: Evidence from Alsace-Lorraine', *American Economic Journal: Applied Economics*, vol. 14(1), pp. 261–92.
- Dell, M., Jones, B.F. and Olken, B.A. (2012). 'Temperature shocks and economic growth: Evidence from the last half century', *American Economic Journal: Macroeconomics*, vol. 4(3), pp. 66–95.
- De Luca, G., Hodler, R., Raschky, P.A. and Valsecchi, M. (2018). 'Ethnic favoritism: An axiom of politics?', *Journal of Development Economics*, vol. 132, pp. 115–29.
- Depetris-Chauvin, E., Durante, R. and Campante, F.R. (2020). 'Building nations through shared experiences: Evidence from African football', *The American Economic Review*, vol. 10(5), pp. 1572–602.
- Dickens, A. (2018). 'Ethnolinguistic favoritism in African politics', *American Economic Journal: Applied Economics*, vol. 10(3), pp. 370–402.
- Donaldson, D. and Storeygard, A. (2016). 'The view from above: Applications of satellite data in economics', *Journal of Economic Perspectives*, vol. 30(4), pp. 171–98.
- Easterly, W. and Levine, R. (1997). 'Africa's growth tragedy: Policies and ethnic divisions', *The Quarterly Journal of Economics*, vol. 112(4), pp. 1203–50.
- Eifert, B., Miguel, E. and Posner, D.N. (2010). 'Political competition and ethnic identification in Africa', *American Journal of Political Science*, vol. 54(2), pp. 494–510.
- Esteban, J., Mayoral, L. and Ray, D. (2012). 'Ethnicity and conflict: An empirical study', *American Economic Review*, vol. 102(4), pp. 1310–42.
- Esteban, J. and Ray, D. (2011). 'Linking conflict to inequality and polarization', *American Economic Review*, vol. 101(4), pp. 1345–74.
- Fafchamps, M. (2000). 'Ethnicity and credit in African manufacturing', *Journal of Development Economics*, vol. 61(1), pp. 205–35.
- Fafchamps, M. (2003). 'Ethnicity and networks in African trade', *Contributions in Economic Analysis and Policy*, vol. 2(1), pp. 1–51.
- Fenske, J. and Zurimendi, I. (2017). 'Oil and ethnic inequality in Nigeria', *Journal of Economic Growth*, vol. 22(4), pp. 397–420.
- Gehring, K. and Schneider, S.A. (2020). 'Regional resources and democratic secessionism', *Journal of Public Economics*, vol. 181, article ID 104073.
- Gennaioli, N. and Rainer, I. (2007). 'The modern impact of precolonial centralization in Africa', *Journal of Economic Growth*, vol. 12(3), pp. 185–234.
- Girard, V., Kudebayeva, A. and Toews, G. (2021). 'Inflated expectations and commodity prices: Evidence from Kazakhstan', in (S. Tsani and I. Overland, eds.), *Handbook of Sustainable Politics and Economics of Natural Resources*, pp. 76–94, Cheltenham: Edward Elgar.
- Green, E. (2013). 'The political economy of nation formation in modern Tanzania: Explaining stability in the face of diversity', in (E.T. Woods, R.S. Schertzler and E. Kaufmann, eds.), *Nationalism and Conflict Management*, pp. 77–98, New York: Routledge.
- Green, E. (2020). 'Ethnicity, national identity and the state: Evidence from sub-Saharan Africa', *British Journal of Political Science*, vol. 50(2), pp. 757–79.
- Habyarimana, J., Humphreys, M., Posner, D.N. and Weinstein, J.M. (2007). 'Why does ethnic diversity undermine public goods provision?', *American Political Science Review*, vol. 101(4), pp. 709–25.
- Habyarimana, J., Humphreys, M., Posner, D.N. and Weinstein, J.M. (2009). *Coethnicity: Diversity and the Dilemmas of Collective Action*, New York: Russell Sage Foundation.
- Harris, A., Nielson, D.L., Medina, L., Berlin, W., Correia, C.B.M., Findley, M.G., Weinstein, J.M., Habyarimana, J., Humphreys, M. and Posner, D.N. (2018). 'Experimental evidence from Uganda, South Africa, and the United States on ethnic identification and ethnic deception', Conference paper prepared for the APSA Annual Conference, Boston.
- Harris, I., Osborn, T.J., Jones, P. and Lister, D. (2020). 'Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset', *Scientific Data*, vol. 7(1), pp. 1–18.
- Iwanowsky, M. (2018). 'The effects of migration and ethnicity on African economic development', Ms., [http://www.mathiwanowsky.com/Migration\\_and\\_Trade.pdf](http://www.mathiwanowsky.com/Migration_and_Trade.pdf) (last accessed: 5 March 2023).
- Kling, J.R., Liebman, J.B. and Katz, L.F. (2007). 'Experimental analysis of neighborhood effects', *Econometrica*, vol. 75(1), pp. 83–119.
- Klor, E. and Shayo, M. (2010). 'Social identity and preferences over redistribution', *Journal of Public Economics*, vol. 94(3–4), pp. 269–78.
- Lessmann, C. and Steinkraus, A. (2019). 'The geography of natural resources, ethnic inequality and civil conflicts', *European Journal of Political Economy*, vol. 59, pp. 33–51.
- Li, X., Zhou, Y., Zhao, M. and Zhao, X. (2020). 'A harmonized global nighttime light dataset 1992–2018', *Scientific Data*, vol. 7(1), pp. 1–9.
- Loayza, N. and Rigolini, J. (2016). 'The local impact of mining on poverty and inequality: Evidence from the commodity boom in Peru', *World Development*, vol. 84, pp. 219–34.

- Lowes, S., Nunn, N., Robinson, J.A. and Weigel, J. (2015). 'Understanding ethnic identity in Africa: Evidence from the implicit association test (IAT)', *American Economic Review*, vol. 105(5), pp. 340–5.
- Mamo, N. and Bhattacharyya, S. (2018). 'Natural resources and political patronage in Africa: An ethnicity level analysis', Working paper, University of Sussex Business School.
- Mamo, N., Bhattacharyya, S. and Moradi, A. (2019). 'Intensive and extensive margins of mining and development: Evidence from sub-Saharan Africa', *Journal of Development Economics*, vol. 139, pp. 28–49.
- Michalopoulos, S. and Papaioannou, E. (2013). 'Pre-colonial ethnic institutions and contemporary African development', *Econometrica*, vol. 81(1), pp. 113–52.
- Michalopoulos, S. and Papaioannou, E. (2020). 'Historical legacies and African development', *Journal of Economic Literature*, vol. 58(1), pp. 53–128.
- Monteiro, J. and Ferraz, C. (2010). 'Does oil make leaders unaccountable? Evidence from Brazil's offshore oil boom', Ms., PUC-Rio.
- Morelli, M. and Rohner, D. (2015). 'Resource concentration and civil wars', *Journal of Development Economics*, vol. 117, pp. 32–47.
- Moretti, E. (2010). 'Local multipliers', *American Economic Review*, vol. 100(2), pp. 373–7.
- Müller-Crepon, C., Pengl, Y. and Bormann, N.C. (2022). 'Linking ethnic data from Africa (LEDA)', *Journal of Peace Research*, vol. 59(3), pp. 425–35.
- Murdock, G.P. (1959). 'An atlas of African history: JD Fage', *American Anthropologist*, vol. 61(3), pp. 530–1.
- Must, E. (2018). 'Structural inequality, natural resources and mobilization in southern Tanzania', *African Affairs*, vol. 117(466), pp. 83–108.
- Nunn, N. (2008). 'The long-term effects of Africa's slave trades', *The Quarterly Journal of Economics*, vol. 123(1), pp. 139–76.
- Nunn, N. and Wantchekon, L. (2011). 'The slave trade and the origins of mistrust in Africa', *American Economic Review*, vol. 101(7), pp. 3221–52.
- Olzak, S. (2011). 'Does globalization breed ethnic discontent?', *Journal of Conflict Resolution*, vol. 55(1), pp. 3–32.
- Pelzl, P. and Poelhekke, S. (2021). 'Good mine, bad mine: Natural resource heterogeneity and Dutch disease in Indonesia', *Journal of International Economics*, vol. 131, article ID 103457.
- Pengl, Y.I., Roessler, P. and Rueda, V. (2022). 'Cash crops, print technologies, and the politicization of ethnicity in Africa', *American Political Science Review*, vol. 116(1), pp. 181–99.
- Raleigh, C., Linke, A., Hegre, H. and Karlsen, J. (2010). 'Introducing ACLED: An armed conflict location and event dataset', *Journal of Peace Research*, vol. 47(5), pp. 651–60.
- Ravallion, M. (2014). 'Poor, or just feeling poor? On using subjective data in measuring poverty', in (A.E. Clark and C. Senik, eds.), *Happiness and Economic Growth: Lessons from Developing Countries*, pp. 140–78, Oxford: Oxford University Press.
- Robinson, A.L. (2014). 'National versus ethnic identification in Africa: Modernization, colonial legacy, and the origins of territorial nationalism', *World Politics*, vol. 66(4), pp. 709–46.
- Rohner, D., Thoenig, M. and Zilibotti, F. (2013). 'Seeds of distrust: Conflict in Uganda', *Journal of Economic Growth*, vol. 18(3), pp. 217–52.
- Shayo, M. (2009). 'A model of social identity with an application to political economy: Nation, class, and redistribution', *American Political Science Review*, vol. 103(2), pp. 147–74.
- Van der Ploeg, F. (2011). 'Natural resources: Curse or blessing?', *Journal of Economic Literature*, vol. 49(2), pp. 366–420.
- Venables, A.J. (2016). 'Using natural resources for development: Why has it proven so difficult?', *Journal of Economic Perspectives*, vol. 30(1), pp. 161–84.
- Vogt, M. (2017). 'Escaping the resource curse: Ethnic inclusion in resource-rich states in West Africa', in (T. Falola and J. Abidogun, eds.), *Issues in African Political Economies*, pp. 199–225, Durham, NC: Carolina Academic Press.
- Vogt, M., Bormann, N.C., Rügger, S., Cederman, L.E., Hunziker, P. and Girardin, L. (2015). 'Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family', *Journal of Conflict Resolution*, vol. 59(7), pp. 1327–42.
- Wucherpfennig, J., Weidmann, N.B., Girardin, L., Cederman, L.E. and Wimmer, A. (2011). 'Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset', *Conflict Management and Peace Science*, vol. 28(5), pp. 423–37.