MiningLeaks

Water Pollution and Child Mortality in Africa

Job Market Paper [Please click here for the latest version]

Mélanie Gittard* and Irène Hu†

December 2, 2023

Abstract

In the midst of Africa's mining boom, communities downstream from industrial mines face increased exposure to toxic waste. Yet, the effects of induced water pollution on the local population's health have not been quantified at the continental scale of Africa, due to data limitation and non-random exposure. This paper investigates this question using a new quasi-experimental design and a novel dataset detailing the location and opening dates of all known industrial mines, obtained through intensive manual data collection. We combine geo-coded information on 2,016 industrial mines with health outcomes from the Demographic and Health Surveys (DHS) from 1986 to 2018 in 26 African countries. Through a staggered difference-in-difference strategy, we compare villages downstream and upstream of mines before and after their opening and find a 25% increase in 24-month mortality rates downstream. The effect is mainly observed among children who were no longer breastfed, confirming that water pollution drives the results. Our analysis rules out other mechanisms like fertility changes, access to facilities, in-migration, conflicts and income effects. The impact intensifies during mine operation and high international mineral prices, is higher in densely mined regions, and fades out with distance. From a public policy perspective, this paper underscores the significant local costs of mine openings on the environment and the health of the surrounding populations.

Keywords: Africa; Industrial Mining; Health; Water Pollution; Natural Resource; Environmental Degradation

JEL codes: I1; L72; Q53; Q32; Q33.

---

*Paris School of Economics (PSE), Centre International de Recherche sur l'Environnement et le Développement (CIRED) and Ecole Nationale des Ponts et Chaussées (ENPC), France. Contact author's email: melanie.gittard@psemail.eu

†Paris School of Economics (PSE) and Université Paris 1 Panthéon-Sorbonne, France. Contact author's email: irene.hu@psemail.eu. We are grateful to our advisors Denis Cogneau, Fabrice Etilé, François Libois and Philippe Quirion and would like to thank Liam Wren-Lewis, Thierry Verdier, Oliver Vanden-Eynde, David Margolis, Luc Behaghel, Sylvie Lambert, Victoire Girard, Mathieu Couttenier, Dominic Rohner, Francis Cottard, Aurore Stéphant, Jean-Alain Fleurisson, Laurent Polidori, Eléonore Lèbre, Cédric Rebolho and our many colleagues at Paris School of Economics and the CIRED for helpful comments and discussion. We are grateful to the participants of the CFDS at PSE, the *Atelier Theia*, the CSAE Conference, the JMA, and the EAERE. This research received the support of the CEPREMAP, the EHESS, the EUR-PGSE, and the GPET thematic group of PSE. We thank the support of the EUR grant ANR-17-EURE-0001. The usual disclaimer applies. Declaration of conflicts of interest: none.

1 Introduction

Metal mining contamination of rivers and floodplains affects 23 million people worldwide by exposing them to high levels of toxic waste [Macklin et al., 2023]. The surge in commodity prices since the 2000s has led to a significant increase in industrial mining activity across Africa, raising alarming concerns regarding environmental degradation [Taylor et al., 2009; Edwards et al., 2013] and adverse health effects among local communities. The ore extraction processes release toxic waste and heavy metals, posing a significant risk of contaminating nearby water resources. Human health can be directly affected through ingestion, inhalation and absorption of metal-contaminated water and indirectly through the food derived from soil-based agriculture. High blood metal levels can cause newborn malformation, neurological issues, heart problems, organ dysfunction, and raise cancer risk. They are especially harmful for children at a stage of rapid biological development. However, the impact of water pollution on the health of local African populations remains unquantified due to limited data availability and non-random exposure.

This paper investigates the local impacts of industrial mining-induced water pollution on local communities' health in Africa. We exploit the temporal and geo-spatial sources of variation and find evidence of water-pollution exposure through a staggered difference-in-difference (DiD) strategy. By comparing health outcomes downstream and upstream of mines before and after their opening, we indirectly assess mining-induced water pollution. We mainly focus on infant mortality, given the significant effects of pollutants absorption during early biological development. We also look at the effects on other children's health outcomes such as anthropometric measures and anemia, as well as women's health and fertility outcomes. We provide the most extensive dataset, to the best of our knowledge, on the location and timing of all known industrial mines in Africa. Using the SNL Metals and Mining database, we supplemented by intensive manual data collection the mines' opening dates by reading mining companies' reports and looking at historical satellite images. We combine geo-coded information on 2,016 industrial mines with health outcomes from the Demographic and Health Surveys (DHS) spanning from 1986 to 2018 over 26 African countries. For each mine and surveyed area, we determine the topographical relationship using the geocoded information from HydroSHEDS, which delineates water basins and builds the upstream-downstream network of each basin.

Our paper's main finding is to show that the opening of a mine leads to detrimental impacts on downstream child mortality, mainly due to increased exposure to water pollution. Our main result reveals an increase from 8.7 to 10.9% in the 24-month mortality rates for children born downstream after a mine opening, compared to those born upstream where mortality rates remain stable. This 2.2 percentage points (p.p.) increase represents a 25% higher mortality rate in downstream areas following a mine opening. We find an increase in the downstream 24-month mortality rates only for children who were no longer breastfed after 6 months. This result underscores the protection provided by breastfeeding [VanDerSlice et al., 1994; Fängström et al., 2008] and confirms that the results are mainly driven by water pollution. Additionally, we find no effects on other surviving children's health outcomes.

What drives these detrimental effects of industrial mining activity on health? We demonstrate that the results are the impact of water pollution and run several analyses to exclude other potential mechanisms. First, we establish that the impact on child mortality is independent from changes in women's health outcomes or fertility patterns. After a mine opening, we in fact observe no significant effects on fertility, pregnancy, miscarriage, anemia or sexually transmitted infections (STIs) in both downstream and upstream areas, and with no significant differences between the two areas. Then, we show that our results are not explained by an improvement in local welfare in upstream areas, by looking at different facilities such as access to piped water, flushed toilets, electricity, or health centers. Our results remain robust when including children living neither upstream nor downstream of the mine in the control group. This strengthens the control for income effects associated with mining activity and allows for a more accurate identification of the impact of water pollution. We identify an increase in the proportion of migrants downstream compared to upstream after a mine opens. Yet, our findings on mortality rates remain robust even after controlling for in-migration, indicating that the impact is primarily driven by water pollution. We also show that our results are robust when controlling for the presence of conflicts.

We conduct heterogeneity analyses showing that the more intense the mining activity, the higher the intensity of water pollution and its effects on mortality. The effect of the opening of a mine on downstream populations increases with the intensity of production, proxied by yearly international mineral prices, and mainly occurs while the mine is active, with downstream mortality rates reaching a 40% increase. We also find that the effect on mortality increases with surrounding mine density and that the effect fades out with the distance between villages and the industrial mining site. Moreover, in line with intensive extraction processes, we find that the increase in downstream mortality is mainly driven by the pollution caused by open-pit mines and by foreign-owned only mines. Additionnally, the heterogeneity analysis reveals a higher impact in areas dependent on the agricultural sector, where crucial water resources intersect with extensive and intensive extraction processes. We observe a larger effect in rural areas, with a 40% increase in downstream mortality rates. Lastly, the event study shows that the effects are significant in the medium run, and up to a decade after the mine opening.

Our findings are supported by a battery of robustness checks. Our results are robust when using a balanced sub-sample of DHS repeated cross-sections and when using the [de Chaisemartin and d'Haultfœuille, 2020] estimator to address negative weights and heterogeneous treatment effects in a staggered adoption design. To account for measurement errors, we implement corrections for DHS random displacement and restrict the analysis to the mines with precise coordinates. Additionally, we control for spatial correlation, and run spatial and temporal randomization inference tests.

Our job market paper provides a quantitative foundation to draw policy implications to help resource-rich developing countries conduct sustainable mining. It highlights the need for more stringent public policies, pointing out the insufficiency of existing policies. Our study challenges the effectiveness of the Extractive Industries Transparency Initiative (EITI), launched in 2002, which mandates member countries to disclose mining activity information and to improve governance of their extractive industries. Performing a back-of-the envelope calculation, we show that over 9,200 deaths per year can be attributed to water pollution induced by industrial mines in the 26 African countries of our analysis. Since this effect holds even for the 18 countries that participate in the EITI, this result demonstrates the ineffectiveness of the EITI in mitigating the adverse impacts of water pollution. This emphasizes the urgent need for the implementation of more rigorous policies, alongside the necessity to quantify the magnitude of the impact of such policies.

This paper contributes to the environmental literature on the health impacts of industrial activities and natural resource extraction. It enhances the ongoing debate on the health-wealth trade-off of industrial mining by highlighting water pollution as a significant negative externality impacting both the environment and human health in developing countries. We contribute to prior research by adopting a unique quasi-experimental design which enables us to identify water pollution and departs from the distance-based proxy used in the literature to study the effects of exposure to industrial mines. We also create a new database detailing the opening dates of industrial mining sites that complements the most extensive database on the location of industrial mines in Africa. Through these two main contributions, our research introduces a novel continental-scale upstream-downstream comparison in Africa to shed light on the health effects of water-induced pollution.

The remainder of the paper is organized as follows. Section 2 reviews the literature and presents our contributions. Section 3 describes the context and the data. Section 4 details the methodology and the main empirical strategy. Section 5 introduces the main results, while section 6 investigates the mechanisms, and section 7 the heterogeneity of the results. Section 8 looks at the dynamic effects, and section 9 at the intensive margins, digging into the heterogeneity of the results according to the distance to the mine, the mining density, and the production intensity. Section 10 proposes a list of robustness checks and placebo tests. Section 11 discusses the limits of the study and section 12 proposes a policy discussion. Eventually, section 13 concludes.

2 Literature review and contributions

This section first presents the literature on the trade-off of mining activity in developing countries. It then describes the mining-induced pollution literature and the economic literature on the health effects of mines. Lastly, we discuss the issues emerging from this literature and the solutions we propose to tackle them.

2.1 Trade-off of mining activity

Our work is related to the strand of literature analysing the health-wealth trade-off of industrial mining activity in developing countries, the results of which are still under debate. If mining can improve health and well-being through local industrial development, it can also damage health through negative externalities such as conflicts, massive migration waves, and exposure to harmful pollution. Determining which of these effects is predominant is still debated in the literature studying the relevance of a natural resource curse [van der Ploeg, 2011; Cust and Poelhekke, 2015; Venables, 2016].

At a broad scale of analysis, Mamo et al. [2019] look at the effects of the discovery of industrial mining deposits in Sub-Saharan Africa. They find an increase in district-level night light emissions, but no significant effects on household wealth [1].They find

[1]Household wealth was measured using the dimensions of access to electricity, wealth index, urbanization, mortality, and education.

temporary positive effects on public service provisions, but a degradation of the sewage system and piped water supply in the medium and long run. Mining also creates negative effects on the environment and agricultural productivity. Aragón and Rud [2016] find that the expansion of large-scale gold mining in Ghana (1997-2005) is responsible for the agricultural total factor productivity decrease in the vicinity of mines. The use of cross-sectional satellite imagery depicting NO2 concentration suggests that air pollution is the main explanatory factor. Dietler et al. [2021] analyze a panel of 52 mines in Sub-Saharan Africa using the same DHS and SNL databases. They find improvements in access to modern water and sanitation infrastructure after a mine opens when comparing individuals living within 50 kilometers of an isolated mine. Yet, proxying exposure to mining activity with distance and focusing on areas with low mining density raises many identification issues that will be largely discussed in section 2.4.1. Our paper deals with these issues and encompasses a wider sample of mines. Other negative externalities of mining activity are the increase of rapacity and corruption and the trigger of insecurity and conflicts [Berman et al., 2017], migration flows of mine workers fueling the spread of infectious diseases such as HIV [Corno and de Walque, 2012], and the discouragement of educational attainment among children [Atkin, 2016; Ahlerup et al., 2020; Malpede, 2021].

Our paper focuses on industrial mining and does not encompass artisanal and small-scale mining (ASM). Few papers have looked at the effects of ASM, mainly due to data limitations. Bazillier and Girard [2020] compare the local spillovers between artisanal and industrial mining sites in Burkina-Faso. They find that the present of artisanal mining (labor intensive and managed in common) and an absence of industrial mines (capital intensive and privatized) have positive impacts on household consumption. Our paper focuses on the effects of industrial mining pollution. If ASM has severe effects on miners' health due to hazardous working conditions, it is likely to be of a smaller magnitude than industrial mining, which extracts and treats larger volumes of minerals[2]. If ASM is often accused of generating more severe pollution than the industrial sector because of their illegal use of mercury [3], the latter often use cyanide instead. Both chemicals being highly toxic pollutants, focusing on industrial mining only is a lower-bound analysis of the impacts of mining activity on the health of local populations.

[2] Industrial mines are responsible for 80% of gold production and 75% of diamond production McQuilken and Perks [2020].

[3] Mercury has been officially banned in over 140 countries (Minamata Convention on Mercury, adopted in 2013).

2.2 Mining-induced pollution

Each stage of industrial mining activity produces chemicals and minerals likely to pollute the surrounding air, water, and soil [Coelho and Texeira, 2011]. The exploration and prospecting stage can last several years before a mine is considered economically viable and worthwhile to open. Meanwhile, mining companies conduct mapping and sampling, as well as drilling, boreholes, and excavations that require both physical and chemical measurement methods likely to pollute the surface and underground, depending on the nature of the deposit in the targeted area. If found financially viable, the company launches the discovery phase where the design and planning of the construction are undertaken. The feasibility study of the project requires further exploration and engineering studies. Subsequently, the development stage takes place and the mine's infrastructure and processing facilities are constructed. It is only after all these stages that production can start. Once the deposit is exhausted comes the closure and reclamation stage, during which the company is supposed to clean, stabilize and rehabilitate the land and isolate contaminated material. Yet, it is common that waste, tailings, or retention dams are abandoned without care or maintenance, which can constitute a potential disaster if hazardous materials leak and are discharged into the environment. Figure 28 in the Appendix proposes a scheme to explain the life cycle of a mine. Figure 27 displays satellite images of the different stages of the Essakane mine, an open-pit gold mine in Burkina-Faso.

Throughout all these stages, different types of pollution can be engendered. Air pollutants can be carried over long distances by dust, ore transportation and the wind; they can damage the surrounding soil and crops, and be inhaled, mostly by mine workers but also by the local population. The leakage of pollutants into the air can also affect water through acid mine drainage that ends up polluting first the surface and then groundwater. During the digging and processing to extract the targeted ore from waste rocks, rocks are crushed and then go through either heap leaching, froth flotation, or smelting. These techniques require the addition of chemicals, such as cyanide or acid, that can separate the targeted minerals from waste. Moreover, these processes are water-intensive and need access to a water source in competition with the local demand. Last but not least, even without the use of these chemicals, leaching happens through the contact of water and oxygen with sulfide minerals contained in the extracted rocks, which accelerates the acidification process and modifies the pH level of water bodies. Pollutants can be released into the environment during the process by spills or after by leaks of humid waste stored in retention dams, but also through the erosion and sedimentation of solid waste that is piled in the tailings around the mining site and drain into the soil with rain. The waste actively pollutes during the entire life cycle of the mine, starting with its opening and during production, but can also continue to pollute when a mine closes and is left without maintenance. This is the case when retention ponds are not covered and dry, allowing this waste to go directly into the environment.

Few papers have managed to show to what extent industrial mining activity creates negative externalities on the environment. Bialetti et al. [2018] look at the effects of mining industries on deforestation in India. Von der Goltz and Barnwal [2019] have suggested the mechanism of water pollution but without strong empirical evidence (looking at anemia). Yet, in-situ measurements have shown the contamination of potable water sources by harmful levels of nitrate, turbidity, iron, cadmium, manganese, and arsenic due to industrial mining sites [Cobbina et al., 2013]. To our knowledge, we are the first to provide indirect, systematic, and large-scale evidence of the mechanism of water pollution caused by industrial mining activity.

The main toxic metals released by mining sites are arsenic, cadmium, copper, lead, mercury, and nickel. Depending on their blood level concentration, they can be essential or non-essential for human health [El-Kady and Abdel-Wahhab, 2018]. However, heavy metals released by mining activity are non-biodegradable, have long-term impacts on the environment, and are found at abnormally high concentrations in the vicinity of mines, within the soil, water resources, vegetation, and crops [Oje et al., 2010; Dike et al., 2020]. People living in that environment are exposed to high quantities of heavy metals through ingestion, dermal contact, and inhalation of soil particles, which can cause several implications for their health. High blood metal concentrations are associated with neurological effects (which induce behavioral problems, learning deficits, and memory loss, especially among children) [Dike et al., 2020], neurodegenerative diseases, cardiovascular effects, gastrointestinal hemorrhages [Obasi et al., 2020], organ dysfunction (kidney, decrease of the production of red and white blood cells, lung irritation) [Briffa et al., 2020], higher probability of cancer development [Madilonga et al., 2021; Obasi et al., 2020], but also a higher probability of infertility, miscarriages, and malformation of newborns [Briffa et al., 2020]. Thus, exposure to heavy metals has a detrimental effect on human health in general and child health in particular, especially during their first months of development, both in- and ex-utero [Coelho and Texeira, 2011]. Children at an early age are the most sensitive, even to low concentrations of heavy metals, as they are at a stage of rapid biological development, but also as they are more exposed through higher blood concentrations linked to incidental ingestion of urban soil and dirty water (less conscious of their environment and danger, playing with polluted soil, eating and drinking without care [He et al., 2020]).

2.3 Health effects of mining activity

The empirical economic literature on the local effects of mining on local communities has grown during the past decade, yet the debate on the costs and benefits, and the positive and negative impacts of industrial mining activity in developing countries, remains. Diverse results have been found on the effects on health, and there is still uncertainty on the direction and magnitude of the impacts of mines on the health of the local population. Additionally, even if geographical proximity to a mining site is usually used as a proxy for pollution exposure, few papers observe the negative externalities on the environment and its consequences on health.

Papers studying the effects of industrial mines on health proxy the exposure to mining activity by the distance to the mine, and different thresholds and mixed results can be found in the literature. Using cross-section data in the state of Orissa in India, Shubhayu et al. [2011] use the distance to the mine as a proxy to measure environmental effects, and find that individuals living near a mine report higher respiratory illness and more work days lost due to malaria. Cross-sectional data prevent identifying a clear causal relationship and from adjusting to specific time and spatial confounders. Benshaul-Tolonen [2018] uses a DiD strategy, comparing individuals living within 10 kilometers to those living between 10-100 kilometers of a mine, before and after its opening. The paper finds that large-scale gold mining in nine countries of Sub-Saharan Africa [4] decreases infant mortality within 10 km during the opening and operating phases, with no effect on further communities (10-100 km). Cossa et al. [2022] use a similar methodology studying a broader set of countries and find a decrease in child mortality as well.

Von der Goltz and Barnwal [2019] assess the effects of industrial mines in 44 developing countries from 1988 to 2012. The paper also relies on a DiD strategy, comparing households living within 0-5 km with households living between 5-20 km before and after the opening of a mine. They find gains in asset wealth, increased anemia among women, and stunting in young children. As anemia and growth deficiencies are argued to be mainly the consequences of exposure to lead, the observed effects on health are interpreted to be the result of pollution due to metal contamination and lead toxicity. They find that women in mining communities show depressed blood hemoglobin, recover more slowly

[4] Burkina Faso, Ivory Coast, the Democratic Republic of the Congo, Ghana, Guinea, Ethiopia, Mali, Senegal, and Tanzania between 1987 and 2012.

from blood loss during pregnancy and delivery, and that children in mining communities suffer some important adverse growth outcomes from in-utero exposure (stunting).

2.4 Challenges and contributions

The most common way to proxy exposure to mining activity is to rely on the distance to an active or open mine; however, there is no clear consensus on which threshold to use, and the treatment allocation seems arbitrary. The disparity in the results of the literature could be explained by differences in terms of empirical strategies and distance choices. Beyond this, using the Euclidian distance to a mine as treatment raises endogeneity concerns. This subsection discusses the main issues that arise when studying the local impacts of industrial mining activity on health.

2.4.1 Endogeneity issues

In this section, two challenges are discussed: the endogeneity issues that arise: (i) when using the Euclidian distance as a proxy for exposure to mining activity; and those when using (ii) repeated cross-sectional data such as the DHS.

Using the interaction between being close to a mine and the mine's activity status raises endogeneity concerns. For instance, Von der Goltz and Barnwal [2019] use a mine panel and pair each DHS village to its closest mine. This creates an unbalanced treatment and control groups, and such an imbalance might be endogenous to socio-economic outcomes or polluting behaviors. As each village is paired to its closest mine, this *de facto* excludes group villages that are in both distance categories (within 5 km of mine A but 5-20 km of mine B) from control group. Thus, there is a higher probability of being treated in areas with high mining density, which is not a random allocation. As a mine fixed-effect identification relies on a within-mine buffer-area comparison, the estimator is driven by mines that have been paired to villages both in the treated and control areas, which is correlated to the mining density of the region. The estimation endogenously selects mines from regions of low or middle mining importance, which might be correlated with the intensity or the type of pollution and the socio-characteristics of the neighboring population, and thus the way health is affected by pollution. To reduce endogeneity issues, Von der Goltz and Barnwal [2019] instrument the mine location with mineral deposit information from S&P, which are *deposits that are being explored or prepared for exploitation*. However, mining exploration is not a random allocation and raises the same concerns as it is directly correlated with mining density. Benshaul-Tolonen [2018] reduces endogeneity issues linked to the pairing by using an administrative district fixed-effect panel and extending the distance (10-100 km), but the same concern remains.

A second concern is linked to the nature of the DHS data, which are repeated cross-section surveys. The literature argues that the conditions for an industrial mine to settle are the presence of mineral deposits, which is considered random. However, the presence of a mine and of a declared mineral deposit is correlated with the population density. As mining exploration is labor intensive, it is more likely to occur in dense areas where DHS is more likely to have surveyed individuals. A treatment allocation based on geographical proximity to the mine is endogenous: treatment groups close to the mine might not be comparable to control groups located further. As district fixed-effect relies on a within-district comparison, the estimation is driven by districts with both control and treated groups, before and after a mine opening, which is correlated with the probability of being surveyed. As DHS renews the surveyed villages at each wave, and as the probability of being surveyed is determined by the population density, the estimation is driven by specific areas. The regression *de facto* and endogenously selects districts that were already dense before the opening and remained so after. This might be areas that are more stable, well-off, and where individuals might be less affected by pollution. This might bias the estimation upward (i.e., less mortality linked to mining activity), and explain the positive effect of mines that Benshaul-Tolonen [2018] finds on mortality in Africa.

In Appendix section G, we propose a replication analysis of Benshaul-Tolonen [2018], taking advantage of our manual-entry work which extends the SNL database. We find similar results as Benshaul-Tolonen [2018], using the same set of countries and our extended sample of mines (using only, gold mines as in Benshaul-Tolonen [2018]). However, when applied to our more comprehensive sample, meaning when including other African countries and industrial mines, we find that are results are no longer significant, which suggests that the effects are context and regional-dependent.

2.4.2 Upstream-downstream analyses

Using geographical distance to a mine as treatment allocation raises endogeneity concerns. An upstream-downstream analysis, which relies on a topographical comparison, reduces these concerns as individuals from similar distances are compared.

Few papers have dealt with upstream and downstream at the scale of a continent, since it requires much more computational capacity and a complex pairing methodology. Duflo and Pande [2007] study the productivity and distributional effects of large irrigation dams in India and use river networks to calculate gradients computed from digital elevation maps for India. Do et al. [2018] use river networks and data from pollution monitoring stations in India to conduct their upstream-downstream analysis. Unfortunately, this is not possible in our case study due to the absence of water quality data from Africa as a whole. Garg et al. [2018] use river networks in Indonesia and re-calculate the upstream-downstream relationship between village pairs using a 30m resolution Digital Elevation Model. Their very refined level of study is not likely to be undertaken at the scale of the African continent in our case, so we chose secondary data computed by hydrologists (HydroSHEDS). We use systematic and highly disaggregated data on water sub-basins that enable us to encompass a wider set of countries, overcoming the issue of pairing a mine or a village with the closest river, since there is uncertainty about whether this point is located above or below the level of the river in altitude. Strobl and Strobl [2011] studied the distributional effects of large dams on agricultural productivity at the scale of the African continent, using Pfafstetter level 6 with an average area of 4200 km^2. Our study takes into account sub-basins at the Pfafstetter level 12, with an average area of 100 km^2.

3 Data and Context

This section describes the data used for our empirical strategy, and some descriptive statistics in the context of industrial mining and child mortality in Africa.

3.1 Data

In this paper, we match socio-economic data from the Demographic Health Surveys to an industrial mining database provided by SNL Mining and Metals.

3.1.1 Health and socio-economic data

We use all available survey rounds from the Demographic Health Surveys that contain GPS coordinates, from 1986 to 2018, covering 36 out of 54 African countries. We then select all the countries which have at least two survey waves to be able to implement our DiD strategy with a sufficient time length before and after the opening of a mine, and end up with 26 countries[5], 12,442 clusters, and 240,431 children under the age of five.

[5]The list of countries within our sample are: Benin, Burkina Faso, Democratic Republic of Congo, Burundi, Cote d'Ivoire, Cameroon, Ethiopia, Ghana, Guinea, Kenya, Liberia, Lesotho, Madagascar, Mali, Malawi, Nigeria, Niger, Namibia, Rwanda, Sierra Leone, Senegal, Togo, Tanzania Zambia, and Zimbabwe

We consider that doing a DiD strategy on the sample of countries that only have one round of the survey, hence a maximum period of five years, will not enable us to capture the longer-term effects of mining activity[6]. Table 23 in the Appendix displays the DHS survey years and countries that we use for our analysis.

We construct the variables of child mortality based on the DHS child recode database, which has information on the age and death of children under five years old, whose mothers are aged between 15-49 years old. Our dependent variable is the probability of 12-month and 24-month mortality for each DHS cluster (i.e., for each child, we construct a dummy variable equal to 1 if she or he is alive and 0 if not, conditional on having reached 12 and 24 months respectively). We also estimate the effects of mining activity on biomarker variables and other indicators of occurrences of illness (diarrhea, fever, and cough) within two weeks preceding the day of the interview among young children. We extend our analysis to women's fertility behavior and health: current pregnancy, total lifetime fertility, miscarriage, and anemia. Finally, as the aim of this article is to isolate the mechanism of water pollution, we use the questions from the DHS on the main source of drinking water, the presence of flushed toilets, electricity, and the access to healthcare facilities to control for households' sanitary and economic environment.

3.1.2 Mineral resource exploitation data

The industrial mining variables come from the SNL Metals and Mining database, which is privately owned by *S&PGlobal* and on license [7]. The SNL database is the best existing panel of mine production, providing information on the location, the dates of opening and closure, the commodity type, and the yearly production (for some mines). This is a non-exhaustive panel of industrial mines in Africa, yet to our knowledge, it constitutes the most comprehensive sample of mines giving the timing of the industrial activity. This dataset has been intensively used in the literature and argued to be the best product available ([Aragón and Rud, 2016; Berman et al., 2017; Kotsadam and Tolonen, 2016; Benshaul-Tolonen, 2018; Von der Goltz and Barnwal, 2019; Mamo et al., 2019]). We emphasize, here, that this paper focuses on the effects of industrial mining, and that we do not include ASM for which there is no available data in the SNL database.

[6] Please note that our final sample does not include Egypt, which has 7 DHS waves and is a well-known mining country. This is explained by the fact that the SNL database characterized Egypt as part of the Middle East rather than in Africa and thus was dropped from our sample.

[7] We are grateful to CEPREMAP, PjSE, EHESS, and the GPET thematic group of PSE for their financial support and their help in purchasing access to the data.

Overall, the SNL database gathers data for 3,815 industrial mines in Africa from 1981 to 2021. Of these, 2,016 mines were located within 100 km of a DHS cluster from a country with at least two surveys. For our DiD strategy, we need information on the timing of the start of mining production. The SNL database provides this information for 278 mines and we manually retrieved the start-up year for the 1,738 remaining mines. The manual collection of data was realized using information on mining history available in the SNL database, and mine reports (cross-checked with Google Maps and aerial images). We describe the manual collection of data more extensively in Appendix B.2.

We build three main variables from the SNL Mining and Metals database by relying on geocoded information and the opening date of the mine. According to the estimation strategy, we will use a variable of proximity (distance to the closest mine), position (whether individual i is upstream or downstream), and a dummy variable for being open or not. Opening dates that were available in the SNL database were computed by the SNL team, and indicated the actual startup year of the mine, i.e., when production first began. We used the same criteria for the data we collected. Finally, we restrict the main analysis that is associated with heavy metal mines (metals with density higher than $5g/cm3$ ([Briffa et al., 2020]), which are the metals listed in Table 27 in Section C.2 of the Appendix). We also include coal mines, as their extraction is associated with mercury and arsenic, which are highly toxic heavy metals.

3.1.3 Water basins

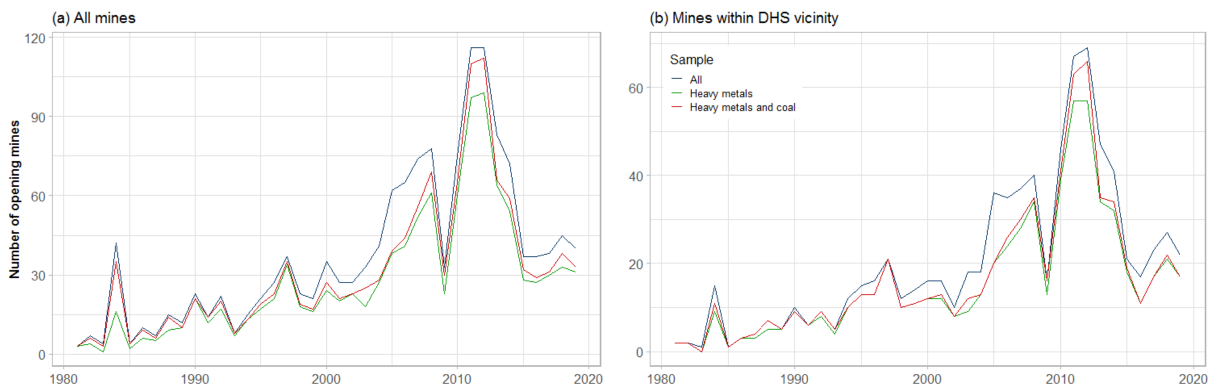
We consider the topographical relationship of water basins where mines and villages are located. A water basin is an area where all the surface water converges towards the same point. We use the HydroBASINS sub-basins geographical information provided by HydroSHEDS, which delineates water basins consistently and subdivides sub-basins into multiple tributary basins to the network of nested sub-basins at different scales. Following the topological concept of the Pfafstetter coding system, each polygon of the sub-basin has a unique direction flow and provides information on up- and down-stream connectivity. We take the finest Pfafstetter level (12 out of 12) that breaks sub-basins down to an average area of 100 km^2. See Figure 6a for an example. We conduct our analysis taking into consideration the three closest sub-basins to each industrial mine, meaning that we take each mine's sub-basin A and tag the one just downstream as B; the one just downstream of B is tagged C; and the one just downstream of C is tagged D. Thus, B, C, and D are the three sub-basins closest to A.

3.2 Descriptive statistics

3.2.1 Mining in Africa

Temporal and spatial variation

Figure 1: Temporal evolution of mine openings



Notes: The data lines plot the number of mines opened each year for the 1981-2019 period, for all mines, heavy metal mines including coal (sample of the main analysis), and only heavy metal mines. Panel (a) displays the temporal evolution of the total mine sample, while panel (b) shows mines that are within the sample of the main analysis, meaning mines that have DHS clusters upstream at a distance of at most 100 km and DHS clusters downstream within the three closest sub-basins.
Sources: Authors' elaboration on DHS and SNL data.

Figure 1 shows the evolution of the yearly number of mines that opened in Africa over the 1981-2019 period, Figure 1 (a) for the entire mining sample and Figure 1 (b) for mines that are in the sample of the main analysis. The mining boom since 2000 is shown in the Figure, with the first peak in 2007, which is in line with the peak in exploration activity that occurred in 2003 ([Taylor et al., 2009]) (as the exploration phase takes place on average a couple of years before a mine opens), and the second peak in 2012. For instance, about 120 industrial mines opened in 2012 (based on the non-exhaustive SNL database). The Figure also presents the evolution according to the characteristics of the mines: it distinguishes the pattern for all mines, heavy metal mines, and heavy metals including coal mines. We observe no differences in timing patterns between Figure 1 (a) and (b), nor between mine types.

What is striking in Figure 1 is that the evolution of mine openings follows the same pattern as the evolution of industrial metal prices, as plotted in Figure 29 in Section C.2 of the Appendix. The mining boom since 2000 follows the increase in real prices of copper, tin, lead, aluminum, zinc, nickel, and other heavy metals, while the sharp fall around 2008/2009 corresponds to the financial crisis. The local minimum around 2016 corresponds to the drop in commodity prices in June 2014 ([Khan et al., 2016; Glöser et al., 2017]). This similar evolution suggests that heavy metal prices are good Instrument Variables for the variable year of mine opening, such as Berman et al. [2017] and Bazillier and Girard [2020] used in their analyses. In Section 9.3 we will use it as a proxy for production intensity.

Figure 2 (c) shows the map of the number of mines that opened before 2019, including mines that opened before 1986, averaged at the cell level (160 km cells). Cells in grey represent areas where no mine opened before 2019, but where at least one will open in the future (whether we know from the data that it opened between 2019-2021, or if the opening is planned for the future). The main mining countries in the SNL database are Guinea, Sierra Leone, Ivory Coast, Ghana, Niger, Burkina Faso, Zimbabwe, Tanzania, Zambia, and the north of South Africa. Please note that since we have excluded countries with only one DHS wave in the main sample of our analysis (cf. Tables 22 and 23), in order to avoid comparing areas with too many differences in terms of temporal variations, we did not undertake the manual-entry work for these countries, which explains why South Africa (which is not in the final sample) does not appear as a major mining country in Figure 2 (c). Figure 3 shows both the temporal and spatial variation of mine openings in Africa (for all the mines sample, not the restricted one of our main analysis), as it plots the number of mines that opened over different periods of our analysis per grid cell. The cells in red are areas where no mines opened during the specified period, but where at least one mine had opened before, whereas cells in grey are areas where no mines have been opened but at least one will open in the future. We observe that the increase in mine openings was higher during the third period 2008-2019 (which is in line with Figure 1), and was particularly important in West Africa.

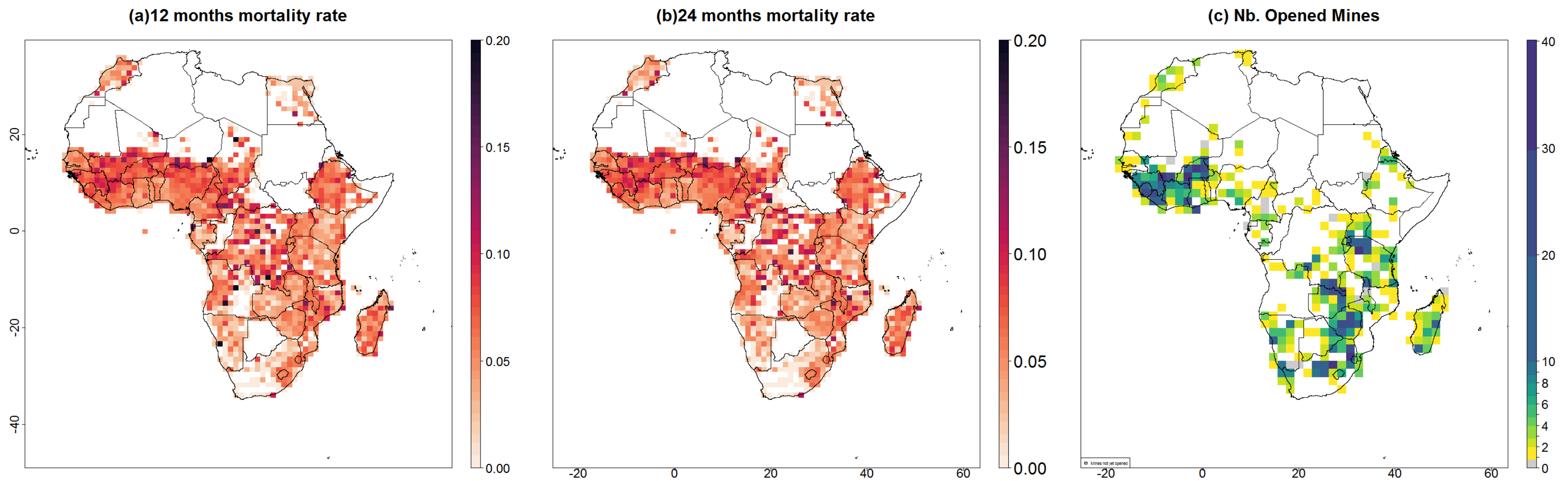
3.2.2 Health risks

Africa faces high infant mortality rates as the average 12-month mortality rate is 6.4% and the average 24-month mortality rate is 8.3% according to DHS data (cf. Table 24). Figures 2 (a) and (b) plot the average mortality rates for all DHS from 1986-2019 averaged at the grid level, and show the spatial variation of mortality rates [8]. Figures 4 and 5 map both spatial and temporal variation of mortality rates as they show the average mortality

[8]Please note that the higher the DHS cluster density, the more accurate the average. The spatial variation is endogenous to the DHS sample.

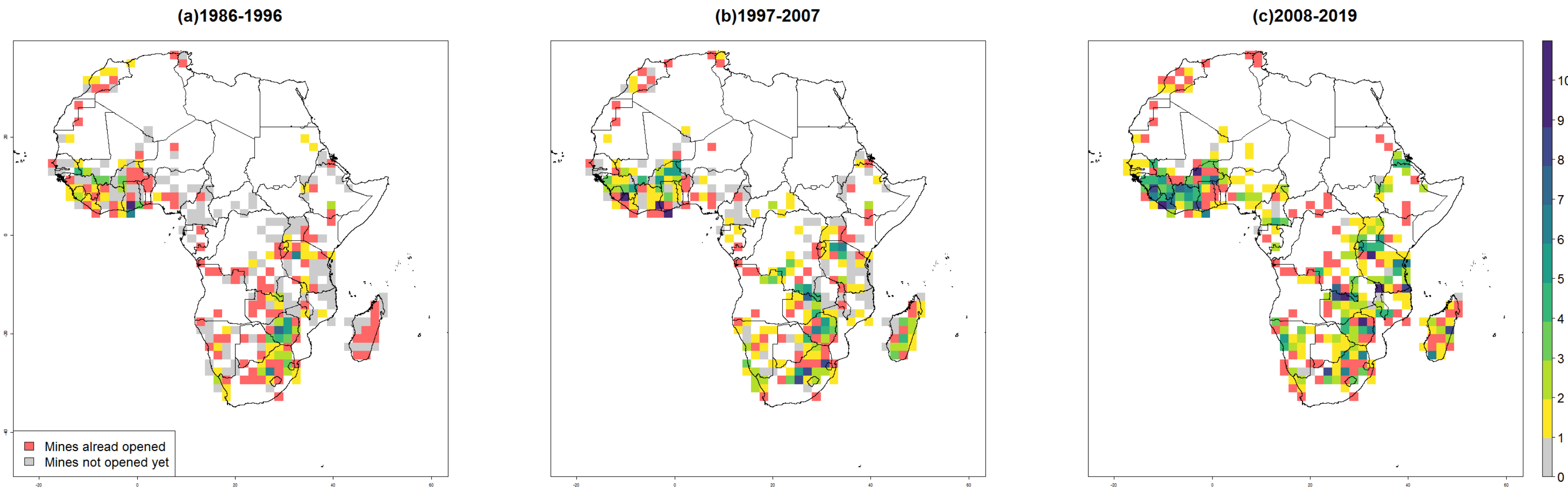
rates for the three main time periods of our DHS sample. We observe the global reduction of mortality over the time period and also the DHS cluster distribution. Figures 31, 32 and 33 plot the same maps for the sample restricted to the one used in the main regression.

Figure 2: Outcomes spatial distribution



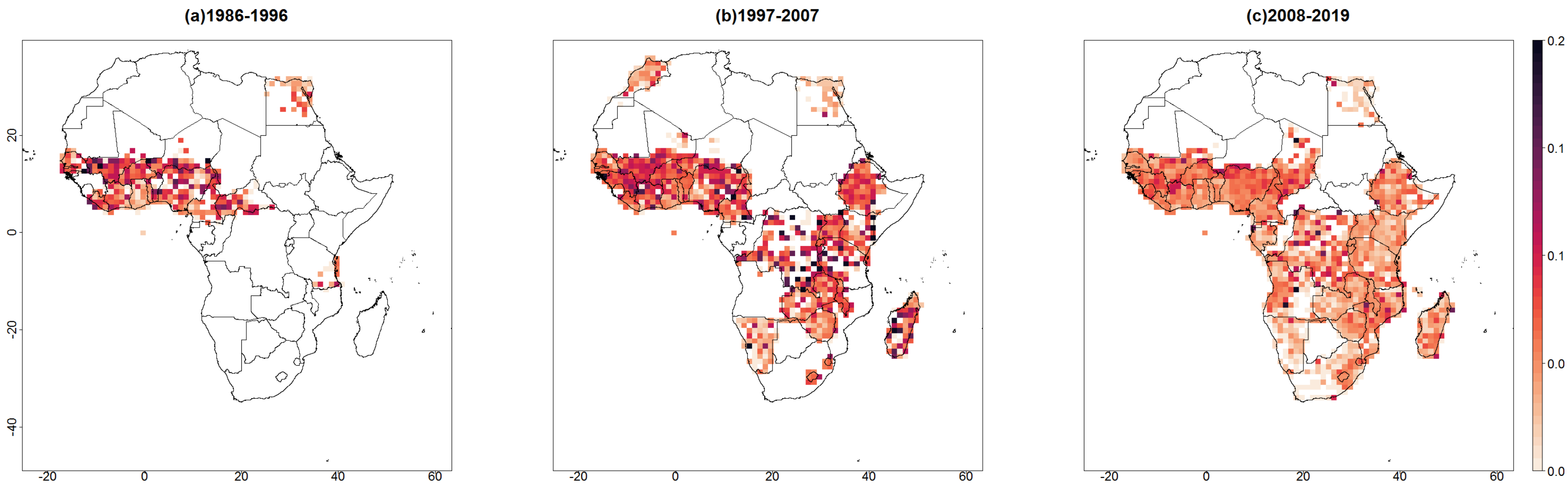
Notes: Panels (a) and (b) represent the means of 12- and 24-month mortality rates for each DHS wave available (listed in table 23) from 1986 to 2019. Means are computed at the grid level (100 km mean size). The mortality rates are estimated without the children that did not reach 12/24 months at the time of the survey. Panel (c) displays the stock of mines that opened before 2019 (including mines that opened before 1986). Means are computed at the grid level (100 km mean size).
Sources: Authors' elaboration on DHS and SNL data.

Figure 3: Spatial variation of mine openings per period



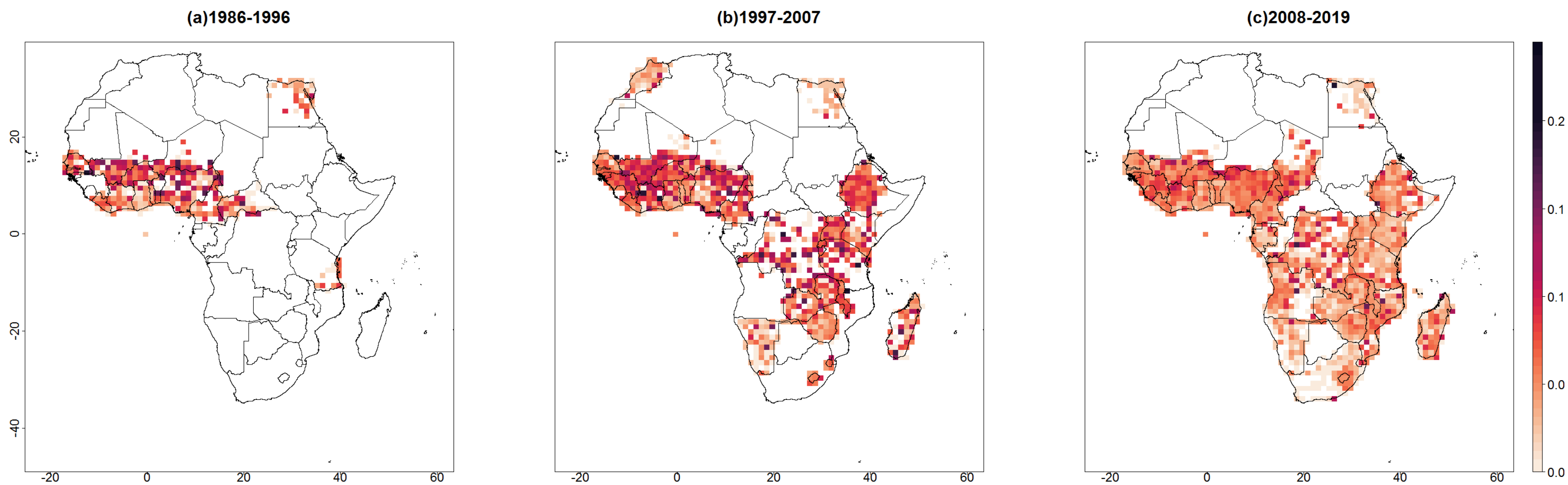
Notes: This figure shows the number of mines that opened over the grid area (160 km on average) during each period. A red grid cell represents an area where no mine opened during the period, but where at least one mine open before the period. A grey cell represents an area where no mine opened during the period, but where at least one mine will open in the future.
Sources: Authors' elaboration on SNL data.

Figure 4: Spatial variation of 12-month mortality rates per period



Notes: This figure shows the means of 12-month mortality rates averaged at the grid level during each period. The mortality rates estimated do not include children that did not reach 12 months of age at the time of the survey.
Sources: Authors' elaboration on DHS data.

Figure 5: Spatial variation of 24-month mortality rates per period



Notes: This figure shows the means of 24-month mortality rates averaged at the grid level during each period. The mortality rates estimated do not include children that did not reach 24 months of age at the time of the survey.
Sources: Authors' elaboration on DHS data.

4 Empirical strategy

The main empirical strategy of this paper uses the relative topographical position of sub-basins as a proxy for exposure to mining activity pollution. It compares the effects on the health of individuals living downstream of a mine to those living upstream, before and after the opening of at least one site. It is a staggered-design DiD analysis with two-way fixed effects at the mine's sub-basin and birth year level. This upstream-downstream strategy intends to identify the mechanism of water pollution.

As seen in Section 2.4.1, this strategy alleviates some endogeneity issues raised by treatments using the Euclidian distance as a proxy for exposure to the mine. First, it reduces the bias linked to unbalanced samples due to endogenous pairing. Second, it breaks the average effects based on distance buffers and highlights the heterogeneity of the effects of mining activity on health, while isolating the negative externalities linked to water degradation.

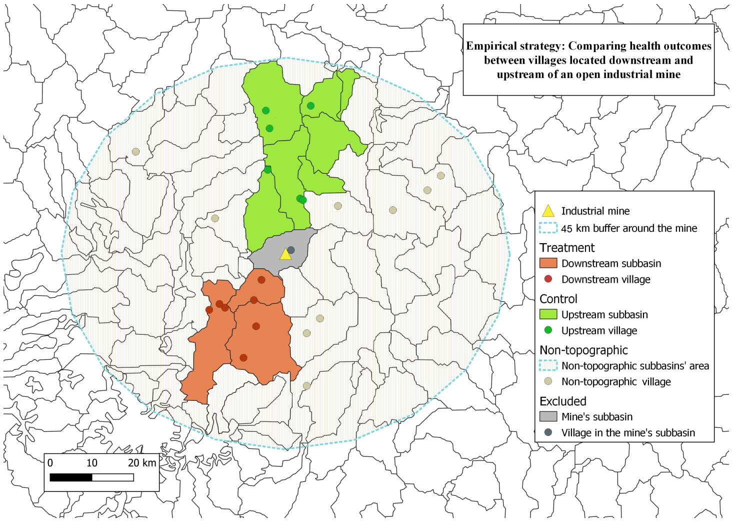
4.1 Measuring exposure to pollution

4.1.1 Pairing strategy

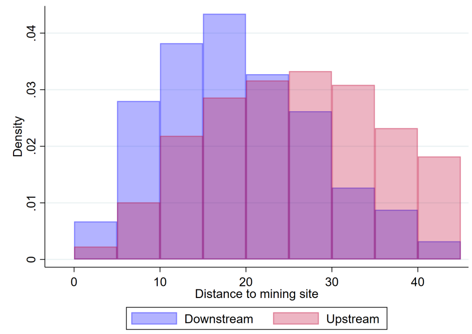
The pairing of DHS clusters to mines represents a significant challenge, as each DHS cluster can be downstream of and close to several industrial sites in major mining areas. It introduces endogeneity in the sample selection and raises the issue of unbalanced samples. In this analysis, we propose the following pairing to overcome this issue and thus be able to measure the exposure to pollution of each DHS cluster.

First, we construct a 100 km buffer around each DHS cluster and register all mines within this buffer (independently of their activity status). We then categorize the topographical position of the DHS cluster relative to the industrial site using a dummy variable equal to 1 if the cluster is downstream of the mine and 0 if it is located upstream. This topographical position is defined using the relative position of each sub-basin. As each cluster and site have GPS coordinates, they lie in a specific sub-basin, and we used the relative position of each sub-basin to classify the DHS according to the paired mine. Through such a process we also have pairs that are located in the same sub-basin, and for which it is impossible to say exactly whether the cluster is downstream or upstream of the mine. At this stage, for these specific couples, we consider the DHS to be downstream. Please note that, as explained in section 3.1.3, we used the finest Pfafstetter level 12 that

Figure 6: Pairing Strategy



(a) Scheme for pairing strategy



(b) Density of distance to the mine

Notes: Panel (a) is a scheme that illustrates the pairing, giving the example of a mine, its main sub-basin, its three closest downstream and upstream sub-basins, and DHS clusters that are in the treatment and control areas within 45 km. Panel (b) plots the density of the distance (in km) to the mining site for DHS clusters across their upstream-downstream position.
Sources: Authors' elaboration on DHS, SNL, and HydroSHEDS data.

breaks down sub-basins at an average area of $100km^2$ (the size of the sub-basin varies according to their shape, cf. Figure 6a). At this stage, some villages can be paired with several mines and can have more than one occurrence in the sample. The difficulty of the strategy lies in choosing the mine that will be paired with the cluster.

Second, we restrict the group of downstream DHS clusters to the ones that lie within one of the three closest sub-basins downstream of the mine's sub-basin, to focus on the most potentially contaminated areas. Third, we pair each cluster with only one mine, proceeding as follows: if a DHS was in both groups (i.e., downstream a mine A and upstream a mine B), then it is automatically assigned to the downstream group, and it is paired to the mine from which it is downstream (i.e., it is paired to mine A) regardless of its activity status. At this stage, some clusters may still be counted twice, as they can be upstream of several mines, or in the three closest sub-basins downstream of several mines. To complete the uniqueness of the pairing, we paired each cluster to the nearest mine regardless of its activity status.

In conclusion, the DHS clusters are attached to the nearest mine from which they are downstream up to the third sub-basin level, or else attached to the nearest mine upstream up to a radius of 100 km. The final remaining problem relates to the clusters that are in the mine's same sub-basin, which we have so far identified as being downstream. We eliminated from the main analysis all DHS villages which are located in the same sub-basin of the mines to which they were paired. Also, this reduces the noise linked to the random displacement of DHS villages (cf. Section 10.3.1) and avoids allocating villages as being downstream whereas they are upstream due to the displacement, as it drops the closest areas around the mine.

Once the pairing is done, we restrict the control group to upstream villages that are within 45 km of the mine to ensure the comparability of upstream and downstream villages. To choose this distance cut-off, we have calculated the mean of the maximal distance between a mine and the furthest extremity of its third downstream sub-basin, which was 44.7 km. Figure 6b plots the distribution of the distance to the mine for both upstream and downstream villages. As downstream villages are prioritized in the pairing strategy, they are slightly closer to the mine, but the two distributions are comparable.

The pairing is illustrated in Figure 6a. It gives the example of a mine, its main sub-basin (grey), the downstream sub-basins (orange), and upstream sub-basins (green), up to 45 kilometers. The dashed area displays the sub-basins within 45 kilometers with no topographical relationship to the mine, meaning they are neither downstream nor upstream. In the main strategy, we compare the villages within the green area to those in the orange area. In section 9 we run robustness tests to check whether the results hold, allowing for further sub-basins and heterogeneity effects by distance to the mining site. In section 10.2.1 we discuss the results when including non-topographical sub-basins.

4.2 Identification Strategy

4.2.1 Main estimation

The main analysis relies on a DiD strategy using the topographical position of a DHS cluster relative to a mine deposit to indirectly identify the channel of water pollution. We propose a staggered DiD specification, with a sub-basin fixed effects panel for each mine. We isolate the mechanism of water pollution by building the treatment and control groups using an upstream-downstream comparison. We restrict our analysis using the pairing strategy explained in section 4.1.1. We compare health outcomes in upstream-downstream areas both before and after the opening of the paired mine. The empirical strategy can be formally written as follows:

$$
\begin{aligned}
Death_{i,v,c,SB} = & \alpha_0 + \alpha_1 Opened_{birthyear,i,v} + \alpha_2 Downstream_{v,SB} \\
& + \alpha_3 Opened_{birthyear,i,v} \times Downstream_{v,SB} + \alpha_4 X_i \\
& + \gamma_{SB} + \gamma_{SB-trend} + \gamma_{c,birthyear} + \epsilon_v
\end{aligned} \tag{1}
$$

$Death_{i,v,c,SB}$ a dummy variable equal to 1 if child i from DHS village v of country c has reached the n^{th} month and has died (n being 12 for 12-month mortality, same for 24 months). $Opened_{birthyear,i,v}$ is a dummy variable equal to 1 if the mine, which is located in sub-basin SB, has opened before child i's year of birth. $Downstream_{v,SB}$ is a dummy variable of relative position (equal to 1 if village DHS v is located in a sub-basin downstream of mine sub-basin SB, and 0 if it is upstream), and X_i a vector of child and mother level controls (mother's age, age square, years of education, urban residency). Finally, γ_{SB} is a mine sub-basin fixed effects, $\gamma_{SB-trend}$ a mine sub-basin linear birth year trend, and $\gamma_{c,birthyear}$ a country birth year fixed effects. This analysis is a staggered design as the treatment shock (mine opening) does not occur at the same time for each DHS cluster.

The main regression is run without the DHS clusters that lie within the same sub-basin as the mine they are coupled with (as discussed in the previous section). The list of countries and survey years used in the main regression are given in Table 23, and the list of metals in Table 27.

4.2.2 Identification assumption

The key assumption of a DiD is that the downstream group would have evolved like the upstream group in the absence of the opening of a mine. As we cannot test if those upstream and downstream areas would have followed the same time trends, we test in Section 8 the common trend assumption using pre-treatment data.

However, the fact that pre-treatment data are parallel is neither a necessary nor a sufficient condition for the identification. Past trends can be identical, but the upstream group may be affected by a group-specific shock during the period of the treatment. The estimation of this paper relies on the fact that the comparison between downstream and upstream villages is a proxy for exposure to water pollution. The major identification assumption is that the opening of a mine affects upstream and downstream areas differently only through the decrease in water quality. Throughout the paper we will try to address the concerns of unobservable factors that might not be orthogonal to our treatment. Section 11 displays a final general discussion of the threats to the identifying assumption and how they have been solved in the analysis.

4.3 Descriptive statistics

In this section, we describe the balance tables for the outcomes that play a key role in our analysis out of parsimony.

Table 1: Balance Table

	Before Mine Opening					After Mine Opening					Within Up.	Within Dwn.	Within
	Upstream		Downstream		Diff	Upstream		Downstream		Diff			
	N	Mean /(SD)	N	Mean /(SD)	(4-2) /(p.v)	N	Mean /(SD)	N	Mean /(SD)	(9-7) /(p.v)	(7-2) /(p.v)	(9-4) /(p.v)	(12-11) /(p.v)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Dth<12													
All	23,547	0.073	7,875	0.074	0.001	12,319	0.055	4,738	0.051	-0.004	-0.018	-0.023	-0.005
		(0.261)		(0.262)	(0.83)		(0.228)		(0.219)	(0.256)	(0)	(0)	(0.468)
Mines	244		237			179		183					
Dth<24													
All	17,726	0.096	5,928	0.098	0.002	8,664	0.068	3,330	0.072	0.004	-0.028	-0.026	0.002
		(0.294)		(0.297)	(0.618)		(0.252)		(0.259)	(0.428)	(0)	(0)	(0.671)
Mines	244		236			168		168					

Notes: Standard errors and p-values in parentheses. Descriptive statistics of 12- and 24-month mortality outcomes, for villages located upstream and downstream of mining sites, for individuals born before and after the opening of the mine.

Balance Table 1 compares the changes in infant mortality before and after the opening of a mine for places upstream *vs* downstream of the mining site, following the pairing strategy. It also displays the number of individuals and paired mines in each group of the analysis. On average, upstream and downstream areas have non-significant differences in terms of 12- and 24-month mortality (columns (5) and (10)). For both upstream and downstream clusters, the opening of a mine significantly decreases the mortality probability (columns (11) and (12)), which is in line with the results of Benshaul-Tolonen [2018], and with the fact that mortality rates decrease over time in Africa as trends are not included (Figures 31 and 32). Table 1 shows that this reduction is overall slightly more important in upstream areas than in downstream areas for under 24-month mortality (0.002), while the opposite holds for under 12-month mortality (-0.005), but the differences are not significant (column (13)). Table 1 does not include any controls and is only descriptive. Table 28 in Section D.1 of the Appendix replicates this exercise for control variables.

Figure 30 in the Appendix identifies the country with the biggest stock of open mines in our sample (Ghana, Zimbabwe, and Tanzania with the highest density of open mines nearby DHS), as well as insights on the variation in mine openings per country over the period.

5 Main results

This section displays the results of our main analysis. The first section describes the overall effects of mine opening on child mortality among the villages living downstream compared to those living upstream. Section displays the effects of being downstream of an open mine on other child health outcomes, while the section focuses on women's outcomes.

5.1 Child mortality

This section presents the main results of this paper derived from equation 1 for the 12- and 24-month mortality rate. Table 2 gathers our main results with mine sub-basin and country birth year fixed effects. We also include mine sub-basin and birth year linear trends adjusting for spatial and period-specific co-founders and trends, and commodity fixed effects. Columns (1) to (4) give the results for the 12-month mortality rate, while columns (5) to (8) for the 24-month mortality rate. Columns (1), (2), (5), and (6) display the results for the total population, while columns (3), (4), (7), and (8) focus on the rural population. Control variables are birth order, mother's age, mother's age square, mother's years of education, urban, and the intensity of the river network. [9]. Even columns include the number of open mines within 45 km of the DHS cluster as control, which controls for the mining density.

The results show that being downstream of an open mine increases the 24-month mortality rate [10] by 2.18 percentage points (p.p). This corresponds to an increase of 25% as the average 24-month mortality increases from 8.7% to 10.9%. The results are higher in terms of magnitude in rural areas, as being downstream an open mine increases the 24-month mortality rate by 3.8 p.p , which is associated with an increase of 40%, as the

[9] The variable intensity of the river network using the HydroRIVERS product. It is a continuous variable, which takes into account the area of the catchment that contributes directly to a river reach, and the Strahler order of the specific river segment. In our sample, the Strahler spans from 3 to 10.

[10] 95 % Confidence interval: [0.000595; 0.042993]

Table 2: Effects of industrial mining opening on child mortality

	12-month mortality				24-month mortality			
	Total Population		Rural Population		Total Population		Rural Population	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Down×Open	-0.00352	-0.00506	0.00517	0.00510	0.0231**	0.0218**	0.0379***	0.0380***
	[0.00824]	[0.00831]	[0.0102]	[0.0102]	[0.0105]	[0.0108]	[0.0130]	[0.0130]
Downstream	-0.0140**	-0.0152**	-0.0203***	-0.0204***	-0.0202***	-0.0211***	-0.0287***	-0.0284***
	[0.00655]	[0.00665]	[0.00743]	[0.00762]	[0.00731]	[0.00739]	[0.00795]	[0.00810]
Open	0.0121*	0.00963	0.0106	0.0102	-0.00302	-0.00496	-0.00650	-0.00588
	[0.00722]	[0.00754]	[0.00858]	[0.00952]	[0.00986]	[0.0101]	[0.0115]	[0.0122]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nb open mines	No	Yes	No	Yes	No	Yes	No	Yes
Birthmth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ctry-bthyr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MineSB-bthyr trd	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	48,472	48,472	33,231	33,231	35,638	35,638	24,544	24544
R2	0.0378	0.0378	0.0476	0.0476	0.0511	0.0511	0.0633	0.0633
Outcome Mean	0.0666	0.0666	0.0716	0.0716	0.0873	0.0873	0.0945	0.0945

Notes: Standard errors clustered at the DHS village level, ${}^{*}p < 0.1, {}^{**} p < 0.05, {}^{***} p < 0.01$. The variables Downstream and Opened are dummies that indicate whether an individual lives in a village downstream of at least one mining site and whether the site opened before the year of birth. Each DHS village is paired to only one mining site so that each individual appears only once in the regression. Other variables are control variables. The sample focuses on heavy metal mines. Columns (1-3) give the results for the total population while columns (4-6) display the results for rural villages. Columns (2, 4, 6, 8) control for the number of open mines within 45 km. Control variables are birth order number, mother's age, mother's age square, mother's years of education and urban, number of open mines, and a continuous variable indicating the presence of rivers and their order.

mortality increases from 9.4% to 13.2%. This is in line with the fact that rural populations have less access to facilities and infrastructure and are more exposed to unsafe water. The results are not significant concerning the 12-month mortality rate, and are very close to zero, showing no difference between individuals living upstream to those living downstream. This lag in the effect of water pollution on children's health may be explained by the higher probability of children under 12 months to be breastfed compared to children under 24 months, hence their decreased exposure to contaminated water and limited direct ingestion [VanDerSlice et al., 1994; Fängström et al., 2008]. This mechanism explaining the different results on 12- *vs* 24-month mortality is explored in Section 6.1.

5.2 Other health effects

Table 3 represents the effect of industrial mining on child health outcomes other than mortality. Columns (1)-(3) present the results on anthropometric measures of children who were still living at the time of the survey and are measured at the time of the survey. A child is affected by stunting if her height-for-age z-score is below minus 2 standard deviations below the mean on the World Health Organization Child Growth Standards. The same definition applies to underweight (weight-for-age) and wasting (weight-for-height). We find a negative and significant effect of industrial mining on underweight but not on stunting and wasting: living downstream of an open mine decreases wasting by 3.9 pp. This result could potentially be explained by the death of the most vulnerable children and the survival of the heaviest ones. We find no results on other diseases among living children: anemia (measured), diarrhea, cough, or fever (reported within the two weeks preceding the interview). We find no effect either of industrial mining on low weight at birth (below 2.5 kg), nor reported size at birth (reported as small or very small by the mother).

Table 3: Effects of industrial mining opening on other child health outcomes

	Surviving children							All births	
	Stunting	Underweight	Wasting	Anemia	Diarrhea	Cough	Fever	$< 2.5kg$	Small
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Down×Open	-0.0167	-0.0389**	-0.00479	-0.0265	0.00146	-0.00824	-0.00219	-0.00925	-0.00337
	[0.0199]	[0.0169]	[0.0109]	[0.0277]	[0.0130]	[0.0176]	[0.0154]	[0.0180]	[0.0120]
Downstream	-0.0126	-0.00293	0.00330	0.0428**	0.00287	-0.0115	0.0141	-0.00951	0.00673
	[0.0172]	[0.0152]	[0.00897]	[0.0188]	[0.0108]	[0.0138]	[0.0144]	[0.0163]	[0.0107]
Open	-0.00618	0.0257*	0.00769	0.00582	-0.00545	0.00547	-0.00571	0.00571	0.0191
	[0.0162]	[0.0146]	[0.0106]	[0.0246]	[0.0111]	[0.0124]	[0.0127]	[0.00954]	[0.0159]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ctry-bthyr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MineSB-bthyr trd	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	37,393	37,043	37,903	19,331	55,162	54,958	54,955	29,162	58,338
R2	0.155	0.124	0.0895	0.215	0.0900	0.104	0.111	0.0608	0.0532
Outcome mean	0.308	0.246	0.0893	0.660	0.165	0.237	0.246	0.171	0.157

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. Columns (1-3) focus only on surviving children (due to variable construction in DHS), while the others encompass all children, including those who died before the survey. The same sample and controls as Table 2 Column 2 apply.

5.3 Women's outcomes

We make sure that the results regarding child mortality are not due to a change in fertility among women [11]. We find no significant effect of industrial mining on whether women ever had a child (Table 4, column 1), or on the total number of children she had (column 2). We find no effect on whether women were pregnant during the time of the survey (column 3). Table 4 displays results on women's other health outcomes: we find no effect of industrial mining on whether women ever had a miscarriage or on their anemia status.

Table 4: Effects of industrial mining opening on women outcomes

	Fertility			Health	
Outcome	Ever had a child	Total lifetime fertility	Currently pregnant	Ever had a miscarriage	Anemia
	(1)	(2)	(3)	(4)	(5)
Down × Open	0.0156 [0.00977]	-0.0164 [0.0720]	-0.0171 [0.0111]	-0.00507 [0.0138]	0.000806 [0.0261]
Downstream	0.00886 [0.00952]	0.0841 [0.0731]	0.0160 [0.0118]	0.00894 [0.0134]	-0.00928 [0.0233]
Open	-0.00161 [0.00916]	0.0663 [0.0595]	0.00607 [0.00833]	-0.00136 [0.0119]	-0.00417 [0.0285]
Controls	Yes	Yes	Yes	Yes	Yes
Ctry-survey year FE	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes
MineSB-svey year trd	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes
N	82,406	82,406	82,373	72,423	31,587
R2	0.510	0.659	0.0422	0.0906	0.122
Outcome mean	0.737	2.912	0.0939	0.136	0.396

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. Control variables are birth order number, woman's age, woman's age square, woman's years of education, urban, number of open mines and presence of rivers.

[11] The analysis has been made using DHS Women Recode, with a population sample of all women aged 15-49 years old.

6 Mechanisms

6.1 Early life characteristics

We want to better understand the mechanism behind the results on 12- and 24-month mortality, and focus on children's breastfeeding and access to healthcare.

6.1.1 Breastfeeding

Table 5 gives the results of several triple interactions investigating the heterogeneity of the effect according to breastfeeding practices. The adverse effect on mortality is mainly driven by children who were no longer breastfed after 6 months. In columns (1) and (2), we analyze the effects on mortality rates between the ages of [6;12] months and [6;24] months. We interact the DiD estimator with a dummy variable, which equals one if the child was not breastfed after 6 months and zero otherwise. The results reveal a significant increase, especially among children who were not breastfed after 6 months. We find a 10 p.p increase in the [6-24] month mortality rate for individuals downstream after a mine opening, who were no longer breastfed after six months, compared to their breastfed counterparts, who experienced a 1 p.p increase. This result highlights the key role of breastfeeding in shielding children from the effects of water pollution VanDerSlice et al. [1994]; Fängström et al. [2008] and supports the argument that our results are driven by water pollution.

We find no significant effect on whether the child was ever breastfed (columns 3 and 4) or the number of months during which the child was breastfed (columns 5 and 6). This absence of results can be explained by the low variation as 98% of the children in the sample were breastfed.

6.1.2 Access to healthcare

In Table 6, none of the triple interactions concerning access to healthcare show significant effects. Our findings are not influenced by factors such as maternal prenatal care (as indicated in columns 1 and 2) or the child's vaccination status (as shown in columns 3 and 4)

Table 5: Effects of industrial mining opening - Breastfeeding

	Breastfeeding					
Var.	No longer breastfed		Ever breastfed		Breastfeed months	
	(1)	(2)	(3)	(4)	(5)	(6)
Mortality outcome	[6,12]	[6,24]	12m	24m	12m	24m
Downstream× Open × Var.	0.0720	0.107*	0.000597	-0.0450	0.00106	0.00106
	[0.0533]	[0.0634]	[0.0519]	[0.0481]	[0.00176]	[0.00210]
Downstream	-0.0017	-0.0065	-0.0104	-0.0161	-0.0125	-0.0293
	[0.0037]	[0.0055]	[0.0250]	[0.0305]	[0.0237]	[0.0300]
Open	-0.0126	-0.0228***	0.0172	0.0163	-0.0785***	-0.134***
	[0.0044]	[0.0064]	[0.0266]	[0.0345]	[0.0231]	[0.0324]
Downstream×Open	0.00218	0.014**	0.00423	0.0702	-0.0213	-0.000897
	[0.0048]	[0.007]	[0.0523]	[0.0488]	[0.0373]	[0.0516]
Var.	0.0474**	0.0526**	-0.921***	-0.880***	-0.0179***	-0.0201***
	[0.0227]	[0.0274]	[0.0115]	[0.0134]	[0.000595]	[0.000657]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes
N	43,628	33,666	45,168	33,022	29,015	18,323
R2	0.0382	0.047	0.208	0.174	0.330	0.355

Table 6: Effects of industrial mining opening - Access to healthcare

	Child's access to health care			
Var.	No prenatal care		Ever vaccinated	
	(1)	(2)	(3)	(4)
Mortality outcome	12m	24m	12m	24m
Downstream× Open × Var.	0.0165	0.0288	0.0114	0.0134
	[0.0272]	[0.0402]	[0.0191]	[0.0231]
Downstream	-0.0140**	-0.0200**	0.0117	0.00258
	[0.00613]	[0.00871]	[0.0127]	[0.0126]
Open	0.00842	0.000268	0.00433	0.00929
	[0.00770]	[0.0117]	[0.0105]	[0.0175]
Downstream×Open	-0.0114	-0.000663	-0.0119	-0.0105
	[0.00823]	[0.0121]	[0.0178]	[0.0211]
Var.	0.0299***	0.0398***	-0.0201***	-0.0257***
	[0.00774]	[0.0110]	[0.00728]	[0.00828]
Controls	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes
Commodity FE		Yes	Yes	Yes
Yes				
N	31,656	19,543	17,372	13,638
R2	0.0558	0.0822	0.239	0.306
Outcome mean	0.0479	0.0639	0.00835	0.0121

6.2 Households' access to water and facilities

We deepen our analysis by studying whether the effects found on child mortality are indeed due to water pollution downstream of mines and not driven by improved access to water and sanitation or facilities upstream. Under 24-month mortality is still significantly increased by 2 p.p. when adding the triple interaction with several facilities variables: whether a household has piped water as the main drinking source (Table 7, column (1)), whether a household has a flushed toilet (column (2)), whether it has access to electricity (column (3)), and whether the mother had visited health facilities during the 12 months preceding the survey (column (4)). We find no significant heterogeneity across the four facilities, which suggests that our results are not explained by an improvement in facilities upstream, which contradicts the findings of Dietler et al. [2021]. Table 29 in Section D.1 of the Appendix looks at the DiD estimator using access to piped water and electricity as dependent variables, and shows no difference after the opening of a mine between upstream and downstream villages.

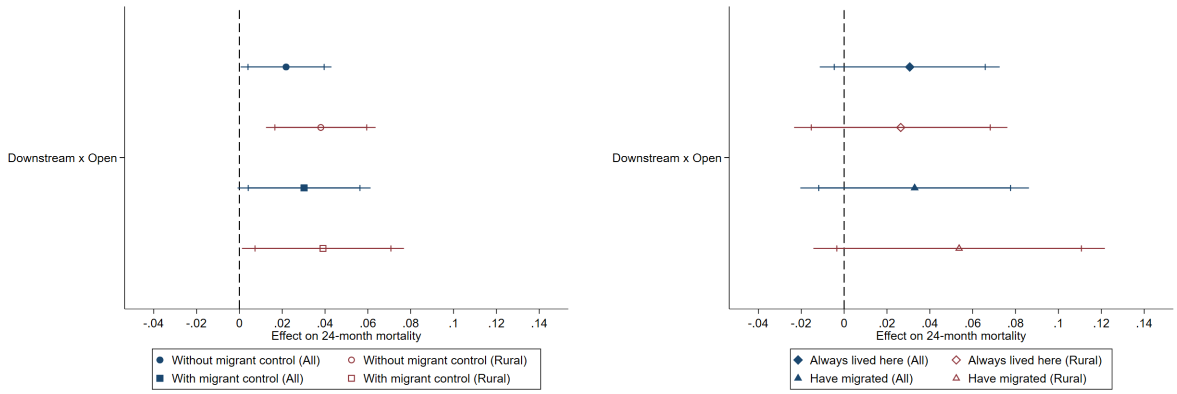
Table 7: Effects of industrial mining opening on access to water, sanitation and facilities

Outcome	24-month mortality			
Var.	Has piped water	Has flushed toilet	Has electricity	Visited health facilities
	(1)	(2)	(3)	(4)
Downstream×Open × Var.	0.000283	-0.0356	-0.0114	-0.0140
	[0.0195]	[0.0263]	[0.0192]	[0.0165]
Downstream×Open	0.0210*	0.0236**	0.0243**	0.0292*
	[0.0118]	[0.0110]	[0.0115]	[0.0159]
Var.	-0.00266	-0.0165*	-0.0157**	-0.00491
	[0.00637]	[0.00920]	[0.00695]	[0.00541]
Downstream	-0.0229***	-0.0215***	-0.0223***	-0.0170*
	[0.00770]	[0.00744]	[0.00764]	[0.00983]
Open	-0.000958	-0.00482	-0.00715	-0.00543
	[0.0103]	[0.0101]	[0.0103]	[0.0122]
Controls	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes
N	35,638	35,536	35,423	32,018
R2	0.0512	0.0513	0.0512	0.0512
Outcome mean	0.0873	0.0873	0.0873	0.0857

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$. The same sample and controls as Table 2 Column 2 apply.

6.3 Migration

Figure 7: Migration analysis



(a) With and without migrant control

(b) Migrant/Always lived here samples

Notes: Panel (a) plots the coefficients associated with being downstream of an open mine on under 24-month mortality when controlling for mothers' migration status or not, across the entire rural sample. Panel (b) plots the coefficients of the same interaction but across the sample of mothers who have migrated or always lived here. *Sources:* Authors' elaboration on DHS and SNL data.

We pursue our analysis by making sure that our results on child mortality are not due to migration and do not suffer from selection bias. The migration information is retrieved from the variable indicating whether mothers have ever migrated to the actual place of residency, or if they have always lived there. This information is available for 60% of our sample (cf. Table 25) and controls for in-migration, which is an important effect of the opening of a mine that attracts new working populations (cf. Section D.1 for more discussion on bias linked to migration). Figure 7 displays the coefficient associated with the interaction term of being downstream of an open mine. The top coefficient in Figure 7a is our main specification when we do not control for migration across the whole. We plot the same focusing on the rural sample. The bottom two coefficients are when we control for migration across our whole and rural sample. We find that all are statistically positive and significant. We further our analysis by splitting the sample across mothers who have ever migrated and mothers who have always lived there (Figure 7b). The estimation suffers from a lack of statistical power (see Appendix Table 8 for the drop of observations) but suggests industrial mining has no differential effect across the two samples.

Table 8: Effects of industrial mining activity, migration analysis

Outcome	Mortality under 24 months							
Spec.	Without migrant control		With migrant control		Migrant sample		Always lived here	
Sample	All	Rural	All	Rural	All	Rural	All	Rural
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Downstream x Open	0.0218**	0.0380***	0.0302*	0.0390**	0.0329	0.0537	0.0306	0.0264
	[0.0108]	[0.0130]	[0.0158]	[0.0193]	[0.0272]	[0.0347]	[0.0214]	[0.0254]
Downstream	-0.0211***	-0.0284***	-0.0249**	-0.0332***	-0.0229	-0.0304	-0.0301**	-0.0418***
	[0.00739]	[0.00810]	[0.0106]	[0.0119]	[0.0190]	[0.0226]	[0.0139]	[0.0147]
Open	-0.00496	-0.00588	-0.0255	-0.0235	0.0116	0.0187	-0.0409**	-0.0425*
	[0.0101]	[0.0122]	[0.0159]	[0.0194]	[0.0260]	[0.0335]	[0.0198]	[0.0256]
Migrant			0.00850*	0.00348				
			[0.00449]	[0.00594]				
N	35638	24544	22231	15060	8658	6007	13503	8982
R2	0.0511	0.0633	0.0634	0.0770	0.112	0.132	0.0797	0.107
Outcome mean	0.0873	0.0945	0.0946	0.104	0.0892	0.102	0.0983	0.107

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. The same sample and controls as Table 2 Column 2 apply.

7 Heterogeneity

7.1 Individual characteristics

Table 9: Effects of industrial mining opening across children's location and gender

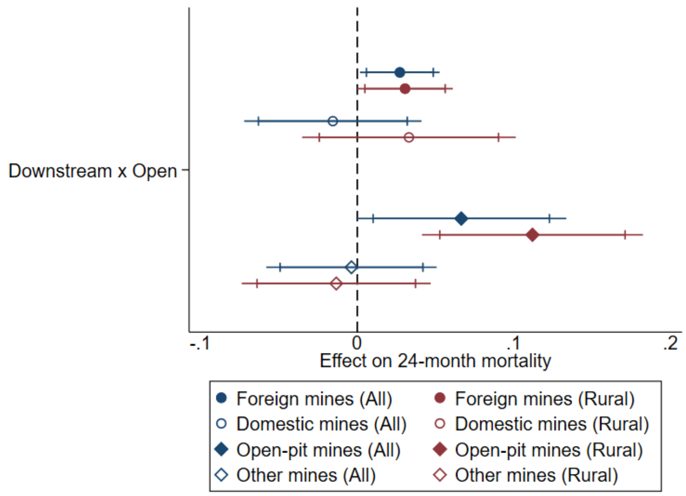
Outcome	24-month mortality					
Sample	All: urban and rural			Rural		
	All	Girls	Boys	All	Girls	Boys
	(1)	(2)	(3)	(4)	(5)	(6)
Downstream× Open	0.0218**	0.0120	0.0334**	0.0380***	0.0204	0.0677***
	[0.0108]	[0.0151]	[0.0167]	[0.0130]	[0.0186]	[0.0199]
Downstream	-0.0211***	-0.0203*	-0.0206*	-0.0284***	-0.0292**	-0.0292**
	[0.00739]	[0.0114]	[0.0111]	[0.00810]	[0.0126]	[0.0117]
Open	-0.00496	0.00364	-0.0178	-0.00588	0.00419	-0.0200
	[0.0101]	[0.0155]	[0.0143]	[0.0122]	[0.0186]	[0.0181]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes
N	35,638	17,452	18,142	24,544	12,009	12,481
R2	0.0511	0.0758	0.0762	0.0633	0.0942	0.0972
Outcome mean	0.0873	0.0805	0.0938	0.0945	0.0883	0.101

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$. The same sample and controls as Table 2 Column 2 apply.

We conduct a heterogeneity analysis across children's location and gender (Table 9). We find that being downstream of an open mine is more critical in rural areas (column 4), as it increases the 24-month mortality by 3.8 p.p, which corresponds to a 40% increase in mortality rate. The heterogeneity by gender shows that our results are mainly driven by the mortality of males (columns (3) and (6)), which remains consistent in rural areas.

7.2 Mining activity's characteristics

Figure 8: Heterogeneity across mines' characteristics



Notes: This graph represents the coefficients associated with the interaction of living downstream of an open mine when splitting the sample between mines that are owned by foreign companies and mines that are owned by at least one domestic company (in blue), and between mines that are open-pit and not open-pit (underground, placer, and in-situ leach) (in red).
Sources: Authors' elaboration.

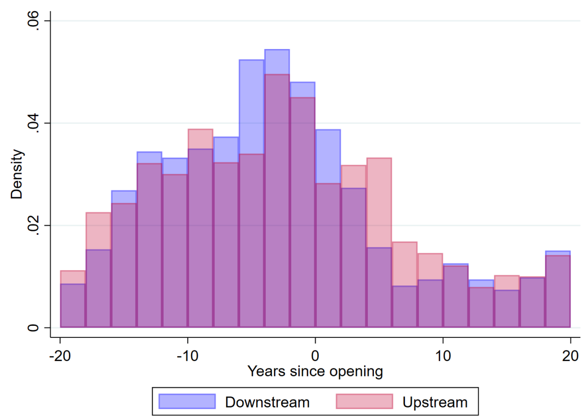
We pursue the heterogeneity analysis across mines' ownership and extraction methods. A mine was considered domestic if at least one of the owning companies is from the same country as the country of location of the mine. Domestic mines represent 17.8% percent of our mine sample. Figure 8 represents the coefficients associated with the interaction term of living downstream of an open mine. We find no effect of a mine opening when we restrict the sample to domestically-owned mines, whereas our results hold when we restrict only to the foreign-owned mines (in blue). This could potentially be explained by improved management of a mine or better consideration for the surrounding populations if a national company is involved. We then look at the open-pit nature of the industrial site, which concerns 21.6% of our mine sample. We find that our results hold when restricting to open-pit mines but not when we only look at the sample of other extraction methods (underground, placer, and in-situ leach) (in red). This is consistent with the fact that open-pit mines are the most polluting mines, due to the generation of large amounts of waste kept in tailing storage facilities.

8 Dynamic effects

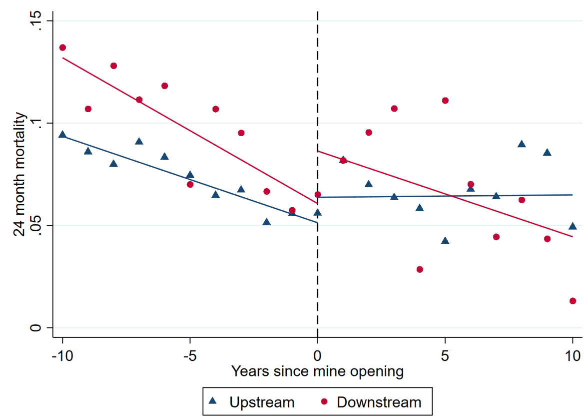
In this section we investigate the dynamic effects of the opening of an industrial mine looking at pre-trends and at whether the effects on 24-month mortality occur within a short or long time, and during the mining activity.

8.1 Pre-trends and event-study

Figure 9: Linear trends of 24-month mortality



(a) All sample



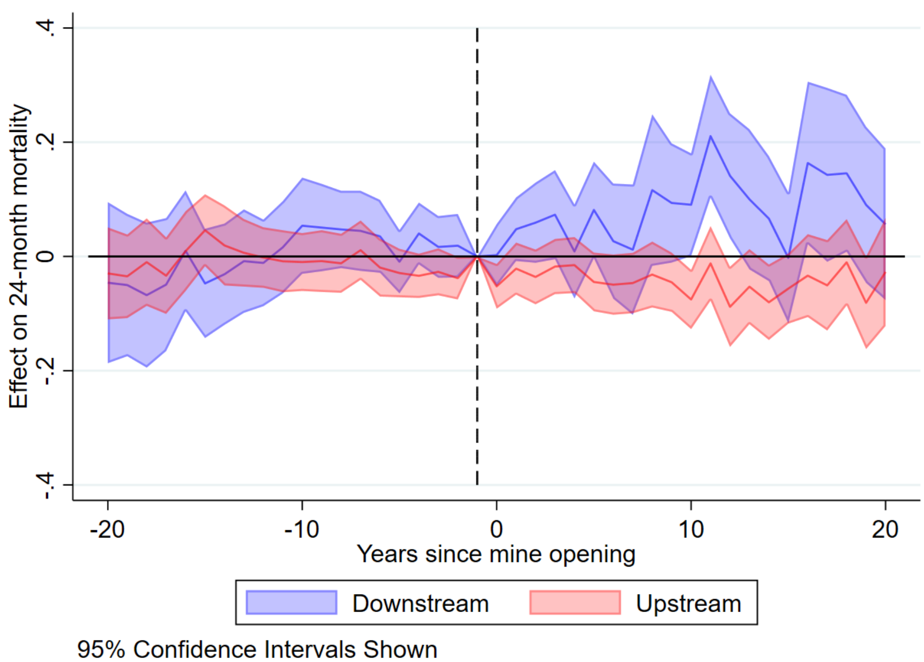
(b) All sample

Notes: Panel (a) shows the distribution of the number of observations per opening year. Panel (b) plots the trends of the 24-month mortality rates according to the oepning year. The figures are made for the whole sample and include neither control variables nor fixed effects. The reference point is -1, the year before the mine opening.
Sources: Authors' elaboration on DHS and SNL data.

The key assumption of the DiD strategy is that the outcome - the 24-month mortality - would follow the same time trend in the absence of the mine opening both in upstream and downstream areas. The common trend assumption cannot be tested. However, we can observe the pre-treatment data and the evolution of mortality rates before each mine opening according to the topographical position. Figure 9b plots the linear trends of the 24-month mortality rates and distinguishes between upstream and downstream DHS clusters, before and after the opening of the paired mine. For each year it plots the average mortality rates over the sample with no control nor fixed effects. Figure 9a plots the distribution of the years before and after the mine opening. Figure 9b shows non-exact parallel trends but is only descriptive, and we can see by looking at the scatter plot that the downstream and upstream areas seem to follow a similar pattern of decreasing mortality before a mine opens. This decreasing pattern is triggered by temporal trends, as the years closest to the mine opening are more likely to be recent years, and the mortality rates are declining overall over the recent decades (cf. Figure 5). This pattern is corrected in Figure 10.

Figure 10 plots the event study of the effect of mine opening for both the upstream and downstream samples. It includes the same controls and fixed effects as the main analysis (cf. Table 2), and thus corrects for previous trends as we include country and mine sub-basin trends. Both upstream and downstream villages do not display any pre-trends, which suggests that the common trend assumption is verified. We observe almost no effect of a mine opening on the mortality rates upstream, only a slight decrease which is significant 10 years after the mine opening. Figure 10 shows that the infant mortality downstream increases once the mine opens. This effect is significant in the medium run, around a decade after the mine opening.

Figure 10: Event study - dynamic effect of mine opening on 24-month mortality



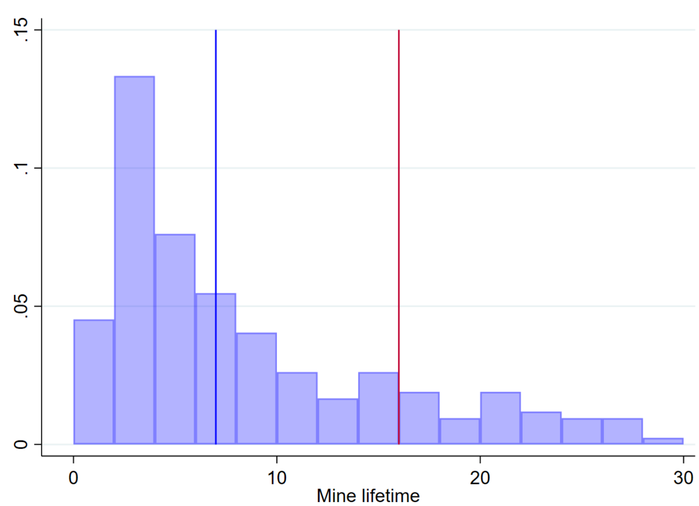
Notes: This Figure plots the event study of the effect of mine opening for downstream and upstream DHS villages 10 years before the mine opening and 10 years after. The year before the mine opening, -1, is taken as the reference point.
Sources: Authors' elaboration on DHS and SNL data.

8.2 Mine closure

In the main analysis, we focus on mine opening without taking into account the closure year, due to the difficulty in manually obtaining this information. In this section we focus on a restricted sample of mines, for which the SNL database directly provides this closure dates. This allows us to investigate whether the closure date has an impact on our main effect.

Figure 11 gives the distribution of the mine life for the restricted sample of mines for which closure years are available. On average, a mine lasts 16 years, but the distribution is skewed to the right and the majority of mines close before 10 years. Please note that the closing date available in the SNL database for this restricted sample is not exact. Over its lifetime, a mine can be put on hold several times for political or economic reasons. In Table 10 we look at the effects of being downstream of a mine that is active at the year of birth of the child. Columns (1) and (2) show no effect on the 12-month mortality rates. Column (3) shows that being downstream of a mine that is active in the year of birth increases the mortality rate by 4 p.p, which corresponds to an increase of 40% in the mortality rate. This result suggests that the harmful effects of mining activity on individuals living downstream are critical mainly while the mine is active.

Figure 11: Distribution of a mine life



Notes: This figure gives the distribution of mine life. The red line y=16 plots the mean of mine life, while the blue line y=7 plots the median. The maximum mine life in our sample is 138 years.
Sources: Authors' elaboration on SNL data.

Table 10: Average effects of mine activity on infantile mortality

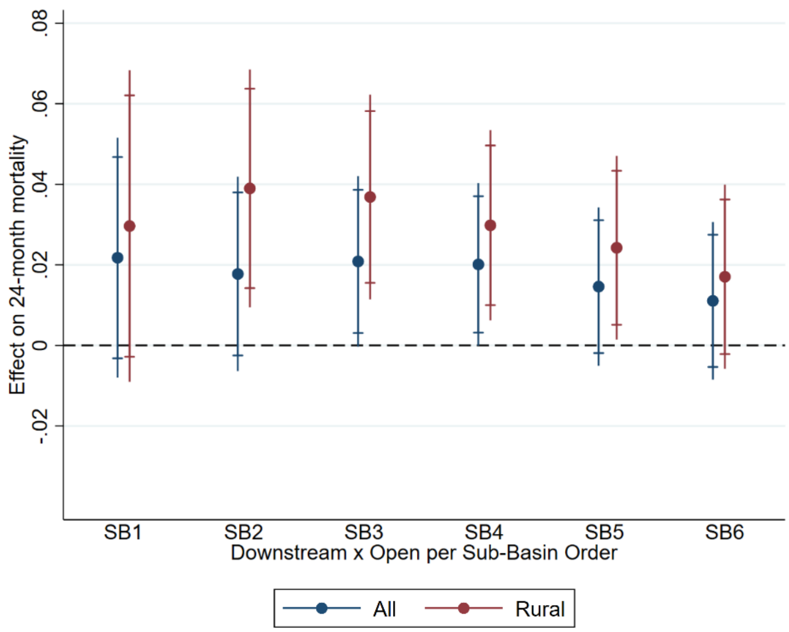
	Mortality under 12 months		Mortality under 24 months	
	All	Rural	All	Rural
	(1)	(2)	(3)	(4)
Downstream×Active	0.0112	0.0321	0.0409*	0.0488**
	[0.0222]	[0.0254]	[0.0248]	[0.0242]
Downstream	-0.0264*	-0.0434**	0.00877	-0.00721
	[0.0139]	[0.0184]	[0.0145]	[0.0184]
Active	0.000600	0.00420	-0.00842	-0.00345
	[0.0140]	[0.0173]	[0.0172]	[0.0193]
Controls	Yes	Yes	Yes	Yes
Nb open mines	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes
N	7,231	5,589	5,270	4,082
R2	0.0899	0.0960	0.0984	0.108
Outcome Mean	0.0756	0.0825	0.0981	0.104

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$. The variables Downstream and Active are dummies that indicate whether the individual lives in a village downstream of at least one mining site and whether the site is active during the year of birth. The same sample and controls as Table 2 Column 2 apply.

9 Intensive margin

9.1 Spatial intensive margin

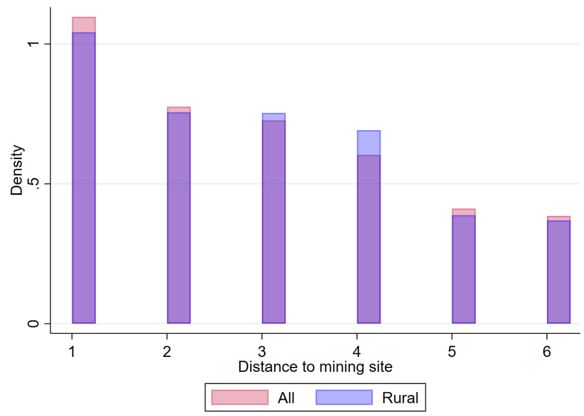
Figure 12: Effect of industrial mine opening according to the downstream sub-basin order



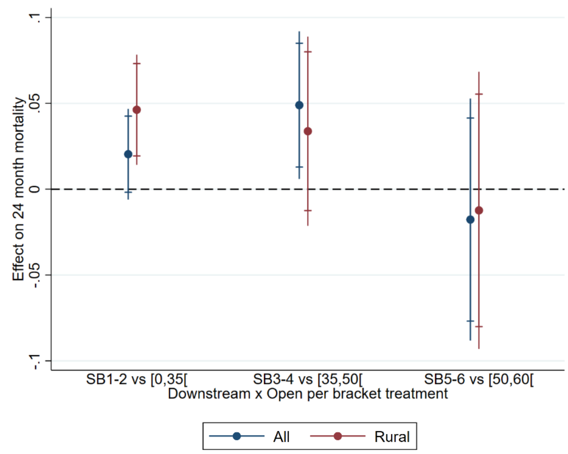
Notes: This Figure plots the treatment effect when changing the treatment group. SB1 gives the DiD estimator when the control group includes DHS villages up to the first sub-basin, SB2 up to the second, and SB6 up to the sixth sub-basin. SB3 provides the main results of this paper.
Sources: Authors' elaboration on DHS and SNL data.

In this section, we change the cut-off for being treated and test the effect on different orders of downstream sub-basins. In Figure 12, we test whether the effect holds when allowing for further sub-basins downstream. It plots the coefficient on *Downstream* $\times$ *Open* for six different models. SB1 corresponds to the model where the treatment group includes only the first neighboring downstream sub-basin, SB2 up to the second, and SB6 up to the sixth. SB3 gives the main results from column (6) in Table 2. The Figure shows an attenuation of the magnitude of the effect when including further sub-basins. The results are significant at the 5% level for all the individuals up to the third sub-basin and up to the fourth, while for all rural areas it is significant from the second up to the fifth sub-basin. We interpret the non-significance of the results in the first sub-basins as being the consequence of statistical power (as the sample size is relatively low up to SB1 and SB2).

Figure 13: Intensive margin - effect of the number of mine openings on under 24-month mortality



(a)



(b)

Notes: Panel (a) plots the distribution of observations downstream per sub-basin order. Panel (b) plots the interaction term on 24-month mortality rate per distance brackets. The first coefficient on the left gives the effect for individuals living within the first and second sub-basins compared to individuals living withing 35 km of the mine. The second coefficient gives the effect for individuals living downstream within the third and fourth sub-basins compared to those living upstream between 35 and 50 km. The third coefficient gives the effect for individuals living downstream within the fifth and sixth sub-basins compared to those living upstream between 50 and 60 km. The distances are chosen based on the mean of the distance between the mine and the extremity of the XXth sub-basin. [a]
Sources: Authors' elaboration on DHS and SNL data.

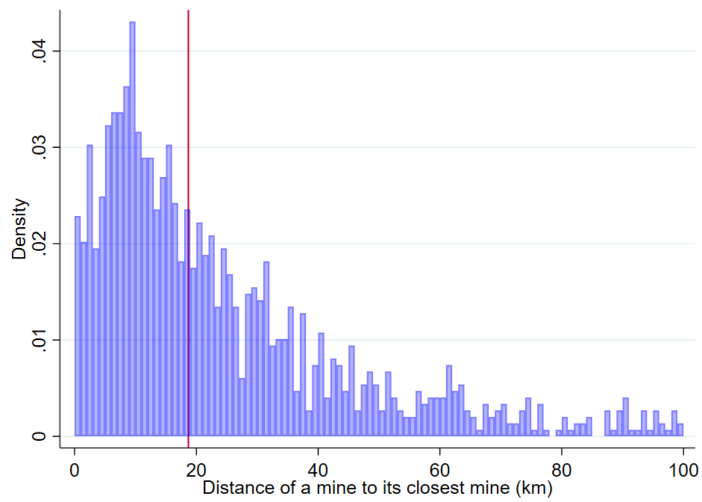
[a]The mean of the distance between the mine and the extremity of the first sub-basin is 27 km. It is 37 km for SB2, 45 km for SB3, 46 km for SB4, 58 km for SB5 and 59 km for SB6

Figure 13 looks at the effect per distance brackets. To build the control group for each sub-basin, we determined which distances correspond to which sub-basin order. We calculated the mean of the maximal distance between the mine and the furthest extremity of each sub-basin order. On average, a mine is 27 km away from the furthest extremity of sub-basin 1, 37 km from sub-basin 2, 45 km from sub-basin 3, 46 km from sub-basin 4, 58 km from sub-basin 5 and 59 km from sub-basin 6. Following this indicator, we compare in Figure 13 the individuals living downstream within the first and second sub-basins to those living upstream within 35 km of the mine (coefficient SB1-2). We then compare individuals living downstream within the third and fourth sub-basins to those living upstream within 35 to 50 km of the mine (coefficient SB3-4). Finally, we compare individuals living downstream within the fifth and sixth sub-basin to those living within

50 to 60 kilometers of the mine. Figure 13 shows that our results are mainly driven by the effect within the third and fourth sub-basins, while in rural areas the effect is only significant within the closest sub-basins. This shows that the effect is critical close to the mine, where pollution is supposed to be the highest. The difference between the whole and rural samples might be explained by the fact that the locations of mines close to urban areas suffer from lower precision (cf. section 10.3.2).

9.2 Mine density

Figure 14: Distance of mines to each other

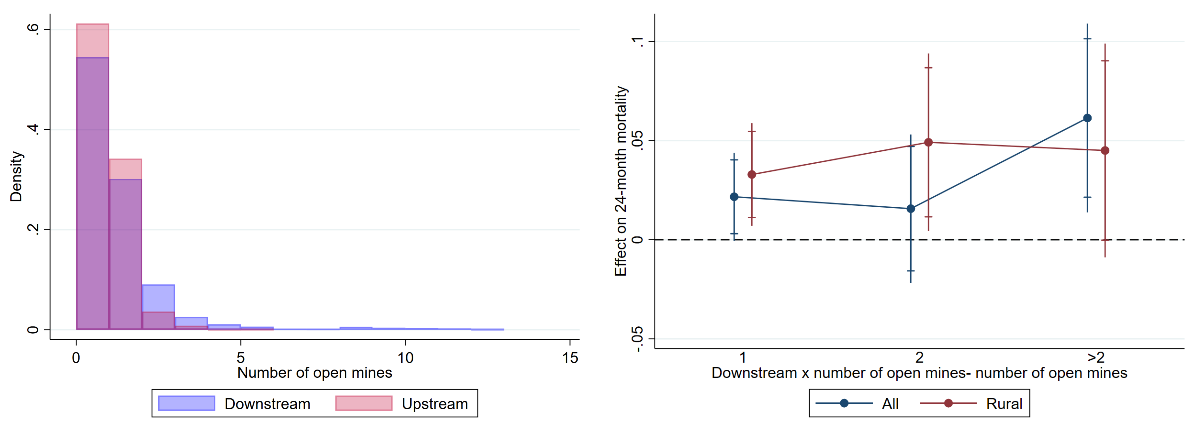


Notes: This Figure gives the distribution of the distance between each mine and its closest industrial mining site, providing insights into how far these mines are located from each other. The median distance (18 km) is plotted in red, and the graph presents distances under 100 km (please note that the maximum distance is up to 474 km).
Sources: Authors' elaboration on DHS and SNL data.

In this section, we explore the intensive margin of our results according to mine density. Figure 14 gives insights into how far these mines are located from each other. It shows the distribution graph of the distance to the closest mine for each mine. It is made with no regard for the activity status of the mining site. On average, a mine is located at least 31 km from its closest mining site, while the median is 18 km. The distribution is skewed to the right, showing that the majority of mining sites are located in areas with a high density of mining activity. Few mines are isolated, up to 100 km from the closest mining site. This graph shows the necessity first to control for the number of open mines within the area in the main analysis (Table 2), and second to look at heterogeneous effects according to the intensity of the mining activity within the area.

Figure 15a plots the frequency of the number of open mines within our main sample, both for upstream and downstream areas, within 45 km and up to the third sub-basin. As the number of observations falls starting at 3 open mines, we investigate in Figure 15b and Table 11 the statistical difference between being downstream of a single open site, two open sites, or more than two. For the whole sample, being downstream of one open mine increases the 24-month mortality by 2 p.p. The effect increases when the number of open mines increases, as being downstream of more than 2 open mines increases the mortality rate by 6 p.p in comparison to being downstream of only one open mine. Table 11 column (1) gives the DiD interaction term when open becomes a continuous variable and not a dummy, being the number of open mines. It shows that being downstream one additional mine that opens increases the mortality by 1.3 p.p, and by 2 p.p within rural areas.

Figure 15: Intensive margin - effect of the number of mine openings on 24-month mortality



(a) (b)

Notes: Panel (a) plots the distribution of the number of open mines across downstream and upstream villages. Panel (b) plots the interaction variable on the 24-month mortality rates. It gives the average treatment effects of the number of open mines on 24-month mortality.
Sources: Authors' elaboration on DHS and SNL data.

Table 11: Effects of the number mine opening on infantile mortality according to the number of open mine

	24-month mortality			
	All		Rural	
	(1)	(2)	(3)	(4)
Downstream×Nb open	0.0130***		0.0205***	
	[0.00484]		[0.00744]	
Downstream×Nb open=1		0.0217*		0.0329**
		[0.0113]		[0.0132]
Downstream×Nb open=2		0.0157		0.0492**
		[0.0191]		[0.0228]
Downstream×Nb open >2		0.0614**		0.0451
		[0.0243]		[0.0275]
Downstream	-0.0214***	-0.0208***	-0.0295***	-0.0308***
	[0.00723]	[0.00767]	[0.00823]	[0.00837]
Nb Open	-0.00419		-0.00842	
	[0.00526]		[0.00613]	
Nb Open =1		-0.0000134		-0.00328
		[0.00785]		[0.00895]
Nb Open =2		-0.00850		-0.0220
		[0.0129]		[0.0141]
Nb Open >2		-0.0407**		-0.0329
		[0.0190]		[0.0219]
Controls	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes
N	35,638	35,638	24,544	24,544
R2	0.0511	0.0511	0.0633	0.0634

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$. Columns (1-3) give the results for the total population while columns (4-6) display the results for rural villages. Columns (2, 4, 6, 8) control for the number of open mines within 45 km. The same sample as Table 2 Column 2 applies. Control variables are birth order number, mother's age, mothers' age square, mother's years of education and urban, and a continuous variable indicating the presence of rivers and their order.

9.3 Production intensive margin

Table 12: Effects of industrial mining opening, across each commodity's price evolution.

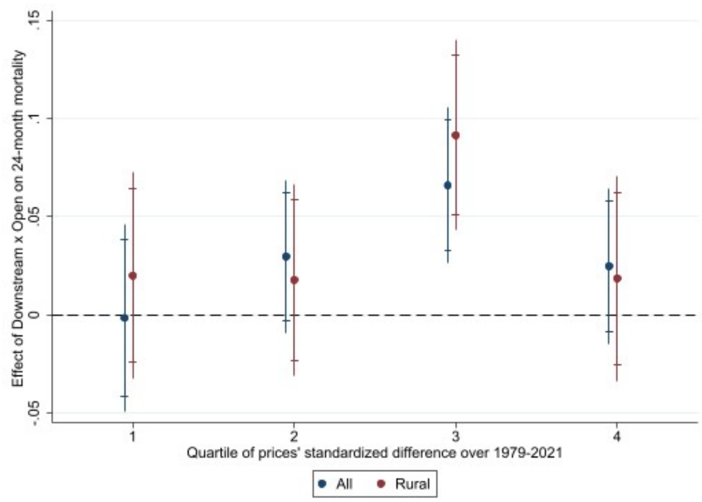
Outcome	24-month mortality	
	(1)	(2)
Price var.	Standardized difference	Z-score
Downstream × Open× Price var	0.0160*	0.0101**
	[0.00823]	[0.00463]
Downstream	-0.00244	0.00102
	[0.0104]	[0.0110]
Controls	Yes	Yes
Country-survey year FE	Yes	Yes
Mine SB FE	Yes	Yes
Mine SB-survey year trend	Yes	Yes
Commodity FE	Yes	Yes
N	31,517	31,517
R2	0.0509	0.0509
Outcome mean	0.0907	0.0907

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$. Standardized difference and Z-score of each commodity's price calculated over 1979-2021. The same sample and controls as Table 2 Column 2 apply.

We proxy the production intensity of each mine by using the global price of each mine's primary commodity as done in Berman et al. [2017] and Girard et al. [2022]. We presume that the higher the prices, the more intense the production. Annual prices were retrieved from the SNL data (coal, gold, lead, nickel, platinum, silver, and zinc) and the World Bank Pink Sheet (copper). For each commodity, we calculate the average price over 1971-2021, and for each year we calculate the standardized difference ($\frac{Price_t - \overline{Price_{[1971-2021]}}}{\overline{Price_{[1971-2021]}}}$) and Z-score ($\frac{Price_t - \overline{Price_{[1971-2021]}}}{\sigma_{Price_{[1971-2021]}}}$). We find a positive and significant effect of the triple interaction (Table 12) with both price variables (column (1) for the standardized difference and column (2) for the Z-score).

We then plot the coefficients associated with the interaction term of being downstream of an open mine across the quartiles of the change in the z-score (Figure 16). For both total and rural samples, we find that an increase in prices (going from the second to the third quartile) leads to an even higher effect of industrial mining on 24-month mortality.

Figure 16: Effect of living downstream of an open mine across the evolution of the price of the mine's primary commodity



Notes: This graph represents the coefficients associated with being downstream of an open mine across the evolution of the price of each mine's primary commodity (available for coal, copper, gold, lead, nickel, platinum, silver, and zinc). The standardized difference to the mean over the 1979-2021 period was calculated For each commodity and then split across quartiles to grasp the relative price evolution specific to each type of commodity.
Sources: Authors' elaboration using DHS, SNL, and World Bank Pink Sheet data.

10 Robustness checks

10.1 Balanced sample and de Chaisemartin and d'Haultfœuille [2020]

10.1.1 Balanced sample

One issue of working with DHS data is dealing with repetitive cross-sections instead of an exact panel. In this section, we define a balanced sample as a restricted sample for which each mine has observations before and after its opening, both upstream and downstream. In this sense, it is a balanced panel of mine, if we consider only two points in time which are: (1) the period before the mine opens, and (2) the period after its opening. Please note that, in this paper, it is possible to restrict the analysis to a balanced sample by the extension of the sample size of mines, and it underlines the limits of analyses looking at dynamic effects while using only a few mines. In this section, we first define the balanced sample and then we replicate the main analysis on this restricted sample.

First, let us define the construction of the balanced sample. A staggered DiD is driven by changes in mortality rates of switchers, which are observations that change treatment status, in comparison to those that do not change treatment status. In the case of a balanced sample, the design of this paper distinguishes three groups of observations:

- Group 1, "the switchers": the subgroup for which a mine has opened between two different years (for which there are DHS observations) and thus for which the treatment status changes from 0 to 1.
- Group 2, "the always treated": the subgroup of areas for which the mine has always been opened and are thus always treated, i.e., the treatment variable, which is an interaction, is always equal to 1.
- Group 3, "the never treated": the subgroup of areas where mines have not yet opened. The treatment variable is equal to 0. The third group comprises subgroups where mines have not opened as of 2022 but their opening is planned for the future (the mine is projected to open). This group also includes mines where no DHS cluster was surveyed after the mine opened. Though this group is called "the never treated", it includes both DHS villages that will never be treated or have not been treated because they will be treated in the future.

The balanced sample makes it possible to identify the three groups. Formally, it is defined as the following, for each group:

Let's consider observations that can be divided into G groups and T periods. For every $(g, t) \in \{1, ..., G\} \times \{1, ...T\}$, let $N_{g,t}$ denote the number of observations in the group g and period t, and let $N = \sum_{g,t} N_{g,t}$ be the total number of observations. For all (g, t) $\in \{1, ..., G\} \times \{1, ..., T\}$, let us call $D_{g,t}$ the $Downstream_{g,t}$ variable and $O_{g,t}$ the $Opened_{g,t}$ variable.

Definition 1 (Group 1 - Balanced sample of "switchers")**.** *Let us call* $G_1 = \{g_0, ..., g_{n_1}\}$ *the set of Group 1. Group 1 is defined as the following:*
For all $g \in G_1, \exists(v_1, v_2, v_3, v_4) \times (t_1, t_2, t_3, t_4) \in g \times T$ *such that:*
(i) $N_{v_1,t_1} > 0 \wedge D_{v_1,t_1} = 0 \wedge O_{v_1,t_1} = 0$
(ii) $N_{v_2,t_2} > 0 \wedge D_{v_2,t_2} = 1 \wedge O_{v_2,t_2} = 0$
(iii) $N_{v_3,t_3} > 0 \wedge D_{v_3,t_3} = 0 \wedge O_{v_3,t_3} = 1$
(iv) $N_{v_4,t_4} > 0 \wedge D_{v_4,t_4} = 1 \wedge O_{v_4,t_4} = 1$
In our setting, g *is the whole area associated with a mine, including both upstream and downstream observations, and is made up of* $k \in N$ *DHS clusters such as* $g = \{v_1, ..., v_k\}$.

Definition 2 (Group 2 - Balanced sample of "always treated"). *Let us call* $G_2 = \{g_0, ..., g_{n_2}\}$ *the set of Group 2. Group 2 is defined as the following:*
For all $g \in G_2, \exists (v_1, v_2) \times (t_1, t_2) \in g \times T$ *such that:*
(i) $N_{v_1,t_1} > 0 \wedge D_{v_1,t_1} = 0 \wedge O_{v_1,t_1} = 1$
(ii)$N_{v_2,t_2} > 0 \wedge D_{v_2,t_2} = 1 \wedge O_{v_2,t_2} = 1$

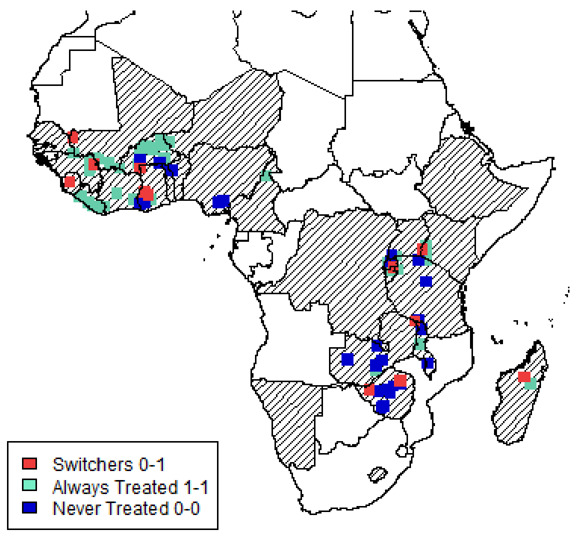
Definition 3 (Group 3 - Balanced sample of "never treated"). *Let us call* $G_3 = \{g_0, ..., g_{n_3}\}$ *the set of Group 3. Group 3 is defined as the following:*
For all $g \in G_3, \exists (v_1, v_2) \times (t_1, t_2) \in g \times T$ *such that:*
(i) $N_{v_1,t_1} > 0 \wedge D_{v_1,t_1} = 0 \wedge O_{v_1,t_1} = 0$
(ii)$N_{v_2,t_2} > 0 \wedge D_{v_2,t_2} = 1 \wedge O_{v_2,t_2} = 0$

Definitions 1, 2, and 3 define the treatment group (Group 1) and the control groups (Groups 2 and 3) in the setting of the mine-balanced sample. For Group 1 switchers, we restrict the sample to areas that have DHS clusters surveyed both downstream and upstream, both before and after the opening of the mine. This means that we select areas that have been surveyed in at least four different locations by the DHS. For Groups 2 and 3, meaning that the surveys occurred only after the opening (Group 2) or before the opening (Group 3), we restrict to areas that have observations both upstream and downstream.

Figure 17 plots the different groups from the balanced sample and displays the group of "switchers", the "always-treated" and the "never-treated" groups. It displays the areas that are key to the main estimation, which are located in Western Africa, Zimbabwe, Western Kenya, Rwanda, Tanzania, and Madagascar. Table 30 in the Appendix plots our main estimator across the African sub-regions, and shows that our results are mainly driven by Western Africa and remain significant in Eastern Africa. Table 13 gives the size of the three groups in the balanced sample, as well as the associated number of mines. It shows that Group 1 "switchers" account for around 12% of the total balanced sample, and corresponds to the neighborhood of 13 mines. In total, the control groups gather 75 mines. We observe that the average mortality rates have decreased over time, before and after the opening of the mining site in Group 1. This is linked to the decrease in infant mortality in Africa over time, which is consistent with balance Table 1 and Figures 4 and 5, and which highlights the importance of controlling for trends.

Table 14 displays the balance table for the restricted sample. This table shows the within comparison before and after a mine opens both downstream and upstream. It shows only descriptive statistics and neither accounts for control variables nor for fixed effects. The table shows that there is a significant difference between upstream and downstream areas after a mine opens concerning both the 12- and 24-month mortality rates. From a descriptive point of view, being downstream of a mine increases the 12-month mortality rate by 2.7 p.p, and the 24-month mortality by 2.4 p.p (column (13)). This difference is explained by a significant decrease in mortality rate within upstream areas after a mine opening (column (11)).

Figure 17: Balanced panel - group identification



Notes: The Figure plots the areas of the groups across the three groups of the balanced panel, for the 24-month mortality rates.
Sources: Authors' elaboration on DHS and SNL data.

Table 13: Balanced Sample - Descriptive Statistics

	Group 1 : Switchers 0-1						Groups 2+3		Group 2 : 1-1		Group 3 : 0-0	
	All		Before Opening		After Opening							
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dth<12												
All	1,191	0.07	894	0.069	297	0.071	8,423	0.07	2,368	0.056	6,055	0.076
		(0.255)		(0.254)		(0.257)		(0.256)		(0.23)		(0.265)
Mines	13		13		13		75		31		44	
Dth<24												
All	1,191	0.089	894	0.091	297	0.084	8,423	0.091	2,368	0.072	6,055	0.099
		(0.285)		(0.287)		(0.278)		(0.288)		(0.258)		(0.298)
Mines	13		13		13		75		31		44	

Notes: Standard errors and p-values are in parentheses. Outcomes' descriptive statistics of under 12-and 24-month mortality, for villages within the Group 1 Switchers for individuals born before and after the opening of the mine, then Group 2 always treated and Group 3 never treated.

Table 14: Balance Table

	Before Mine Opening					After Mine Opening					Within Up.	Within Dwn.	Within
	Upstream		Downstream		Diff	Upstream		Downstream.		Diff			
	N	Mean /(SD)	N	Mean /(SD)	(4-2) /(p.v)	N	Mean /(SD)	N	Mean /(SD)	(9-7) /(p.v)	(7-2) /(p.v)	(9-4) /(p.v)	(12-11) /(p.v)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Dth<12													
All	5,272	0.079	1677	0.063	-0.016	1,812	0.054	853	0.066	0.012	-0.025	0.002	0.027
		(0.27)		(0.243)	(0.025)		(0.226)		(0.248)	(0.248)	(0)	(0.814)	(0.005)
Mines	54		56			38		37					
Dth<24													
All	5,272	0.101	1677	0.088	-0.012	1,812	0.07	853	0.081	0.011	-0.031	-0.007	0.024
		(0.301)		(0.284)	(0.123)		(0.254)		(0.273)	(0.306)	(0)	(0.527)	(0.03)
Mines	54		56			38		37					

Notes: Standard errors and p-values in parentheses. Descriptive statistics of 12-month and 24-month mortality outcomes, for villages upstream and downstream of mining sites, for individuals born before and after the opening of a mine, over the balanced sample.

10.1.2 Heterogeneous treatment effects with two-way fixed effects: de Chaisemartin and d'Haultfœuille [2020]

The main result of this paper estimates the effect of being downstream of an open mine by using standard DiD designs. However, recent developments in the estimation of DiDs in staggered adoption designs ([Borusyak et al., 2021; Goodman-Bacon, 2018; Callaway and Sant'Anna, 2021; de Chaisemartin and d'Haultfœuille, 2020]) show that the estimated ATT[12] is a weighted sum of different ATTs with weights that may be negative. The negative weights are an issue when the treatment effect is heterogeneous between groups over time, as one could have the treatment coefficient in those regressions as negative while the treatment effect is positive in every group and time period. Using treated observations as controls creates these negative weights. In our design, the effect on Group 1 "switchers" is compared to two control groups, Group 2 "always treated" and Group 3 "never treated". The negative weights might come from the comparison of the effect of the Group 1 "switchers" to the Group 2 "always treated". This biases the DiD estimator as it is an average of local treatment effects. In this section, we use the de Chaisemartin and d'Haultfœuille [2020] estimator, which deals with the issue of negative weights in a staggered adoption design.

Table 15 compares the two-way fixed effects (TWFE) used in the main results (odd columns), to the de Chaisemartin and d'Haultfœuille [2020] estimator (dCDH) (even columns) [13]. Columns (1), (2), (5) and (6) give the results for the entire sample, while columns (3), (4), (7) and (8) for the balanced sample, defined in Section 10.1.1. Columns (1)-(4) give the results for the 12-month mortality rates, while columns (5)-(8) for the 24-month mortality rates.

First, let us look at the 24-month mortality rates. When looking at the TWFE estimator, we see that the results are stable on the balanced sample, even though it only represents 27% of the whole sample. This is consistent with the fact that the balanced sample keeps the villages that drive the main estimation's results. When focusing on the balanced sample, being downstream of an opened mine increases the 24 month-mortality rates by 3.19 p.p, which represents an increase of 36%. Column (6) gives the dCDH estimator for the whole sample, while column (4) is for the balanced sample. We observe an increase in terms of the magnitude of the effect when correcting for negative weights, as being downstream of an open mine increases the 24-month mortality rate by 11 p.p,

[12] Average Treatment on the Treated

[13] The Stata command *did_multiplegt* is used to run the dCDH estimator.

which represents an increase of 129%. If these magnitudes seem high, it is reassuring to observe the stability of the direction and significance of our main effect when using the dCDH estimator. Regarding the 12-month mortality rate, we observe that the restriction to the balanced sample displays a 2.8 p.p increase.

Table 15: Effects of industrial mining opening on 24-month mortality de Chaisemartin and d'Haultfœuille [2020]

	12-month mortality				24-month mortality			
	Whole Sample		Balanced Sample		Whole Sample		Balanced Sample	
	TWFE	dCDH	TWFE	dCDH	TWFE	dCDH	TWFE	dCDH
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Downstream×Open	-0.00506	0.0249	0.0286*	0.1457	0.0218**	0.1109**	0.0319**	0.1667*
	[0.00831]	[0.0262]	[0.00222]	[0.1132]	[0.0108]	[0.0405]	[0.0162]	[0.1112]
Downstream	-0.0152**		-0.0242***		-0.0211***		-0.0283***	
	[0.00665]		[0.00826]		[0.00739]		[0.00866]	
Open	0.00963		0.0347		-0.00496		0.0238	
	[0.00754]		[0.0320]		[0.0101]		[0.0347]	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	48, 472	48, 472	9, 606	9, 606	35, 638	35, 638	9, 606	9, 606
R2	0.0378		0.0523		0.0511		0.0599	

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. Column (1) gives the result of the main anlysis, the Two Way Fixed Effect (TWFE) for the whole sample, while Column (2) gives the de Chaisemartin and d'Haultfœuille [2020] estimator. Columns (3) and (4) give the TWFE and de Chaisemartin and d'Haultfœuille [2020] estimators for the balanced sample.

10.2 Sensitivity analysis

10.2.1 Including non-topographical sub-basins

In this section, we replicate the main analysis from Table 2, adding to the control group individuals living in a sub-basin with no topographical relation to the mine sub-basin, within 45 km. This test can have several readings.

First, it strengthens the control for income effects linked to mining activity, enabling us to more precisely isolate the channel of water pollution, and exclude other potential mechanisms. Indeed, villages close to the mine, but located in a sub-basin with no topographical relationship to the mine, are allegedly less exposed to mining-induced water pollution and would be as exposed to income or labor effects, conflicts, or migration.

Yet, it also leads to the comparison of villages that do not necessarily share the same water resources, and this could blur the interpretation of our estimation. For example, other activities such as more intensive agriculture or livestock farming could aggregate around the mining site and could be responsible for other types of pollution. If these activities are located in a sub-basin with no topographical relationship to the mine, the estimated comparison would display the difference between the pollution of the mine and the pollution of these activities, rather than the pollution of the mine only, and this would lead to a downward bias in our analysis. Moreover, as mining activity is water intensive, the location of these activities might also be endogenous to the location of the mine, and this could induce an even larger downward bias.

Table 16 displays the results when including the non-topographical sub-basins within the control group. As expected, the table suggests that the main results of the 24-month mortality rates are downward biased and only significant at the 10% level for the rural population.

Table 16: Effects of industrial mining opening on infantile mortality - including DHS with non-topographic relationship

	12-month mortality				24-month mortality			
	Total Population		Rural Population		Total Population		Rural Population	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Downstream×Open	-0.00266	-0.00227	0.00341	0.00381	0.00746	0.00804	0.0172*	0.0165*
	[0.00543]	[0.00544]	[0.00632]	[0.00632]	[0.00752]	[0.00753]	[0.00887]	[0.00888]
Downstream	-0.00490	-0.00472	-0.00904*	-0.00884*	-0.00486	-0.00455	-0.0103	-0.0106*
	[0.00445]	[0.00446]	[0.00506]	[0.00507]	[0.00561]	[0.00561]	[0.00637]	[0.00636]
Open	0.00364	0.00488	0.00318	0.00466	0.00115	0.00308	0.00283	0.000259
	[0.00292]	[0.00302]	[0.00345]	[0.00359]	[0.00377]	[0.00391]	[0.00457]	[0.00440]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nb open mines	No	Yes	No	Yes	No	Yes	No	Yes
Birthmonth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-bthyr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB-bthyr trd	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	168,931	168,931	123,413	123,413	124,670	124,670	91,395	91395
R2	0.0214	0.0215	0.0252	0.0252	0.0305	0.0305	0.0356	0.0355
Outcome Mean	0.0638	0.0638	0.0670	0.0670	0.0824	0.0824	0.0872	0.0872

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. Columns (1-3) give the results for the total population while columns (4-6) display the results for rural villages. The same controls as Table 2 apply. Columns (2, 4, 6, 8) control for the number of open mines within 45 km. The sample includes individuals living in non-topographic sub-basins within 45km.

10.2.2 Dropping fixed effects and other tests

Table 17: Effects of industrial mining opening on 24 months mortality, while dropping fixed-effects.

Outcome	24-month mortality			
	(1)	(2)	(3)	(4)
Downstream×Open	0.0218**	0.0216**	0.0179*	0.0177*
	[0.0108]	[0.0108]	[0.0105]	[0.0104]
Downstream	-0.0211***	-0.0211***	-0.0218***	-0.0219***
	[0.00739]	[0.00738]	[0.00734]	[0.00733]
Open	-0.00496	-0.00494	-0.00459	-0.00466
	[0.0101]	[0.0101]	[0.00972]	[0.00962]
Controls	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	No	No	No
Commodity FE	Yes	Yes	No	No
Mine SB-birthyear trend	Yes	Yes	Yes	No
N	35,638	35,638	35,638	35,638
R2	0.0511	0.0504	0.0503	0.0491
Outcome mean	0.0873	0.0873	0.0873	0.0873

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. The same sample as Table 2 Column 2 apply.

We find stability in our results when dropping fixed effects one by one: birth month primary commodity, and sub-basins birth year trend (Table 17) until keeping the two-way fixed effects (i.e., keeping the mine sub-basin fixed effects and the country birth year fixed effects).

Section F.1 runs other tests. Table 31 shows that our result is stable when controlling for the hand work. Figure 37 shows that the main results are stable when dropping countries one by one, and Figure 36 when dropping metals one by one.

10.2.3 Spatial correlation

As an additional robustness check, we run our main result's specification while taking into account the spatial correlation of DHS clusters. We estimate the standard errors with a spatial HAC correction following the method developed by **?** and using the Stata command introduced by Colella et al. [2019]. Table 18 shows the stability of our results for different cut-off distances of spatial correlation (from 20 km to 200 km). We did not directly include the **?** test in the main analysis, as it does not allow for several fixed effects.

Table 18 corrects for spatial correlation for the results when using only Mine-Subbasin and country birth year fixed effects (result from Table 17, column (4)).

Table 18: Effects of industrial mining activity, Conley spatial correction (acreg)

Outcome	Mortality under 24 months					
Conley spatial correction threshold	20 km	45 km	60 km	80 km	100 km	200 km
	(1)	(2)	(3)	(4)	(5)	(6)
Downstream×Open	0.0177*	0.0177*	0.0177*	0.0177*	0.0177*	0.0177*
	[0.0100]	[0.00999]	[0.0101]	[0.0101]	[0.00996]	[0.00916]
Downstream	-0.0219***	-0.0219***	-0.0219***	-0.0219***	-0.0219**	-0.0219**
	[0.00709]	[0.00746]	[0.00789]	[0.00847]	[0.00913]	[0.0106]
Open	-0.00466	-0.00466	-0.00466	-0.00466	-0.00466	-0.00466
	[0.00941]	[0.00932]	[0.00943]	[0.00960]	[0.00955]	[0.00938]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes
N	35,648	35,648	35,648	35,648	35,648	35,648
R2	0.00262	0.00262	0.00262	0.00262	0.00262	0.00262

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. The variables Downstream and Opened are dummies that indicate whether the individual lives in a village downstream of at least one mining site and whether the site opened before the birth year of the child. Each village DHS is paired to only one mining site so that each individual appears only once in the regression. Other variables are control variables. The sample focuses on heavy metal mines. Control variables are birth order number, mother's age, mother's age square, mother's years of education, urban, number of open mines, and presence of rivers.

10.3 Measurement errors

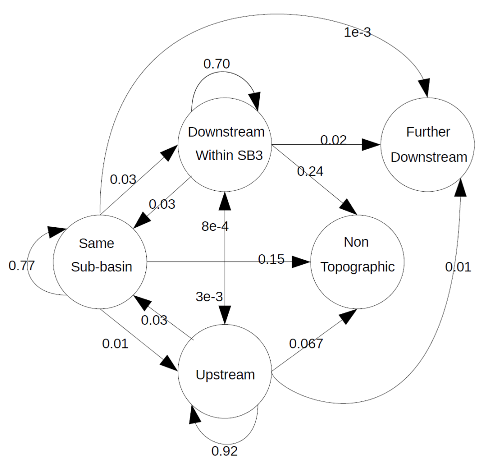
In this section, we deal with measurement errors that come from the nature of the data. Section 10.3.1 tests for the random displacement of DHS villages and section 10.3.2 tests our results according to the precision of the mine location.

10.3.1 DHS random displacement

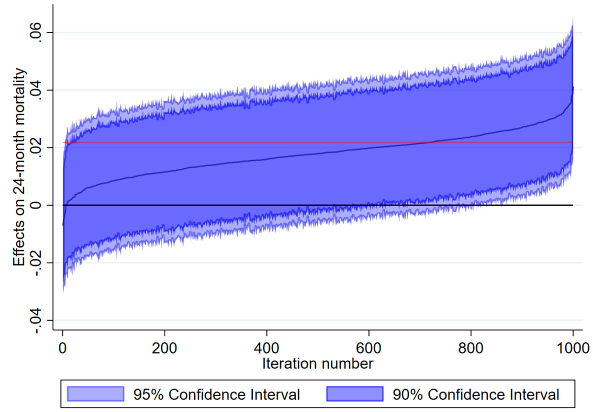
DHS randomly displaces the GPS coordinates of each village to protect the confidentiality of respondents. Urban locations are displaced within 2 km, while rural clusters are displaced within 5 km, with 1% of rural clusters moved up to 10 km. Displacements are made within administrative districts. This random reshuffling of DHS villages introduces measurement errors in our main estimation, all the more important as our treatment allocation depends on the relative position of the DHS villages to the mine.

First, we randomly displace each DHS village 1,000 times within a buffer of 2 km for urban clusters and 5 km for rural clusters. Thus, each displaced village is located in a new sub-basin, which can be the initial sub-basin or not. Then, we determine the topographical relation of this sub-basin to the sub-basin of the mine, which gives the treatment of the DHS village: whether it falls in a sub-basin upstream, downstream, in the same sub-basin as the mine, or in a sub-basin with no topographical relationship with the one of the mine. The topographical relation of the new sub-basin gives the new treatment status of the DHS village. We only reshuffled the position of DHS villages that have a topographical relation with the mine initially, and that were up to the third sub-basin downstream. This means that we can have some DHS villages that exit the main sample, for instance, if their newly assigned sub-basin has no topographical relation, or is downstream in the fourth sub-basin, or if it falls in the same sub-basin as the mine, since these cases are excluded from the main result. The only new observations that come within the sample are DHS villages that were initially within the same sub-basin as the mine and fall upstream or downstream with the new iteration of the random displacement of their location. Please note that it is possible as well, but very rare, that a DHS village falls in the ocean.

Figure 18: DHS random displacement - 1,000 iterations



(a)



(b)

Notes: Panel (a) plots the transition probability graph for 1,000 random displacements of DHS clusters. Panel (b) plots the interaction term for 1,000 different regressions, each done for a new sample where DHS GPS coordinates have been randomly displaced. The red line y=0.0218 plots the coefficient from our main result. The coefficients are ordered, and we plot the 95% and 90% intervals.
Sources: Authors' elaboration.

Figure 18a gives the probability graph showing the transition probabilities of changing treatment status. For instance, after 1,000 iterations, a DHS village initially downstream within the third sub-basin has a 70% chance of remaining downstream up to sub-basin three, has a 0.3% chance of being upstream of the mine, a 24% chance of falling into a sub-basin with no topographical relation (and be out of the sample), a 3.5% chance of being in the same sub-basin as the mine, and finally a 2% chance of being downstream further than the third sub-basin (and be out of the sample). In the end, a DHS village treated in our initial sample has a 25% chance of leaving the sample. Please note that this random reshuffling is not perfect; ideally DHS villages should be reshuffled within administrative level 2 boundaries following DHS procedure.

Figure 18b plots the interaction term $Donwstream \times Open$ of our main estimation for 1,000 random displacements of DHS GPS coordinates. The coefficients are ordered, and we plot the 95% and 90% intervals. As there is a higher probability that a DHS cluster leaves the sample rather than a new enters it, the number of observations varies for each iteration and is likely to be smaller than our main estimation.

10.3.2 Accuracy of mine location

We further test for potential measurement errors by looking at the precision of the mines' location. The SNL database provides information on the accuracy levels of each mine's GPS coordinates and enables us to restrict the analysis to mines with exact coordinates, precise to 1 km. Our main results are positive but no longer significant when restricting to the mines with exact coordinates, but hold when focusing on rural households. This hints towards a higher effect of industrial mining activity on child mortality among rural households, and a lack of precision in the location of mines close to urban areas.

Table 19: Effects of industrial mining opening, restriction to exact GPS coordinates.

Outcome	24-month mortality		
Accuracy level	All	Exact coordinates	
Sample	Urban and rural	Urban and rural	Rural
	(1)	(2)	(3)
Downstream×Open	0.0218**	0.0116	0.0294**
	[0.0108]	[0.0116]	[0.0143]
Downstream	-0.0211***	-0.0204**	-0.0239***
	[0.00739]	[0.00807]	[0.00888]
Open	-0.00496	0.00532	0.00805
	[0.0101]	[0.0104]	[0.0130]
Controls	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes
N	35,638	29,195	20,172
R2	0.0511	0.0517	0.0626
Outcome mean	0.0873	0.0858	0.0920

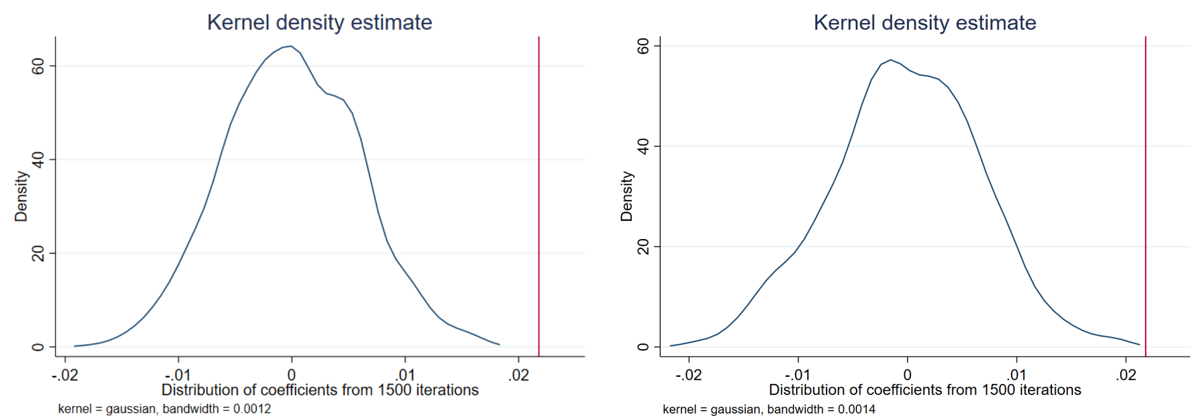
Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. The same controls as Table 2 Column 2 apply.

10.4 Placebo tests

10.4.1 Randomization inference

To make sure that the assignment of each village to its topographical position relative to the mine is indeed what drives our results on child mortality, we run a randomization inference test. We randomly draw 1,500 permutations of the "Downstream" variable without changing the start-up year, and 1,500 permutations of the "Open" variable without changing the downstream position [14]. The simulations show that the distribution of treatment effects (Downstream $\times$ Open) shift around zero (Figure 19). The red line represents the initial treatment effect using our main specification: we are sure at the 1% level that our main model is not misspecified.

Figure 19: Spatial and temporal randomization inference tests



(a) Randomization of the "Downstream" variable

(b) Randomization of the "Open" variable

Notes: The two panels represent the distribution of coefficients associated with the interaction term of being downstream of an open mine and its effect on under 24-month mortality when conducting 1,500 permutations of the "Downstream" position of each DHS sub-basin (Panel (a)), and 1,500 permutations of the "Open" variable (Panel (b)). The red line represents the initial treatment effect using our main specification.
Sources: Authors' elaboration using the Stata *ritest* command.

10.4.2 Placebo diseases

We conduct a placebo test on other potential diseases that could affect women's, and thus children's, mortality. We do not find significant industrial mining on the infection of any sexually transmitted disease among women living downstream of an open mine (Table 20,

[14] The randomization inference of the "Downstream" and "Open" treatment are within the sub-basin level, and are clustered at the DHS village level.

column (1)) or among awareness of tuberculosis (column (2)). This absence of differential results on women's health across upstream and downstream villages is reassuring for our identification of the water pollution channel.

Table 20: Effects of industrial mining opening on women, placebo diseases.

	(1)	(2)
Outcome	Any sexually transmitted infection	Heard of tuberculosis
Downstream × Open	0.00332 [0.00923]	-0.0387 [0.0259]
Downstream	0.00751 [0.00772]	0.0277 [0.0210]
Open	0.00766 [0.00913]	0.00469 [0.0314]
Controls	Yes	Yes
Country-survey year FE	Yes	Yes
Mine SB FE	Yes	Yes
Mine SB-survey year trend	Yes	Yes
Commodity FE	Yes	Yes
N	66,653	14,750
R2	0.0888	0.186
Outcome mean	0.0501	0.938

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. The same sample and controls as Table 4 apply.

11 Discussion and limits

11.1 Selection issues

Our study includes the best available data on mining and child health we have at the scale of a continent, but of course, is neither exhaustive nor represents the whole African continent. First, as we limited our sample to countries with at least two waves of DHS, we had to drop many countries with only one wave. Among these was South Africa, which has intense mining activity. Future work will be possible once other survey waves have been conducted. Our 26 countries represent about two-thirds of the total population of the continent[15].

Second, our mining data is also limited to industrial sites, and do not include artisanal or small-scale mining for which information is much harder to retrieve at the scale of the continent. One possibility for another project is to use data on the suitability of

[15] The proportion is stable between 1981 and 2020.

artisanal gold mining [Girard et al., 2022] and compare its environmental impact to that of industrial mining.

One remaining concern is about the exhaustivity of the SNL, and the heterogeneity of the sampling selection across countries. We tried our best to evaluate the exhaustivity of the SNL data by comparing it with other mining data and social sources (Ministries, USGS, mining websites. . .) but it is out of our feasible means to get an exact proportion of representativity and to compete with the business-oriented activity of SNL.

11.2 Threats to the identifying assumption

11.2.1 Type of pollution

First, our study mainly focuses on water pollution through the lens of water sub-basins, while controlling for rivers, i.e., surface water. Yet, further work could also control for groundwater as its pollution may follow different dynamics than surface water: villages located in areas with low-depth groundwater could pump the water more easily and with more affordable water pumps than those in areas with deeper groundwater. The former villages could therefore be more exposed to mining-induced water pollution than the latter ones. Moreover, groundwater could take more time to be contaminated by mining-induced pollution than surface water, but its contamination could also be more permanent.

This paper does not directly examine mining-induced air pollution. The main hypothesis is that wind direction is less correlated with the topographical position of the village than water pollution, and that the comparison between upstream and downstream villages should exclude the effect of air pollution. Besides, as discussed in Section 2, the effects of air pollution seem to concern the mine workers more than the surrounding population, even though fine particles can be displaced over long distances. However, our main results are not entirely net off the impacts of air pollution. The best control included so far is adding the sub-basins with no topographical relationship to the mine, as they are allegedly exposed to air pollution only, while sub-basins with a topographical relationship would be exposed to both water and air pollution. It is beyond the scope of the current paper to take into account the direction of the wind to disentangle both sources of pollution, but could be empirically feasible for another paper. Our study is most likely an underestimation of the total pollution induced by industrial mining activity.

The same concern remains for soil pollution. An additional heterogeneity analysis would be to take into account areas prone to subsistence agriculture or livestock, as mining-induced water pollution could also contaminate soil and cattle in the long run, as well as the subsequent food produced. We assume more harmful effects of industrial mining on the health of local populations if both water and food are polluted. In another paper, one could study the heterogeneous effects across the global agroecological zones (GAEZ) and crop suitability, and test whether villages located near industrial mining sites in high-yield crop areas are more affected than villages with less suitable soil.

11.2.2 Threats to identification

A major threat to identification is that the opening of a mine may not be orthogonal to unobservable factors that affect health and water quality, in different ways for downstream and upstream areas.

Migration is a major methodological concern, as we show in Table 29 of Appendix D.1 that migrants significantly settle downstream after a mine opening. Section 6.3 shows that our main results are robust when controlling for in-migration. A main violation of the identifying assumption would be if downstream villages anticipate the mine opening and strategically out-migrate within upstream areas to avoid pollution. In this case, there would be a selection bias, as the individuals surveyed downstream after the mine opening would be those that were not able to migrate or anticipate the pollution. Controlling for in-migration in DHS villages, we show that our results are robust to this specific strategic behavior. However, we cannot control for strategic out-migration outside of the study area, meaning individuals out-migrating to avoid pollution somewhere other than the upstream area. In this paper, we made the choice not to use mother fixed-effects and retrospective questions on birth history, to limit endogenous selection due to out-migration, and to account for children born up to five years prior to the year of the survey.

Accordingly, a threat to the identification would be a differed improved access to infrastructure associated with the opening of a mine between upstream and downstream areas. Table 29 shows no difference in terms of access to electricity and piped water between downstream and upstream areas after a mine opening, and Section 6 shows that our result is robust when controlling for improved access to facilities.

An important omitted variable in our current study is the increased presence of conflict and violence around areas with mining activity, as shown by Berman et al. [2017]. This could also explain the increase in child mortality in the vicinity of mines. We do not directly control for conflict, but there would be an upward bias of our estimation only if conflicts systematically happen more downstream than upstream. As our results hold when including non-topographical sub-basins in rural areas, it is a first-step approximation that water pollution is indeed the main explanatory factor of increased child mortality. Further work could include the ACLED data to exclude this mechanism.

Another concern is that other industries could aggregate around mining industry and be partly responsible for pollution. More than a bias, this could be a threat to identification if the location of the industry is correlated with the topographical position of the mine. In another paper, we could look at the correlation between mining activity and other industry implementations. Controlling for them could enable us to isolate the pollution solely linked to mining activity.

12 Policy discussion

In this section, we first try to compute how many deaths were related to the water pollution linked to industrial mining activity in the 26 countries of our sample. Then, we try to assess whether the Extractive Industries Transparency Initiative, a global standard for good governance in the extractive sector, has been successful in reducing this mortality.

12.1 Back-of-the-envelope calculation

In this section, we compute a back-of-the-envelope calculation to grasp how many deaths could have been averted had there been policies implemented to limit water pollution, over the 1981-2020 period and within the 26 Sub-Saharan countries of our analysis.

First, we consider that as that DHS is representative at the national level, it is feasible to calculate the proportion of individuals living within 45 km of a mine, the proportion of those living downstream, etc. Here are the probabilities computed using the DHS database:

- $x = 28\%$: Proportion of individuals living within 45 km of a mine [16]

[16] Please note that this is exactly the proportion of individuals living within 45 km of a mine and

- $x = x_d + x_u + x_{nt} + x_{sb}$, with
  - $x_d = 1.94\%$: Proportion of individuals living downstream [17]
  - $x_u = 5.25\%$: Proportion of individuals living upstream
  - $x_{nt} = 17.93\%$: Proportion of individuals living with no topographical relation
  - $x_{sb} = 2.92\%$: Proportion of individuals living in the same sub-basin as a mine

Our analysis leads to an estimation of $e = 2.18\%$ the increased mortality rate because of industrial mining-induced water pollution. In our sample, 9,258 individuals live downstream of a mine and we count 822 deaths among them. The total number of additional deaths due to mining-induced water pollution is 9,258 $\times$ 2.18% = 202 deaths. We now look at the 880 million children aged 0-2 years during 1981-2020 in our 26 countries [18]. As we assume the representativity of the DHS surveys and the stable proportion of the population living in the vicinity of mines, this would mean that 1.94% $\times$ 880 million = 17 million children lived within 45 km downstream of a mine. This leads to 2.18% $\times$ 17 million = 370,600 deaths due to mining-induced water pollution over 1981-2020 in our 26 countries, i.e., 9,265 deaths per year, or 16 deaths per mine per year.[19] To get a better sense of the magnitude of this figure, there are on average 840,000 births per year per country (average within the 26 countries over 1981-2020), which means that the number of deaths caused by mining-induced water pollution over 26 countries represents 1.1% of the number of births per country[20].

12.2 Extractive Industries Transparency Initiative members

We look at whether there is a significant difference across countries that have signed the Extractive Industries Transparency Initiative, launched in 2002, and which currently encompasses 55 countries. Member countries commit to disclose information along the production value chain of oil, gas, and mining extraction, and to respect a common set of governing standards. We want to see if there is an effect of the EITI Rules signed by member countries on the effect of industrial mining on child mortality. 18 out of the 26 countries included in our sample signed the EITI[21], which includes 76% of our sample of

downstream up to the third sub-basin.

[17] Up to the third sub-basin

[18] Source: World Bank data.

[19] There are 604 mines in total in our main results' regressions.

[20] As sampling weights are not considered in the calculation, we do not give a number per country.

[21] Burkina Faso, Cote d'Ivoire, Democratic Republic of the Congo, Ethiopia, Ghana, Guinea, Liberia, Madagascar, Malawi, Mali, Niger, Nigeria, Senegal, Sierra Leone, Tanzania, Togo, Uganda, and Zambia are EITI members. Benin, Burundi, Kenya, Namibia, Rwanda, and Zimbabwe have not joined the EITI.

Table 21: Effects of industrial mining opening, across EITI membership.

Outcome	24-month mortality					
Sample	All			Rural		
	Not an EITI member	EITI member		Not an EITI member	EITI member	
	(1)	(2)	(3)	(4)	(5)	(6)
Downstream× Open	0.0179 [0.0210]	0.0259** [0.0127]	0.0133 [0.0190]	-0.0428 [0.0518]	-0.0221 [0.0416]	-0.0557 [0.0669]
Surveyed after joining EITI			0.0237 [0.0303]			-0.0259 [0.0315]
D× O× Surv. after joining EITI			0.0243 [0.0236]			0.0221 [0.0705]
Downstream	-0.0552*** [0.0140]	-0.00934 [0.00876]	-0.00446 [0.0107]	0.00399 [0.0500]	0.0356 [0.0381]	0.0118 [0.0641]
Open	-0.00576 [0.0246]	-0.00588 [0.0112]	-0.00656 [0.0163]	-0.0456 [0.0651]	-0.00345 [0.0268]	-0.0114 [0.0443]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes
N	8,434	26,810	26,810	2,251	8,685	8,685
R2	0.0373	0.0548	0.0548	0.0838	0.0677	0.0679
Outcome mean	0.0716	0.0920	0.0920	0.0604	0.0738	0.0738

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. The same sample and controls as Table 2 Column 2 apply.

children.

We estimate our main specification across the sample of countries that are members of the EITI or not. We find that our results hold even for countries who committed to improved governance of their extractive industries (Table 21), but which are also countries heavily relying on this activity in their national economy. We find no significant effect of our results when looking at whether surveys were conducted before or after their country signed the EITI Standards (triple interaction Downstream $\times$ Open $\times$ Surveyed after joining EITI in columns (3) and (6)).

13 Conclusion

This paper identifies a negative externality of industrial mining on local population living standards, as we show that industrial mining sites indirectly increase infant mortality in surrounding villages through the contamination of water resources. We match geocoded repeated-cross sectional household surveys to geocoded data on industrial mine openings obtained through intensive manual data collection. We propose a staggered DiD strategy and isolate the mechanism of water pollution by building the treatment and control groups using an upstream-downstream comparison. We compare the effects on the health of villages located upstream and downstream of a mine deposit, before and after its opening. We are the first, to the best of our knowledge, to take into account the topography of mining areas using an upstream-downstream comparison and to empirically quantify this effect at the scale of 604 mines in 26 countries of Sub-Saharan Africa over 1981-2020.

We find that the opening of industrial mines increases the 24-month mortality rate by 25% among villages located downstream compared to villages located upstream, and thus indirectly isolates the channel of water pollution. We find almost no effects on other child health outcomes, such as anthropometric measures, cough, fever, diarrhea, or anemia. We exploit the variation of the opening of a mine and show that our results are not driven by a change in women's fertility behavior, differential access to piped water, electricity, or health facilities. The heterogeneity across the breastfeeding practices shows that the results are mainly driven by children who were no longer breastfed after 6 months, which corroborates the mechanism of water pollution.

In an additional heterogeneity analysis, we show that our results are mainly driven by the pollution occurring during the time of mining activity. We find that the effects are even more harmful in rural areas, for open-pit and foreign-owned mines, and in places with a high density of mines. We also find that the effects increase with productivity intensity (proxied by international commodity prices). We run manyfold robustness checks and find that our results hold when controlling for in-migration, and when restricting to a balanced sample which deals with the issue of repeated cross-section surveys. Our results are also robust to the heterogeneous treatment effects estimator of de Chaisemartin and d'Haultfœuille [2020], to measurement error tests, and a battery of placebo tests such as spatial and temporal randomization inference tests.

References

Ahlerup, P., Baskaran, T., and Bigsten, A. (2020). Gold mining and education: A long-run resource curse in africa? *The Journal of Development Studies*, 56(9):1745–1762.

Aragón, F. and Rud, J. P. (2016). Polluting industries and agricultural productivity: Evidence from mining in ghana. *The Economic Journal*, 126(597):1980–2011.

Atkin, D. (2016). Endogenous skill acquisition and export manufacturing in mexico. *American Economic Review*, 106(8):2046–85.

Atlas des Conflits pour la Justice Environnementale (2022). Tentative de responsabilité sociale chez iamgold's mine, burkina faso. `https://ejatlas.org/conflict/gold-and-water-rush-in-burkina-fasos-essakane-mine/?translate=fr`. Accessed: 2022-08-11.

Bazillier, R. and Girard, V. (2020). The gold digger and the machine. evidence on the distributive effect of the artisanal and industrial gold rushes in burkina faso. *Journal of Development Economics*, 143:102411.

Benshaul-Tolonen, A. (2018). Local Industrial Shocks and Infant Mortality. *The Economic Journal*, 129(620):1561–1592.

Berman, N., Couttenier, M., Rohner, D., and Thoenig, M. (2017). This mine is mine! how minerals fuel conflicts in africa. *American Economic Review*, 107(6):1564–1610.

Bialetti, A., Page, L., Pande, R., Rowe, K., and Sudarshan, A. (2018). Lease splitting and dirty entrants: The unintended deforestation consequences of india's environmental clearance process reform. *PEDL Research Papers*.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *Revisions Requested from Review of Economic studies*.

Briffa, J., Sinagra, E., and Renald, B. (2020). Heavy metal pollution in the environment and their toxicological effects on humans. *Heliyon*.

Callaway, B. and Sant'Anna, P. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230. Themed Issue: Treatment Effect 1.

Cobbina, Kumi, and Myilla (2013). Small scale gold mining and heavy metal pollution: Assessment of drinking water sources in datuku in the talensi-nabdam district. *International Journal of Scientific and Technology Research*, 2:96–100.

Coelho, P. and Texeira, J. P. (2011). Mining activities: Health impacts. *Elsevier*.

Colella, F., Lalive, R., Sakalli, S. O., and Thoenig, M. (2019). Inference with arbitrary clustering. IZA Discussion Paper n. 12584.

Corno, L. and de Walque, D. (2012). Mines, migration and hiv/aids in southern africa. *Journal of African Economies*, 21.

Cossa, H., Dietler, D., Macete, E., Munguambe, K., Winkler, M., and Fink, G. (2022). Assessing the effects of mining projects on child health in sub-saharan africa: a multi-country analysis. *Globalization and Health*.

Cust, J. and Poelhekke, S. (2015). The local economic impacts of natural resource extraction. OxCarre Working Papers 156, Oxford Centre for the Analysis of Resource Rich Economies, University of Oxford.

de Chaisemartin, C. and d'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Dietler, D., Farnham, A., Loss, G., Fink, G., and Winkler, M. S. (2021). Impact of mining projects on water and sanitation infrastructures and associated child health outcomes: a multi-country analysis of demographic and health surveys (dhs) in sub-saharan africa. *Global Health*.

Dike, I., Onwurah, C., Uzodinma, U., and Onwurah, I. (2020). Evaluation of pb concentrations in selected vegetables and portable drinking water, and intelligent quotients of school children in ishiagu-a pb mining community: health risk assessment using predictive modeling. *Environ Monit Assess*, (192(2):126).

Do, Q.-T., Joshi, S., and Stolper, S. (2018). Can environmental policy reduce infant mortality? evidence from the ganga pollution cases. *Journal of Development Economics*, 133:306–325.

Drechsel, F., Engels, B., and Schäfer, M. (2018). Les mines nous rendent pauvres : L'exploitation minière industrielle au burkina faso. Technical report, GLOCON.

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

Edwards, D., Sloan, S., Weng, L., Dirks, P., Sayer, J., and Laurance, W. (2013). Mining and the african environment. *Conservation Letters*, 7.

El-Kady, A. A. and Abdel-Wahhab, M. A. (2018). Occurrence of trace metals in foodstuffs and their health impact. *Trends in Food Science and Technology*, 75:36–45.

Fängström, B., Moore, S., Nermell, B., Kuenstl, L., Goessler, W., Grandér, M., Kabir, I., Palm, B., Arifeen, S., and Vahter, M. (2008). Breast-feeding protects against arsenic exposure in bangladeshi infants. *Environmental health perspectives,*.

Garg, T., Hamilton, S., Hochard, J., Kresch, E. P., and Talbot, J. (2018). (not so) gently down the stream: River pollution and health in indonesia. *Journal of Environmental Economics and Management*, 92:35–53.

Girard, V., Molina-Millan, T., and Vic, G. (2022). Artisanal mining in africa.

Global Energy Monitor Wiki (2021). Heavy metals and coal. `https://www.gem.wiki/Heavy_metals_and_coal#cite_note-Toppin-3`. Accessed: 2022-04-28.

Glöser, S., Hartwig, J., Wheat, D., and Faulstich, M. (2017). The cobweb theorem and delays in adjusting supply in metals' markets. *System Dynamics Review*, 32.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Working Paper 25018, National Bureau of Economic Research.

He, A., Li, X., Ai, Y., Li, X., Li, X., Zhang, Y., Gao, Y., Liu, B., Zhang, X., Zhang, M., Peng, L., Zhou, M., and Yu, H. (2020). Potentially toxic metals and the risk to children's health in a coal mining city: An investigation of soil and dust levels, bioaccessibility and blood lead levels. *Environ Int*, (141:105788).

Khan, T., Nguyen, T., and Ohnsorge, Franziska an Schodde, R. (2016). From commodity discovery to production. *Policy Research Working Paper*, World Bank, Washington, DC(7823).

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

Macklin, M., Thomas, C., Mudbhatkal, A., Brewer, P., Hudson-Edwards, K., Lewin, J., Scussolini, P., Eilander, D., Lechner, A., Owen, J., Bird, G., Kemp, D., and Mangalaa, K. (2023). Impacts of metal mining on river systems: a global assessment. *Science*, 381(6664):1345–1350.

Madilonga, R., Edokpayi, J., Volenzo, E., Durowoju, O., and Odiyo, J. (2021). Water quality assessment and evaluation of human health risk in mutangwi river, limpopo province, south africa. *Int J Environ Res Public Health*, (18(13):6765).

Malpede, M. (2021). The dark side of batteries: Education and cobalt mining in the democratic republic of the congo. *SSRN*.

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*, 139:28 – 49.

McQuilken, J. and Perks, R. (2020). 2020 state of the artisanal and smallscale mining sector. washington, d.c.: World bank. Technical report, World Bank.

Obasi, N., Obasi, S., Nweze, E., Amadi, S., Aloke, C., and Aloh, G. (2020). Metal pollution and human health risk assessment of soils and edible plants in farmlands around enyigba lead-zinc mining site, ebonyi state, nigeria. *Environ Monit Assess*, (192(5):292).

Oje, O., Uzoegwu, P., Onwurah, I., and Nwodo, U. (2010). Environmental pollution levels of lead and zinc in ishiagu and uburu communities of ebonyi state, nigeria. *Bull Environ Contam Toxicol*, (85(3):313-7).

Porgo, M. and Gokyay, O. (2016). Environmental impacts of gold mininig in essakane site of burkina faso. *Human and Ecological Risk Assessment: An International Journal*, 23.

Shubhayu, S., Pattanayak, Sills, E., and Singha, A. (2011). Under-mining health: Environmental justice and mining in india. *Health Place*, 17(1):140–148. Health Geographies of Voluntarism.

Strobl, E. and Strobl, R. (2011). The distributional impact of large dams: Evidence from cropland productivity in africa. *Journal of Development Economics*, 96(2):432–450.

Taylor, C., Schulz, K., Doebrich, J., Orris, G., Denning, P., and Kirschbaum, M. (2009). Geology and nonfuel mineral deposits of africa and the middle east. *U.S. Geological Survey*.

van der Ploeg, F. (2011). Natural resources: Curse or blessing? *Journal of Economic Literature*, 49(2):366–420.

VanDerSlice, J., Popkin, B., and Briscoe, J. (1994). Drinking-water quality, sanitation, and breast-feeding: their interactive effects on infant health. *Bulletin of the World Health Organization, 72(4), 589–601.*

Venables, A. (2016). Using natural resources for development: Why has it proven so difficult? *Journal of Economic Perspectives*, 30(1):161–84.

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

A Appendix

B Descriptive statistics

B.1 Data

Table 22 displays the number and years of DHS waves, and the total number of DHS clusters and children under 5 years old for each country . Overall, the DHS sample gathers 36 countries overall in Africa, from 1986 to 2018. In our main empirical analysis, we decided to only keep DHS countries that had at least two survey rounds in order to have comparable temporal variation across countries. Our final sample accounts for the following countries (cf. Table 23): Tanzania, Burkina-Faso, Ghana, Zimbabwe, Mali, Democratic Republic of Congo, Guinea, Namibia, Madagascar, Cote d'Ivoire, Sierra Leone, Liberia, Nigeria, Senegal, Ethiopia, Uganda, Botswana, Malawi, Cameroon, Morocco, Niger, Kenya, Mauritania, Rwanda, Burundi, Lesotho, Togo, Eswatini, Algeria, Benin, Eritrea, Republic of the Congo, Guinea-Bissau, Somalia, Sudan, Tunisia, Djibouti, Equatorial Guinea (by order of importance in terms of mining activity according to Figure 23).

Table 22: DHS surveys overall across countries

Countries	Survey Years	Number of clusters	Number of children under 5
AO	2015	625	14,177
BF	1993, 1999, 2003, 2010	1,413	36,744
BJ	1996, 2001, 2012, 2017	1,752	31,884
BU	2010, 2016	930	20,824
CD	2007, 2013	836	27,307
CF	1994	230	2,639
CI	1994, 1998, 2012	674	12,227
CM	1991, 2004, 2011, 2018	1,619	31,279
EG	1992, 1995, 2000, 2003, 2005, 2008, 2014	7,741	75,394
ET	2000, 2005, 2010, 2016	2,313	42,173
GA	2012	334	5,911
GH	1993, 1998, 2003, 2008, 2014	2,037	17,931
GN	1999, 2005, 2012, 2018	1,289	26,588
KE	2003, 2008, 2014	2,391	32,235
KM	2012	252	3,134
LB	1986, 2007, 2013	776	16,224
LS	2004, 2009, 2014	1,199	10,269
MA	2003	480	6,030
MD	1997, 2008	860	15,932
ML	1996, 2001, 2006, 2012, 2018	1,867	52,996
MW	2000, 2004, 2010, 2015	2,655	56,688
MZ	2011	610	10,950
NG	1990, 2003, 2008, 2013, 2018	3,830	106,848
NI	1992, 1998	503	11,332
NM	2000, 2006, 2013	1,290	13,630
RW	2005, 2008, 2010, 2014	1176	21,927
SL	2008, 2013	787	17,483
SN	1993, 1997, 2005, 2010, 2012, 2014, 2015, 2016, 2017	2,572	73,084
SZ	2006	274	2,706
TD	2014	624	18,441
TG	1988, 1998, 2013	768	13,869
TZ	1999, 2010, 2015	1,259	20,520
UG	2000, 2006, 2011, 2016	1,765	37,603
ZA	2017	671	3,397
ZM	2007, 2013, 2018	1,585	29,105
ZW	1999, 2005, 2010, 2015	1,431	19,847

Notes: This table gives the sample size of children under five years old overall DHS surveys.

Table 23: DHS surveys in regression sample across countries

Countries	Survey Years	Number of clusters	Number of children under 5
BF	1993, 1999, 2003, 2010	694	23,846
BJ	2001, 2012, 2017	62	1,911
BU	2010, 2016	317	8,280
CD	2007, 2013	82	5,092
CI	1994, 1998, 2012	196	4,838
CM	1991, 2004, 2011, 2018	90	2,513
ET	2000, 2005, 2010, 2016	100	2,956
GH	1993, 1998, 2003, 2008, 2014	1,217	12,074
GN	1999, 2005, 2012, 2018	360	11,775
KE	2003, 2008, 2014	233	4,130
LB	1986, 2007, 2013	190	7,537
LS	2004, 2009, 2014	336	2,810
MD	1997, 2008	131	3,301
ML	1996, 2001, 2006, 2012, 2018	570	19,147
MW	2000, 2004, 2010, 2015	207	6,651
NG	1990, 2003, 2008, 2013, 2018	105	3,993
NI	1992, 1998	40	1,105
NM	2000, 2006, 2013	138	2,175
RW	2005, 2008, 2010, 2014	713	14,615
SL	2008, 2013	377	13,717
SN	1993, 1997, 2005, 2010, 2012, 2014, 2015, 2016, 2017	363	10,111
TG	1988, 1998, 2013	104	2,187
TZ	1999, 2010, 2015	325	6,866
UG	2000, 2006, 2011, 2016	305	9,031
ZM	2007, 2013, 2018	364	10,966
ZW	1999, 2005, 2010, 2015	468	8,307

Notes: This table gives the sample size of children under five years old that are in our main analysis, meaning within 100 km of an industrial mine.

Tables 24, 25, and 26 present the descriptive statistics for all of our outcome and control variables for the sample of all individuals living within 45 km of an industrial mine regardless of their topographical position, in the 26 countries of Sub-Saharan Africa with at least 2 waves of DHS and for heavy metals and coal mines. These descriptive figures are important to show that our analysis does not suffer from selection biases across the samples we use for our different regressions.

Table 24: Descriptive statistics of children's outcomes

	Mean	SD	Med	Min	Max	N
Mortality rates						
12-month mortality	0.064	.244	0	0	1	189,181
24-month mortality	0.083	.275	0	0	1	139,683
Control variables						
Birth order number	3.655	2.421	3	1	18	240,431
Male	0.508	0.500	1	0	1	240,431
Anthropometric measures						
Stunting	0.319	0.466	0	0	1	137,834
Underweight	0.234	0.423	0	0	1	136,043
Wasting	0.077	0.267	0	0	1	138,222
Weight and size at birth						
Less than 2.5 kg	0.164	0.370	0	0	1	117,651
Small or very small size	0.161	0.367	0	0	1	226,796
Measured anemia level						
Any anemia	0.633	0.482	1	0	1	67,567
Illness in the last 2 weeks						
Diarrhea	0.168	0.374	0	0	1	216,097
Cough	0.260	0.439	0	0	1	214,940
Fever	0.265	0.441	0	0	1	214,913
Nutrition						
Given plain water	0.187	0.390	0	0	1	122,915
Ever breastfed	0.980	0.140	1	0	1	223,039
Months breastfed	14.788	8.917	15	0	59	156,011
Health access						
No prenatal care	0.101	0.301	0	0	1	169,268
Ever vaccinated	0.788	0.409	1	0	1	82,082
Characteristic of paired mine						
Domestic mine	0.177	0.381	0	0	1	240,431
Open-pit mine	0.676	0.468	1	0	1	103,667

Notes: We present the mortality rates at n months, conditionnally on having reached n months, for the whole sample of children living within 45 km of an industrial mine and regardless of their topographic position. The sample is restricted to the 26 Sub-Saharan countries with at least two waves of DHS and to heavy metals and coal mines.

Table 25: Descriptive statistics of mothers' outcomes

	Mean	SD	Med	Min	Max	N
Mother's characteristics						
Mother's age	28.918	6.979	28	15	49	240,431
Years of education	3.985	4.226	3	0	22	240,332
Urban	0.287	0.452	0	0	1	236,966
Migrant	0.594	0.491	1	0	1	161,292
Access to sanitation and health facilities						
Piped water as main drinking water source	0.261	0.439	0	0	1	240,431
Has flushed toilet	0.086	0.280	0	0	1	239,773
Has electricity	0.218	0.413	0	0	1	236,692
Visited health facility in the last 12 months	0.623	0.485	1	0	1	218,053

Notes: The sample is restricted to all mothers of 0-5 years old children living within 45 km of an industrial mine in the 26 Sub-Saharan countries with at least two waves of DHS and to heavy metals and coal mines.

Table 26: Descriptive statistics of women's outcomes

	Mean	SD	Med	Min	Max	N
Fertility behavior and health						
Ever had a child	0.736	0.441	1	0	1	330,889
Total lifetime fertility	2.890	2.785	2	0	18	330,889
Currently pregnant	0.091	0.288	0	0	1	330,744
Ever had a miscarriage	0.127	0.333	0	0	1	296,235
Any anemia	0.378	0.485	0	0	1	115,481
Placebo disease						
Any STD	0.049	0.216	0	0	1	276,924
Heard of tuberculosis	0.935	0.246	1	0	1	88,438

Notes: The sample is restricted to all women aged 15-49 living within 45 km of an industrial mine in the 26 Sub-Saharan countries with at least two waves of DHS and to heavy metals and coal mines.

B.2 Handwork

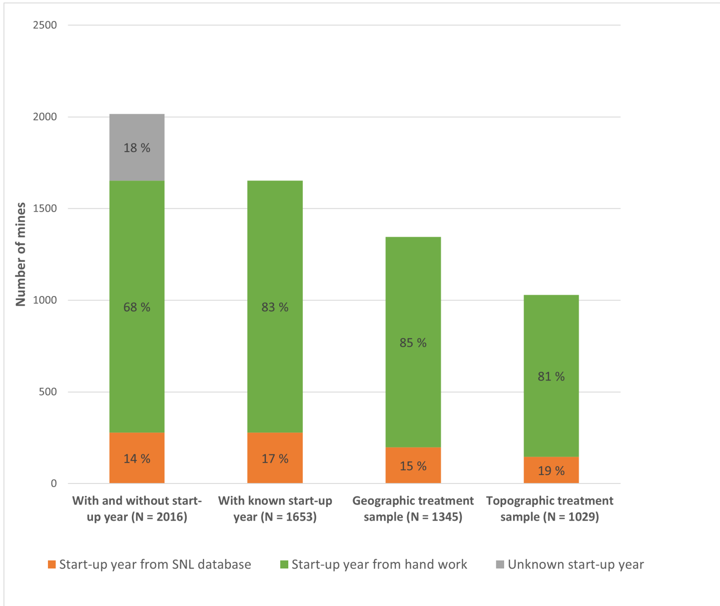
Out of the 3,815 industrial mines recorded by the SNL database in Africa, 2,016 were located within 100 km of a DHS cluster (with at least 2 waves of DHS). Within the database, 278 had information on the opening and closure years. For the 1,738 remaining mines we searched for their opening year [22]. The handwork consisted of reading the reports (comments and work history) available in the database and browsing through the aerial images available on the SNL platform, which provided the exact GPS coordinates and main location labels. This information was corroborated through online research (press releases, mining companies' websites, specialized websites on global mining activities, etc.) as well as Google maps and Google timelapse satellite imagery. A mine opening corresponds to the beginning of production.

The exact startup year could not be determined for 18% of our sample (Figure 20 Bar (1)), and these mines were dropped from our regressions. In total we hand-checked 83% of the mines located within 100 km of a DHS cluster, and for which we knew their startup year (Figure 20 Bar (2)). Among the sample of mines with a startup year, 83.2% opened after 1981 (first year of birth within the DHS child surveys). For each of the following graphs, we studied the entire sample of 2,016 mines and plotted the percentage of mines that were hand-checked, and the percentage of mines that ended up having a startup year and are thus included in our study. We conduct this analysis on all available mines within 100 km of a DHS cluster to be transparent on the creation of our sample as compared to the original one.

The distribution across each mining site's primary commodity of production can be found in Figure 22. Half of our sample consists of gold mining sites. Figure 23 represents the distribution across country of location. Ownership information is available for 65% of our sample and the main owners are from the USA, UK, Canada, Australia, and China (Figure 25).

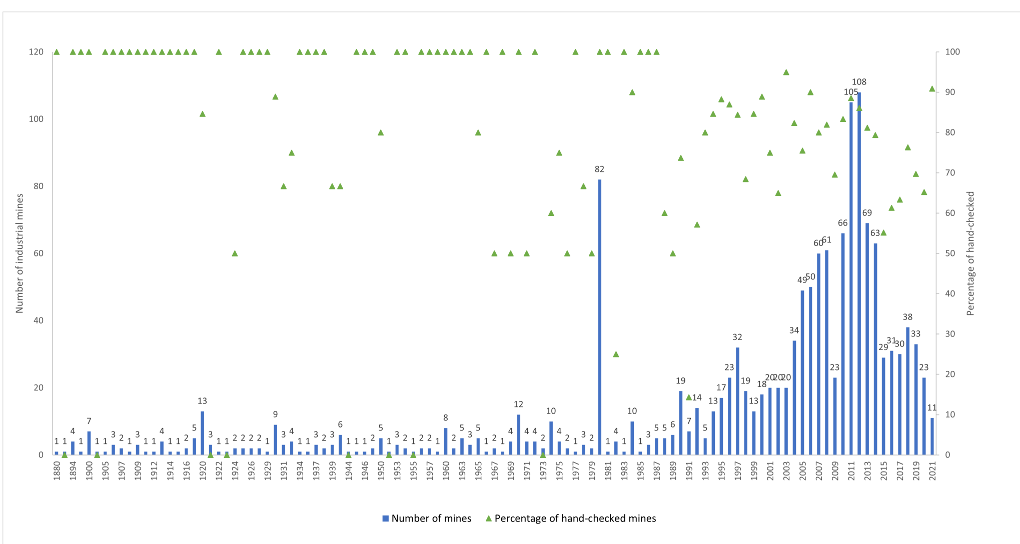
[22] We also looked at their closure date as well as their current activity status, i.e., whether the mining site looked active or inactive. However, this information was harder to retrieve and we finally focused on the opening date.

Figure 20: Description of hand work and industrial mines samples



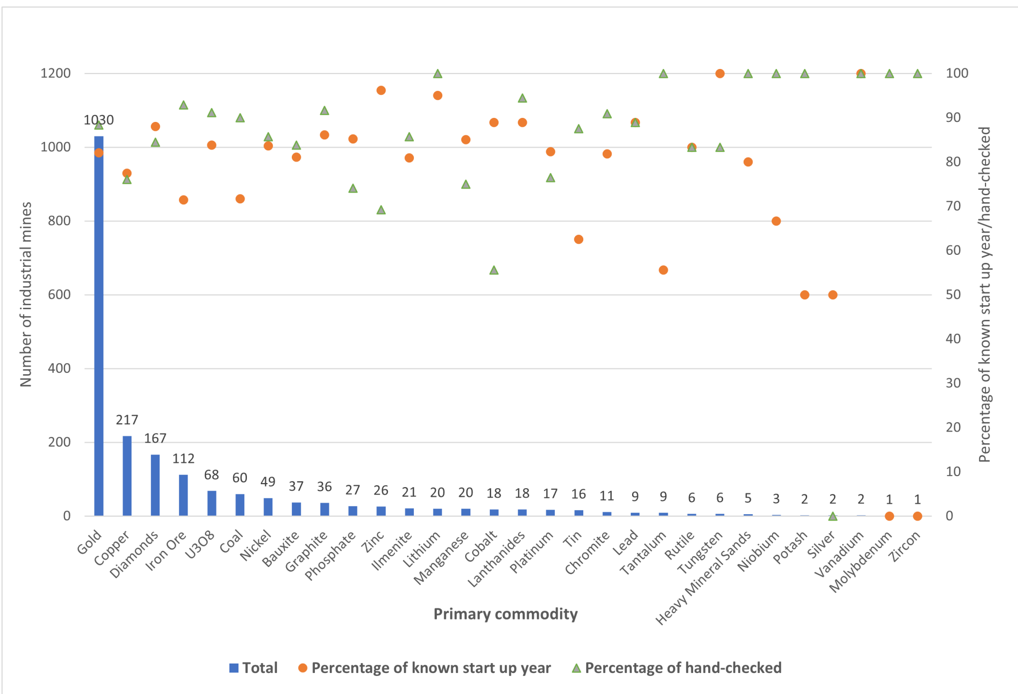
Notes: This Figure gives the number and percentage of mines for which we have manually retrieved the opening year. Bars (1) and (2) correspond to all the mines located within 100 km of a DHS cluster; Bar (3) to the sample associated with the replication of Benshaul-Tolonen [2018] in Section G; and Bar (4) corresponds to the main analysis, i.e., to mines that have at least one DHS cluster upstream within 100 km, and one DHS cluster downstream within the three closest sub-basins (cf. pairing strategy Section4.1.1) .
Sources: Authors' elaboration on DHS and SNL data.

Figure 21: Mines and percentage of hand-checked mines across startup years



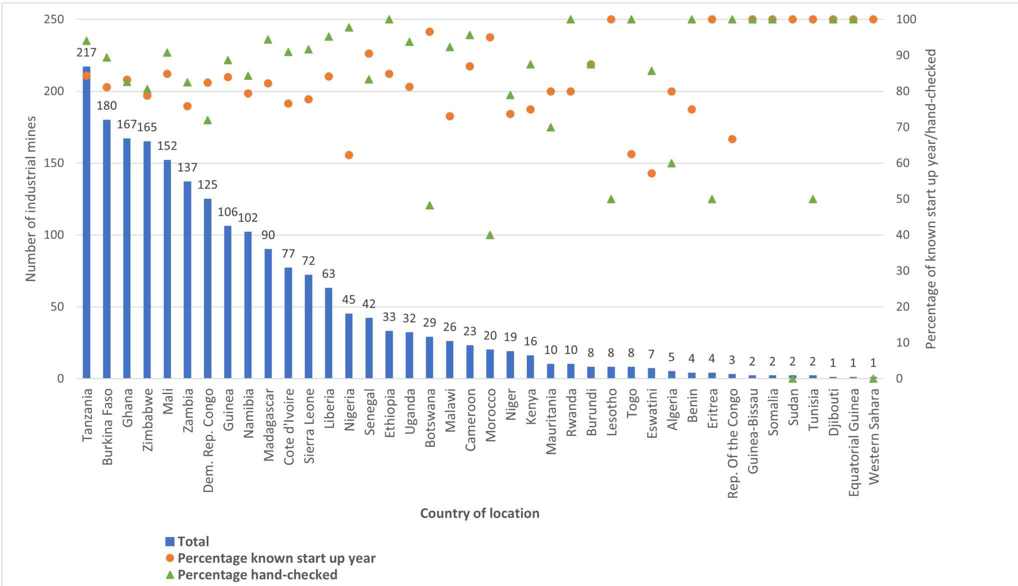
Notes: This graph displays the number of mines that opened during a specific year and the percentage of mines that were checked by hand for the 2,016 mines located within 100 km of a DHS cluster.

Figure 22: Mines and percentage of hand-checked mines across primary commodities



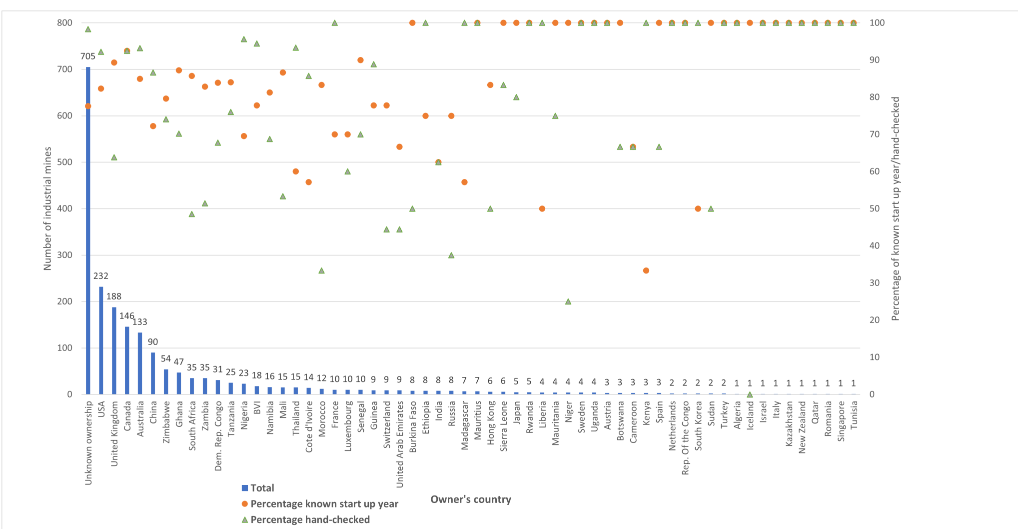
Notes: This graph gives the number of mines for each primary commodity, and the percentage of mines that were checked by hand, for the 2,016 mines located within 100 km of a DHS cluster.

Figure 23: Mines and percentage of hand-checked mines across country of location



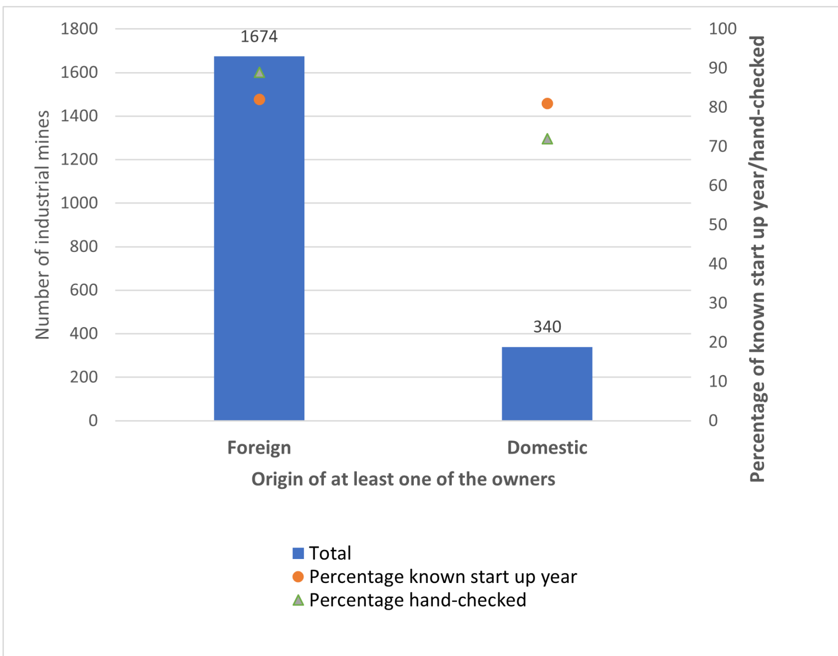
Notes: This graph gives the number of mines for each country of location, and the percentage of mines that were checked by hand, for the 2,016 mines located within 100 km of a DHS cluster.

Figure 24: Mines and percentage of hand-checked mines across owner's country



Notes: This graph gives the number of mines by the owning company's registration country, and the percentage of mines that were checked by hand, for the 2,016 mines located within 100 km of a DHS cluster.

Figure 25: Mines across foreign and domestic ownership

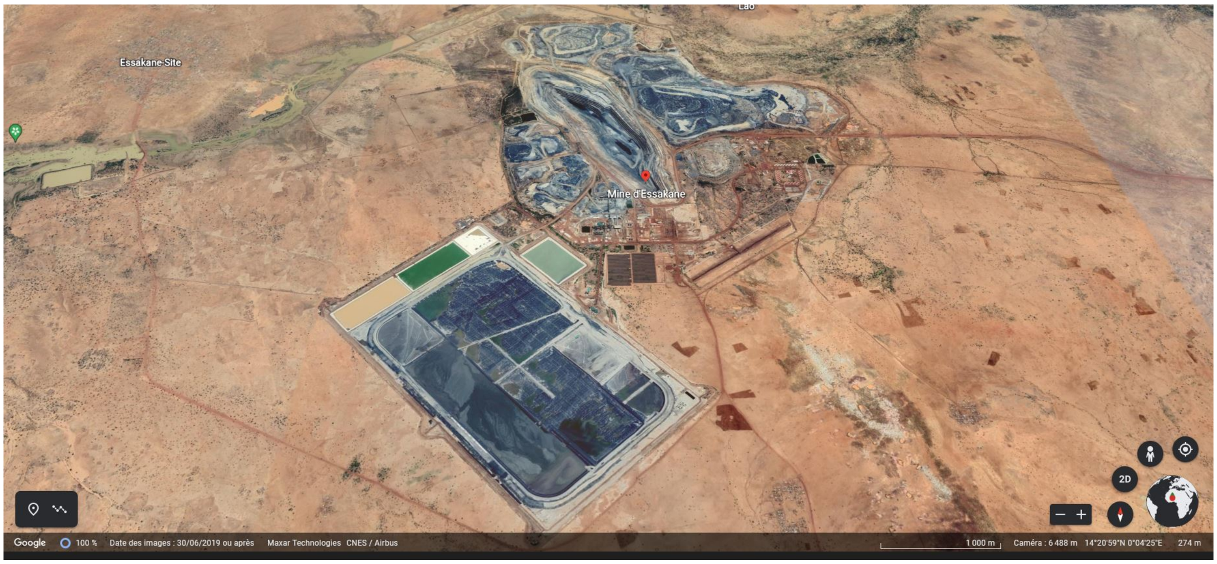


Notes: This graph gives the number of mines across domestic and foreign ownership, and the percentage of mines that were checked by hand, for the 2,016 mines located within 100 km of a DHS cluster.

C Context

C.1 Case study: the Essakane mine

Figure 26: Satellite image of the Essakane Mine in 2019



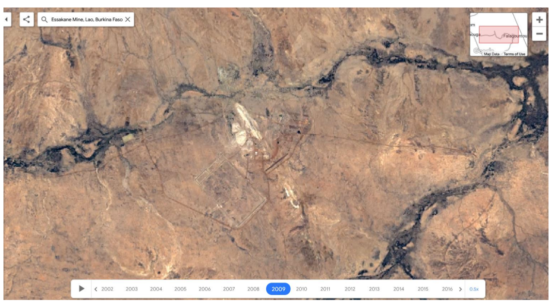
Notes: Satellite image of the Essakane Mine in 2019. Retention dams can be seen.
Sources: Google Earth.

Figure 26 shows the satellite image of a mine from our sample, the Essakane mine, in 2019, and Figure 27 shows the different stages of expansion and construction of the mine. Essakane is the second largest mine in Burkina-Faso and the most productive gold mine still in activity in the country. It is an open-pit gold mine that extends over a 100 km^2 area. It is located in the North-East of Burkina-Faso in the Oudalan province, near the Nigerian and Malian borders, and is hydrologically found in the sub-basin of the Gorouol and Feildegasse rivers. It is exploited by the Burkinabé society Iamgold Essakane and belongs to the Canadian investor IamgoldInc (International African Mining Gold Corporation), who obtained the project in 2007. The installation of the mine in 2009 forcibly displaced five villages, and 16,000 people who had no choice; the promised compensation to the communities for the displacement cost, loss of pastures and common forests has not been fulfilled ([Atlas des Conflits pour la Justice Environnementale, 2022]). Mining at Essakane has been shown to have negative impacts on the environment and the health of the local population, both indirectly and accidentally.

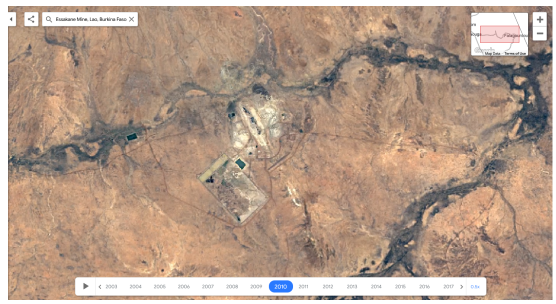
In November 2015, Drechsel et al. [2018] conducted qualitative interviews among the inhabitants of local communities of six active mining zones in Burkina-Faso, including the Essakane zone. If the local population admits to the benefits of the construction of a primary school and educational establishment, of a health center, of roads and electricity, the interviewed people do not find it sufficient to outweigh the negative aspects. The mine does not offer formal employment to the local population who are not educated enough to undertake the required skilled work. On the contrary, due to the loss of agricultural land and the prohibition of practicing gold panning, the local population fell to unemployment and poverty. As a farmer from the Essakane area explains: *"Before the mine arrival, we had better lives, we had animals, we were rich"* [Drechsel et al., 2018]. In 2010, the tailing storage facility of the mine collapsed, causing the death of the surrounding livestock which was poisoned by chemicals used in the mine, and creating tension between the local population and the mine operator. In 2011, a truck carrying two containers of cyanide fell into the water source of the Djibo dam and led to the death of all the fish in the dam. Tensions regarding water scarcity exist as well, as the mine is water-intensive and reinforces the vulnerability of the local population to droughts. Even though the regional government prohibited the mine from using the village water, the national government overruled the decision, and the operator directly uses the water originally intended for the village. This led to protests around the mine by the local population in 2011, but with no success. Finally, miners had a major impact on soil degradation, due to the construction of mining infrastructures, the multiplication of satellite pits andthe abandoned sterile holes devoid of gold [Porgo and Gokyay, 2016].

Environmental pollution has caused a deterioration to the living conditions around the Essakane mine. Porgo and Gokyay [2016] use water sampling and digital calipers to capture water pollution and particle measurement in the Essakane zone and survey the surrounding population to understand the related health conditions. They find high levels of particles at the Essakane site center due to transportation and mining activity, such as the work of perforation, blasting, loading, transportation of ore, crushing, grinding, and energy production based on hydrocarbons. This air pollution mainly concerns mine workers, who develop acute respiratory infections (ARI), incurable lung diseases caused by prolonged and severe inhalation of fine particles. The drilled well-water samples display abnormally high concentrations of arsenic (higher than the WHO standard), which come from the intensive use of acids (low pH) and the release of trace metals. The surrounding population presents diarrhea (13%) and affections of the skin and wounds (11%), reported to be caused by lack of hygiene, and the use of drugs and chemicals. The main important health impact associated with the mine is the increase of malaria (20%), as stagnant water from mining dams attracts infected mosquitoes.

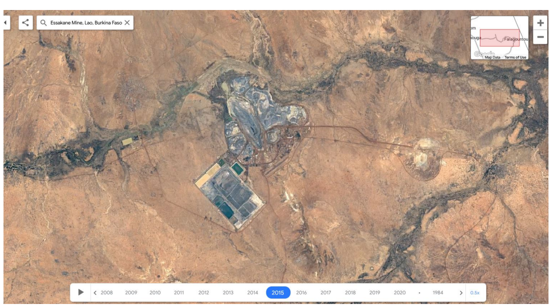
Figure 27: Expansion of the Essakane Mine, 2009-2019.



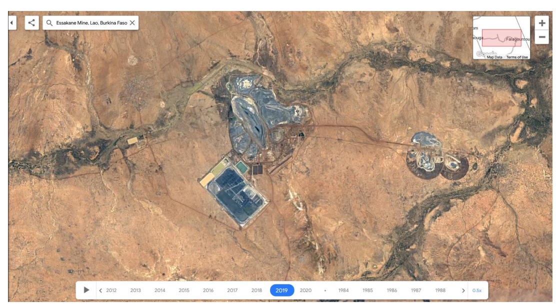
(a) 2009



(b) 2010



(c) 2015



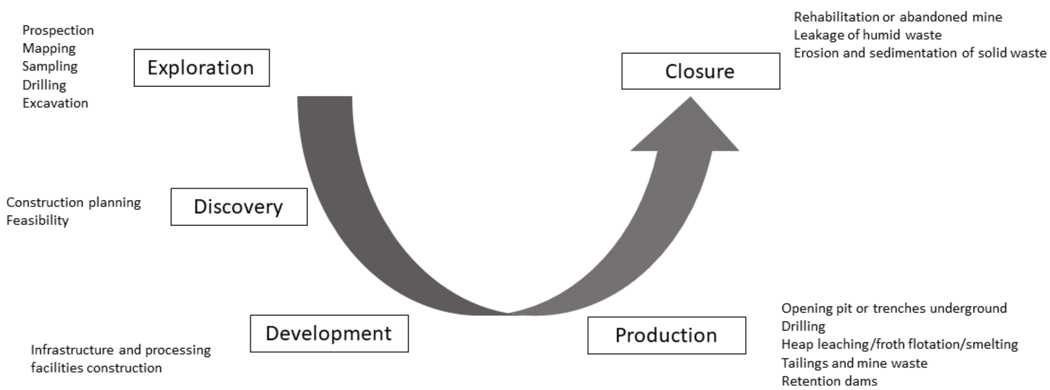
(d) 2019

Notes: The four satellite images represent the expansion of the Essakane plant, in Burkina Faso. Retention dams can be seen.
Sources: Google Earth Engine Timelapse.

C.2 Mine life cycles and types

Figure 28 shows the main stages of an industrial mining project, from the exploration phase to the start of production and the closure of the mine. It is hard to give the average length of each phase, but the average mine life is 16 years from the start of production to its closure (Figure 11). Figure 29 gives the time evolution of international prices for all commodities used in our main sample.

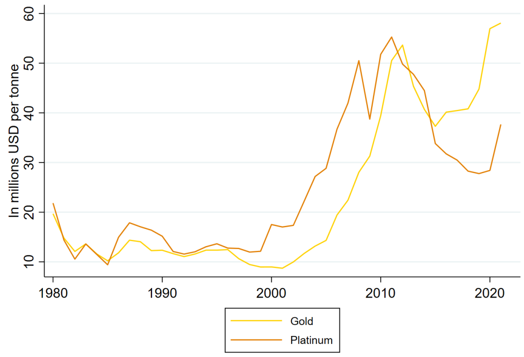
Figure 28: Life cycle of industrial mine



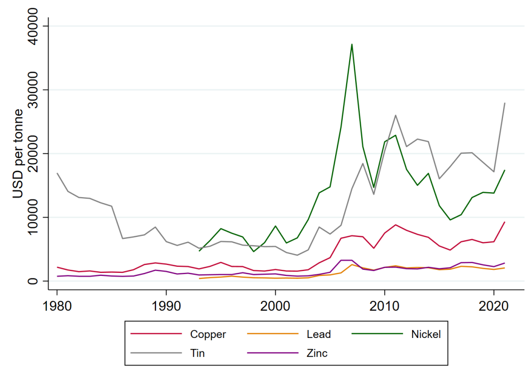
Notes: The figure schematizes the main stages of an industrial mining project .
Sources: Authors' elaboration, largely inspired by Coelho, Teixeira and Goncalves (2011)

Table 27 gives the chemical properties of each metal, including their main chemical compounds (column (1)), their density (column (2)), and displays their share in the main estimation sample (in terms of the number of mines, column (3), and Total Individual Sample, (column (4)). Heavy metals are defined according to their density as being greater than $5gcm^{-3}$ [Briffa et al., 2020]. If small amounts of heavy metals can be mandatory, a high and abnormal concentration of heavy metals may cause health issues due to chronic toxicity. Heavy metals released during mining activity are toxic elements that degrade the environment and human biology. This is also the case for heavy metals released during the mining and burning of coal, which is linked to toxic heavy metals such as lead, mercury, arsenic, and nickel [Global Energy Monitor Wiki, 2021]. The main regression analysis includes heavy metals and coal mines to capture the negative externalities linked to the

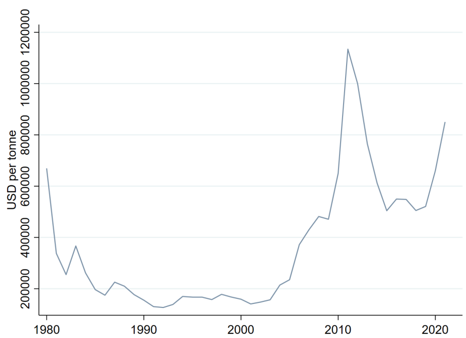
Figure 29: Time evolution of international commodity prices



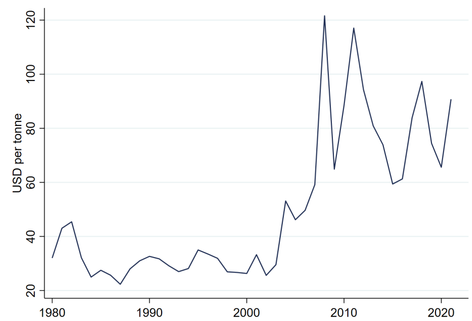
(a) Gold and Platinum



(b) Other metals



(c) Silver



(d) Coal

Notes: The Figure plots the evolution of metal prices from 1980 to 2020.
Sources: Authors' elaboration from SNL data and World Bank Pink Sheet data.

most toxic mines.

Table 27: Metals, chemical properties and sample distribution

Metals	Main chemical coumpounds (1)	density (gcm^{-3}) (2)	Nb. Mines (3)	Total Individual Sample (%) (4)
Heavy Metals				
Gold	Gold	19.3	581	41.88
Copper	Copper	8.96	89	5.03
Iron ore	Iron	7.87	54	8.72
U308	Uranium	8.39	36	1.60
Nickel	Nickel	8.9	25	5.06
Platinum	Platinum	21.45	21	0.43
Zinc	Zinc	7.14	19	2.46
Chromite	Iron Chromium	[4.5,5.09]	16	0.57
Ilmenite	titanium	4.6	14	3.67
Lanthanides	Lanthane(57) Lutecium(71)	[6.1,9.8]	13	1.95
Manganese	Manganese	7.21	12	0.62
Tin	Tin	[5.7;7.26]	10	4.87
Cobalt	Cobalt	8.9	7	0.56
Tungsten	Tugsten	19.25	6	1.06
Tantalum	Tantalum	16.69	5	0.15
Vanadium	Vanadium	6.12	4	0.04
Niobium	Niobium	8.57	3	0.39
Heavy Mineral Sands	Zirconium Titanium Tungsten Thorum	[4.5,17.6]	3	0.16
Silver	Silver	10.49	1	0.00
Lead	Lead	11.29	1	0.06
Non-Heavy Metals				
Diamonds	Carbon	3.5	115	11.73
Coal	Carbon Mercury? Arsenic?	1.35	55	2.19
Bauxite	Aluminium	2.79	23	1.94
Graphite	Carbon	2.26	21	0.82
Phosphate	Phosphate	1.83	14	2.78
Lithium	Lithium	0.53	14	0.80
Rutile	titanium	4.23	2	0.29
Potash(Salt)	Potassium	0.89	1	0.17

Notes: This table gives for each metals the main chemical coumpounds (Column (1)) and their density (Column (2)). Columns (3) gives the number of mines within 100 km of a DHS cluster for which the metal is the main primary commodity, and Column (4) the percentage of children under 5 associated to these mines.

D Empirical Strategy

D.1 Descriptive Statistics

Table 28 gives the balance table for some household and mother characteristics. This is a descriptive table that accounts neither for controls nor fixed effects. Table 29 shows the effect of being downstream of a mine for the same variables. We observe no statistical difference in terms of access to piped water, and electricity, age, and years of education of the mother. However, Table 29 shows that the proportion of urban households increases by 13 p.p once a mine opened in downstream areas, compared to upstream areas. The proportion of mothers that are migrants also increases by 8 p.p. These results suggest that the in-migrants coming after the mine opening, seeking jobs for instance, settle down in downstream areas that become more urban. It suggests that the miner villages are located downstream of the mine. This underscores the need to control for migration and verify that this is not driving our results (cf. Section 6.3).

In Figure 30 we plot the distribution of mines opened within 100 km upstream or within the 3 closest sub-basins downstream during a child's birth year, in order to see which countries gather the highest number of industrial mining activity in the vicinity of surveyed households over 1986-2018. Ghana, Zimbabwe, Tanzania, Zambia, Guinea, and Sierra Leone have the highest density of open mines near DHS clusters, while Benin, Burundi, Cameroon, Lesotho and Niger have the lowest number of open mining sites. This figure also represents the variation in the number of mines that opened between the first and last year of surveys for each country. We can thus capture the evolving context of change in industrial mining activity over our period of interest. Ghana, Tanzania, Guinea, Mali, and Burkina-Faso witnessed the highest number of mine openings between 1986 and 2018.

Figures 31, 32, and 33 give the spatial variation of the infant mortality outcomes and the mine openings for the sample restricted to our main analysis.

Table 28: Balance Table - Double Difference with Topographic Treatment - Descriptive Statistics

	Before Mine Opening					After Mine Opening					Within Up.	Within Dwn.	Within
	Upstream		Downstream		Diff	Upstream		Downstream.		Diff			
	N	Mean /(SD)	N	Mean /(SD)	(4-2) /(p.v)	N	Mean /(SD)	N	Mean /(SD)	(9-7) /(p.v)	(7-2) /(p.v)	(9-4) /(p.v)	(12-11) /(p.v)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Household Charateristics													
% Urban Household													
All	29,399	0.33	9,835	0.175	-0.155	16,285	0.385	6,174	0.291	-0.094	0.055	0.116	0.061
		(0.47)		(0.38)	(0)		(0.487)		(0.454)	(0)	(0)	(0)	(0)
Mines	244		237			190		193					
Has piped water													
All	29,399	0.307	9,835	0.193	-0.115	16,285	0.365	6,174	0.27	-0.095	0.057	0.077	0.019
		(0.461)		(0.395)	(0)		(0.481)		(0.444)	(0)	(0)	(0)	(0.002)
Has electricity													
All	29,399	0.211	9,835	0.14	-0.071	16,285	0.356	6,174	0.226	-0.13	0.145	0.086	-0.059
		(0.408)		(0.347)	(0)		(0.479)		(0.418)	(0)	(0)	(0)	(0)
Mother Charateristics													
Age													
All	29,399	29.106	9,835	29.187	0.081	16,285	28.779	6,174	28.818	0.039	-0.328	-0.369	-0.042
		(7.065)		(7.039)	(0.325)		(6.847)		(6.986)	(0.707)	(0)	(0.001)	(0.764)
Years of Education													
All	29,399	2.406	9,835	2.91	0.504	16,285	4.297	6,174	4.851	0.554	1.891	1.941	0.05
		(3.6)		(3.741)	(0)		(4.417)		(4.2)	(0)	(0)	(0)	(0.001)
% Migrant													
All	18,509	0.615	6,593	0.578	-0.037	9,773	0.597	3,962	0.589	-0.007	-0.019	0.011	0.029
		(0.487)		(0.494)	(0)		(0.491)		(0.492)	(0.421)	(0.002)	(0.278)	(0.094)

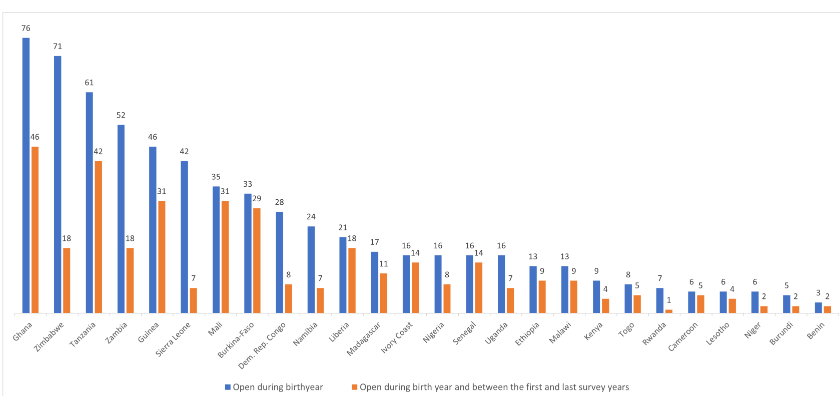
Notes: Standard errors and p-values in parentheses.

Table 29: Average effects of mine opening on control variables

	Household's characteristics			Mother characteristics		
	% urban households	Has piped water	Has electricity	Age	Yers of education	% migrant
	(1)	(2)	(3)	(4)	(5)	(6)
Downstream×Open	0.131***	0.0205	-0.00100	0.0426	-0.0121	0.0881***
	[0.0422]	[0.0277]	[0.0200]	[0.162]	[0.144]	[0.0314]
Downstream	-0.0142	-0.0411*	0.00433	-0.0835	-0.119	-0.0327
	[0.0307]	[0.0239]	[0.0143]	[0.132]	[0.108]	[0.0272]
Open	-0.0455	-0.00637	-0.0104	-0.136	-0.126	-0.0432*
	[0.0293]	[0.0209]	[0.0181]	[0.136]	[0.112]	[0.0233]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Nb open mines	Yes	Yes	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes
N	61690	61690	61179	61690	61690	38834
R2	0.608	0.489	0.547	0.681	0.463	0.185

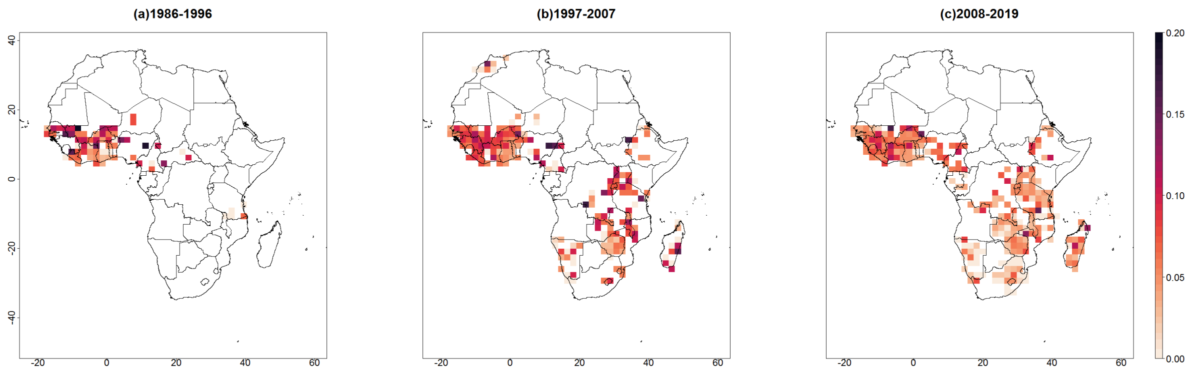
Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. The variables Downstream and Opened are dummies which indicate whether the individual lives in a village downstream of at least one mining site, and whether the site opened before the birth year of the child. Each village DHS is paired to only one mining site, so that each individual appears only once in the regression. Other variables are control variables. The sample focuses on heavy metal mines.

Figure 30: Number of open mines during the birth year, between first and last wave



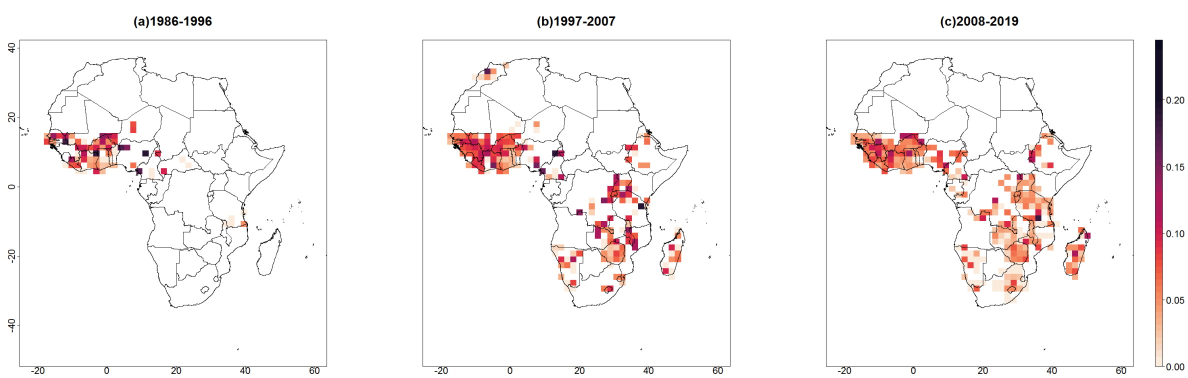
Notes: The figure represents the number of mines that were opened during the birth year of children located within our topographical treatment sample by country, and the number of mines that were opened during the birth year of children located within our topographical treatment sample and which opened between the first and last year of the survey for each country.
Sources: Authors' elaboration on SNL and DHS data.

Figure 31: Spatial variation of 12-month mortality rates per period - restricted sample



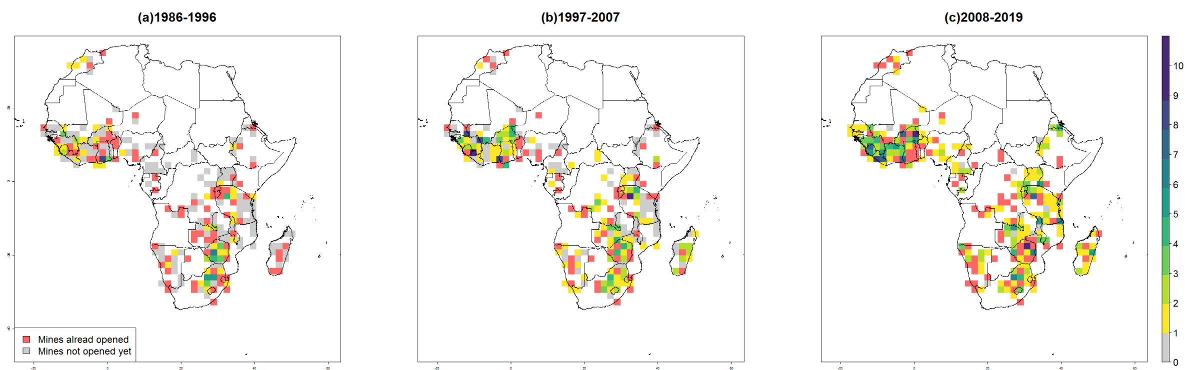
Notes: The figure represents the means of 12-month mortality rates averaged at the grid level over: (a) 1986-1996, (b) 1997-2008, and (c) 2008-2019, for the sample of the main analysis. The mortality rates are estimated without the children that did not reach 12 months of age at the time of the survey.
Sources: Authors' elaboration on DHS data.

Figure 32: Spatial variation of 24-month mortality rates per period - restricted sample



Notes: The figure represents the means of 24-month mortality rates averaged at the grid level over: (a) 1986-1996, (b) 1997-2008, and (c) 2008-2019, for the sample from the main analysis. The mortality rates are estimated without the children that did not reach 24 months of age at the time of the survey.
Sources: Authors' elaboration on DHS data.

Figure 33: Spatial variation of mine opening per period - restricted sample



Notes: The figure represents the number of mines that opened during the periods over the grid area (160 km on average). A red grid cell represents an area where no mine opened over the period, but where at least one mine has opened before the period. A grey cell represents an area where no mine opened over the period, but where at least one mine will open in the future.
Sources: Authors' elaboration on SNL data.

E Heterogeneity

Table 30: Effects of industrial mining activity, across sub-regions

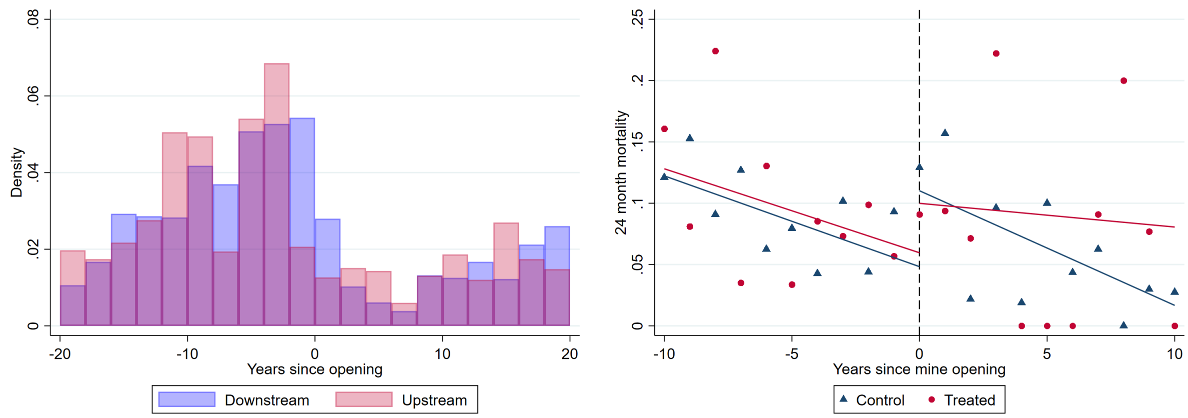
	Mortality under 24 months		
	Western Africa	Eastern Africa	Central and Southern Africa
	(1)	(2)	(3)
Downstream×Open	0.0443**	0.0292*	-0.0304
	[0.0176]	[0.0154]	[0.0225]
Downstream	-0.0153	-0.0369***	0.0538**
	[0.00946]	[0.0116]	[0.0227]
Open	0.00523	-0.0137	0.0120
	[0.0133]	[0.0157]	[0.00857]
Controls	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes
N	21,006	13,484	20,014
R2	0.0521	0.0457	0.0238
Outcome Mean	0.0981	0.0712	0.0836

Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

F Dynamic effects - pre-trends and event study

In this section, we display the appendix figures associated with Section 8, giving the same analysis as in the main section, restricted to the balanced sample. Figure 34 plots the parallel trends, while figure 35 plots the event study for the balanced sample.

Figure 34: Linear trends of 24-month mortality - balanced Sample

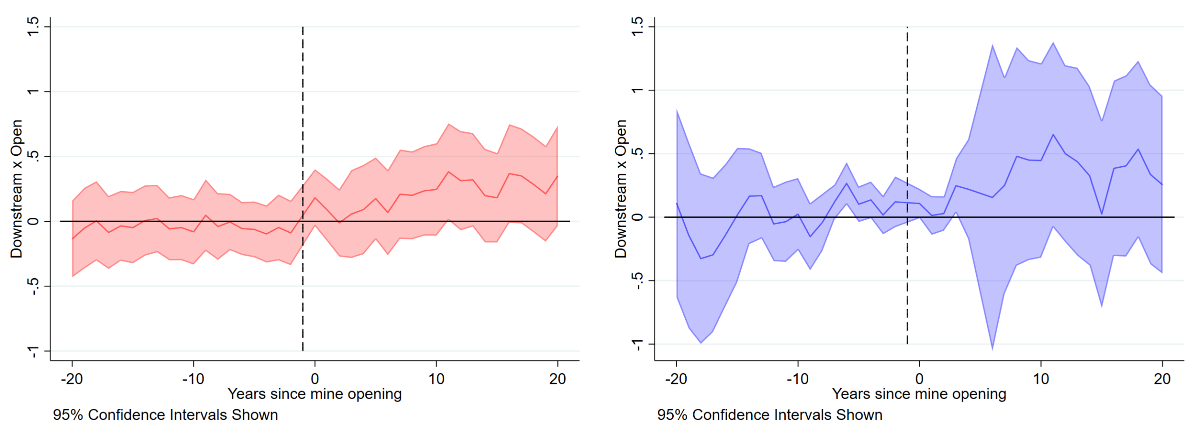


(a) Balanced sample

(b) Balanced sample

Notes: Panel (a) gives the distribution of the number of observations per opening year. Panel (b) plots the trends of the 24-month mortality rates according to the year of the mine opening. The figures are made for the balanced sample and include neither control variables nor fixed effects.

Figure 35: Event study - dynamic effect of mine opening on under 24-month mortality - balanced sample



(a) Balanced sample - upstream

(b) Balanced sample - downstream

Notes: Panel (a) plots the event study for the upstream villages, while panel (b) plots the event study for the downstream villages for the balanced sample. Controls and fixed effects are the same as in the main analysis (column (4), Table 2).

F.1 Sensitivity analysis

Table 31 shows that our results are stable when controlling for a dummy variable indicating whether the mine opening year has been found manually or was given directly by the SNL database (column (2)).

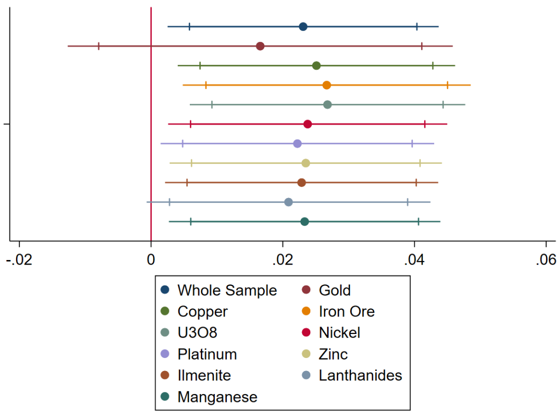
Table 31: Effects of industrial mining opening, controlling for handwork.

Outcome	Mortality under 24 months			
Specification	Main result (1)	Adding control (2)	SNL database (3)	Handwork (4)
Downstream × Open	0.0218** [0.0108]	0.0218** [0.0108]	0.0222 [0.0351]	0.0344** [0.0138]
Dummy handwork		0.0254 [0.0335]		
Downstream	-0.0211*** [0.00739]	-0.0212*** [0.00739]	-0.0167 [0.0246]	-0.0316*** [0.00844]
Open	-0.00496 [0.0101]	-0.00489 [0.0101]	-0.0194 [0.0476]	-0.00973 [0.0126]
Controls	Yes	Yes	Yes	Yes
Birthmonth FE	Yes	Yes	Yes	Yes
Country-birthyear FE	Yes	Yes	Yes	Yes
Mine SB FE	Yes	Yes	Yes	Yes
Mine SB-birthyear trend	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes
N	35,638	35,638	6,702	22,017
R2	0.0511	0.0511	0.0615	0.0641
Outcome mean	0.0873	0.0873	0.0727	0.0954

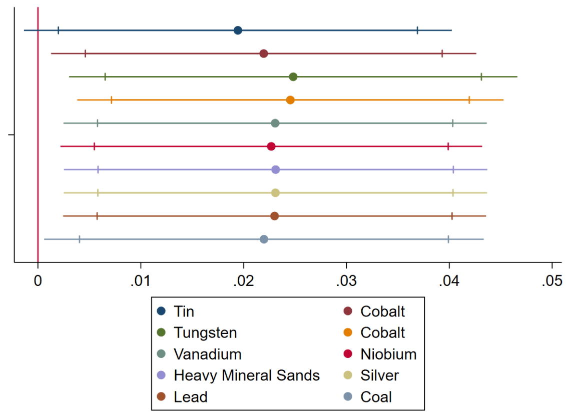
Notes: Standard errors clustered at the DHS village level, $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. Columns (1) and (2) rely on the same sample and controls as Table 2 Column 2. Column (2) controls for the hand-work, while Columns (3) and (4) split the samples.

Figure 36 displays the DiD estimators for different regressions with restricted samples, meaning while dropping each metal one by one, using the sample for the 24-month mortality rates, and the heavy metals and coal mine sample. This suggests that our main results are not driven by a specific metal. Accordingly, Figure 37, plots the DiD estimators while dropping countries one by one to show that our analysis is not driven by a particular country.

Figure 36: Regression results when dropping commodities one by one

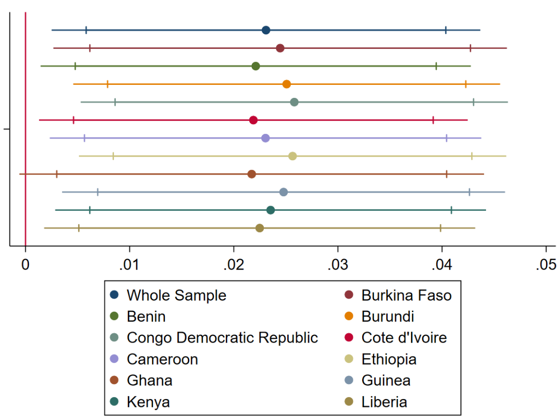


(a)

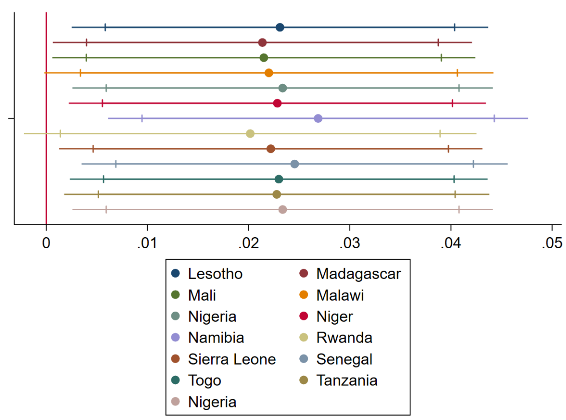


(b)

Figure 37: Regression results when dropping countries one by one



(a)



(b)

Sources: Authors' elaboration on DHS and SNL data.

G Geographical Treatment

In this section, we propose to replicate the empirical strategy of Benshaul-Tolonen [2018], who finds that a mine opening is associated with a 5.5 p.p decrease in 12-month mortality. The identification strategy relies on a treatment based on proximity, comparing individuals living nearby to those living further from an industrial mine. In this estimation, geographical proximity is used as a proxy exposure to industrial mining activity, including both positive and negative externalities, such as exposure to mining pollution. The identification strategy relies on a DiD strategy. It compares, within each district, the infant mortality in areas within 10 km of a mine deposit (treatment group) with infant mortality in DHS clusters further away from a mine deposit (10-100 km, control group), before and after the opening of the mine deposit. As the strategy is a two-way fixed effects, including a district fixed effects, the comparison is made within each district. The identification can be formally written as:

$$\begin{aligned} Death_{i,v,c,m,SB} = &\alpha_0 + \alpha_1 Opened_{birthyear,i,v} + \alpha_2 MineDeposit_{[0;10km]v} \\ &+ \alpha_3 Opened_{birthyear,i,v} \times MineDeposit_{[0;10km]v} + \alpha_4 X_i \\ &\gamma_d + \gamma_{d-bthtrend} + \gamma_{c,birthyear} + \epsilon_v \end{aligned} \quad (2)$$

where $Death_{i,v,c,district}$ is a dummy variable equal to 1 if child i from DHS village v (within district d) of country c has reached the n^{th} month and has died (n being 12 for the 12-month mortality, 24 and so on). $Opened_{birthyear,i,v}$ is a dummy variable equal to 1 if at least one mine is located within 10 km for the treatment group, or within 100 km for the control group, has opened before child i's year of birth (this cohort comparison can be considered here as a source of triple difference).$MineDeposit_{[0;10km]v}$ is a dummy variable of proximity (1 if village DHS v is within 10 km of a mine deposit, 0 if it is within 10-100 km), X_i a vector of child/mother level controls (mother's age and age square, years of education, urban status). Finally, γ_d is a district fixed effects, $\gamma_{d-bthtrend}$ a district birth year linear trend, and $\gamma_{c,birthyear}$ a country birth year fixed effects. Please note that the matching of DHS clusters to mines relies on the same strategy as in Benshaul-Tolonen [2018], and assigns a DHS cluster to the closest mine (without consideration for its opening status). Thanks to this pairing, if a DHS cluster is both in the treatment and control groups of two different mines (i.e., within 10 km of mine A and within 10-100 km of Mine B), we assign it mechanically to the treatment group (so linked to mine A). This creates a bias explained in Section 2.4, which explains the choice for a district fixed effects and reduces the noise linked to DHS random displacements.

Firstly, we give our estimators from the exact replication of Benshaul-Tolonen [2018]'s results, using our own calculation, and find similar impacts (Tables 32 and 33). Second, we propose the replication of the results using our extended sample, including more countries, DHS waves, types of mines, and mines check by hand, and show that the results from Benshaul-Tolonen [2018] are mainly determined by the choice of countries.

G.1 Exact replication of Benshaul-Tolonen [2018]

The geographical treatment proxies exposure to mining activity using the distance to the site and partly follows the analysis from Benshaul-Tolonen [2018], and finds contradictory impacts on infant mortality. To better understand how our results can be compared to the literature, we propose in this section a replication exercise of the main results from Benshaul-Tolonen [2018][23].

For this replication analysis, we used the same mines and DHS survey rounds as Benshaul-Tolonen [2018]. Please note that we have few differences in terms of the whole sample, as Benshaul-Tolonen [2018] counts 37,365 children *vs* 41,902 for us, that might be explained by the way we calculated the 100 km buffer distance [25]. A main difference between our paper and Benshaul-Tolonen [2018] is the independent variable, as we use the opening of the industrial mine as a shock, whereas Benshaul-Tolonen [2018] uses the activity status based on production data given by the SNL product. This accounts for interim years, between the opening and final closiure of the mine, when production has been on hold. In this section, we replicate this exact same variable.

Table 32 displays the replication of the main results from Benshaul-Tolonen [2018], Table 2. We find that a mine opening within 10 km is associated with a 4.7 p.p decrease in infant mortality rate, while Benshaul-Tolonen [2018] found a 5.5 p.p decrease. Our results are slightly less significant than that of Benshaul-Tolonen [2018], and we identify a different impact according to gender, with a significant reduction of girl mortality rate of 7 p.p *vs* a non-significant reduction for boys, which differs from the previous study. To

[23] Please note that a first difference between the two analyses is the sample, as Benshaul-Tolonen [2018] uses 43 gold mines that match with 31 DHS surveys from nine countries (Burkina Faso, Cote D'Ivoire, Ethiopia, Ghana, Guinea, Mali, Senegal, Tanzania, and DRC [24]). However, when pairing the DHS cluster to the same industrial mining sites from Benshaul-Tolonen [2018], no DHS from DRC remained. In the end, the analysis is only on the 8 first countries, in accordance with Figure A6 from the Appendix of Benshaul-Tolonen [2018]), for an entire sample of 48,151 1-year-old children.

[25] In the replication codes of Benshaul-Tolonen [2018], one can observe that the distance has been determined using the Stata command *nearstat* [...] dband(0,25), which relies on different projections (not specified) as ours from *R* libraries, explaining the small sample differences.

Table 32: Replication Benshaul-Tolonen [2018] Main Results

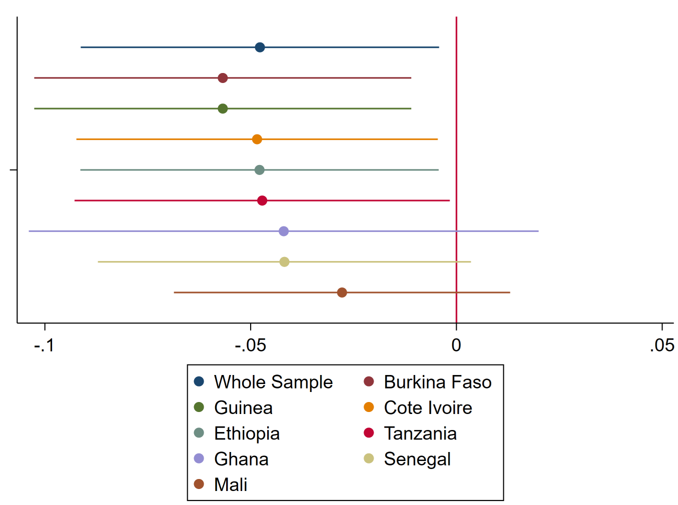
Dependent variable	Infant mortality first 12 months			
Sample :	Children (1)	Children drop spillover (2)	Boys (3)	Girls (4)
Industrial × mine deposit (at birth)	-0.0472**	-0.0474*	-0.0289	-0.0781***
	[0.0230]	[0.0260]	[0.0320]	[0.0301]
Mine deposit [0;10km]	0.0392**	0.0546***	0.0517**	0.0561**
	[0.0169]	[0.0195]	[0.0229]	[0.0231]
Mother's age	-0.0145***	-0.0154***	-0.0155***	-0.0152***
	[0.00190]	[0.00210]	[0.00274]	[0.00297]
Mothers's age × Mother's age	0.000222***	0.000236***	0.000223***	0.000245***
	[0.0000302]	[0.0000335]	[0.0000435]	[0.0000475]
Years edu.	-0.00214***	-0.00230***	-0.00272***	-0.00184**
	[0.000489]	[0.000547]	[0.000827]	[0.000760]
Urban$_h h$	-0.0125***	-0.0120**	-0.00710	-0.0183***
	[0.00428]	[0.00480]	[0.00687]	[0.00659]
Birth-month FE	Yes	Yes	Yes	Yes
Country birth year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
District BirthYear trend	Yes	Yes	Yes	Yes
Drop10-30 km away	No	Yes	Yes	Yes
Drop investment phase	No	Yes	Yes	Yes
Mean of outcome	0.102	0.104	0.110	0.099
Mean(treatment, pre-treatment)	0.154	0.163	0.173	0.153
Observations	41902	34228	17534	16694

Notes: $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. Standard errors clustered a DHS cluster level. The variables Mine deposit [0;10km] and Industrial × mine deposit (at birth) are a replication from Benshaul-Tolonen [2018] and indicate whether the child is born within 10km of at least one industrial mining site and whether this site was active at the time of the birth. All regressions control for mother's age, age square, mother's education and whether the household is urban, for district, birth month and country-birth year. The main outcome is infant mortality in the 12 months since birth. Columns 2-5 drop the two years preceding th opening year, defined as investment phase in Benshaul-Tolonen [2018] and the individuals living within 10-30km of the closest industrial mine. Mean (treatment, pre-treatment) is the sample for the treatment group before the mine were active. dummies which indicate whether the individual lives in a village within at least one mining site, and whether the site opened before the birth year of the child. Each village DHS is paired to the closest opened mining site, so that each individual appears only once in the regression. Other variables are control variables.

follow Benshaul-Tolonen [2018]'s example, we excluded, in columns (2)-(5) of Table 32, individuals born within 10-30 km of the closest industrial mining site and those born in the two years before the opening of a mine, which is a proxy for the investment phase according to Benshaul-Tolonen [2018].

Please note that in accordance with the descriptive statistics from Benshaul-Tolonen [2018] we have a very high mean of 12-month mortality rates (from 10 to 17% according to the groups) in the sample. These are relatively high numbers that do not match with the World Bank data. This is because Benshaul-Tolonen [2018], in order to measure the mortality and avoid mechanically increasing the mortality rate of these cohorts [26] drops all the individuals who are still alive but did not reach the age of 12 months. For replication purposes, we propose keeping this variable and correcting this in Table 33 where we observe average mortality rates of around 7%. Figure 38 replicates Figure A6 from Benshaul-Tolonen [2018], which shows the coefficient estimates of the main regression for *industrial* × *mine deposit* on infant mortality, each regression excluding the sample from one country as indicated by the country name. Figure 38 shows that results are highly sensitive to the presence of Mali, Senegal, and Ghana in the sample (whereas they do not consist of the majority of the sample (5,847, 1,098 and 5,595 respectively).

Figure 38: Regression results when dropping one country at a time



Sources: Authors' elaboration on DHS and SNL data.

[26]We can read in the codes that if the living individuals were dropped, the children from these specific cohorts that died before 12 months were not dropped: mechanically, the mortality rates for all the years preceding the survey rounds are 100%, which explains the high mean of outcomes.

Table 33: Replication Benshaul-Tolonen [2018] Main Results

Dependent variable	Infant mortality first 12 months corrected			
Sample :	Children	Children drop spillover	Boys	Girls
	(1)	(2)	(3)	(4)
Industrial × mine deposit (at birth)	-0.0494**	-0.0471*	-0.0439	-0.0631**
	[0.0229]	[0.0244]	[0.0317]	[0.0298]
Mine deposit [0;10km]	0.0394**	0.0587***	0.0682***	0.0513**
	[0.0179]	[0.0198]	[0.0255]	[0.0235]
Mother's age	-0.0118***	-0.0123***	-0.0120***	-0.0124***
	[0.00175]	[0.00196]	[0.00256]	[0.00283]
Mothers's age × Mother's age	0.000182***	0.000189***	0.000172***	0.000203***
	[0.0000279]	[0.0000312]	[0.0000405]	[0.0000452]
Years edu.	-0.00143***	-0.00152***	-0.00204***	-0.000803
	[0.000455]	[0.000510]	[0.000772]	[0.000715]
Urban$_h h$	-0.0106***	-0.0113***	-0.00501	-0.0196***
	[0.00384]	[0.00436]	[0.00661]	[0.00600]
Birth-month FE	Yes	Yes	Yes	Yes
Country birth year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
District BirthYear trend	Yes	Yes	Yes	Yes
Drop10-30 km away	No	Yes	Yes	Yes
Drop investment phase	No	Yes	Yes	Yes
Mean of outcome	0.079	0.080	0.083	0.077
Mean(treatment, pre-treatment)	0.109	0.118	0.120	0.115
Observations	40386	32873	16823	16050

Notes: $^{*}p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$. Standard errors clustered a DHS cluster level. The variables Mine deposit [0;10km] and Industrial × mine deposit (at birth) are a replication from Benshaul-Tolonen [2018] and indicate whether the child is born within 10km of at least one industrial mining site and whether this site was active at the time of the birth. All regressions control for mother's age, age square, mother's education and whether the household is urban, for district, birth month and country-birth year. The main outcome is infant mortality in the 12 months since birth. Columns 2-5 drop the two years preceding th opening year, defined as investment phase in Benshaul-Tolonen [2018] and the individuals living within 10-30km of the closest industrial mine. Mean (treatment, pre-treatment) is the sample for the treatment group before the mine were active. dummies which indicate whether the individual lives in a village within at least one mining site, and whether the site opened before the birth year of the child. Each village DHS is paired to the closest opened mining site, so that each individual appears only once in the regression. Other variables are control variables.

G.1.1 Replication using an extended sample

Tables 34 and 35 display the results, replicating Benshaul-Tolonen [2018]'s estimation strategy with our overall sample of mines and DHS surveys. Table 34 focuses on the 12-month mortality rates and shows that we find a significant reduction in infant mortality by 0.8 p.p only when controlling for migrants (column (2)). Columns (1) and (2) display the results for the whole sample, while columns (3) and (4) drop the spillovers effects (areas between 10-30 km, and the two years before the opening of the mine, which represents the investment phase in Benshaul-Tolonen [2018]). Columns (5) and (6) replicate the analysis for the male sample while columns (7) and (8) for the female.

Table 35 displays the results for the 12-month (columns (1)-(4)) and 24-month mortality rates (columns (5)-(8)), and compares the estimators when not including the migrant control variable (columns (1), (3), (5), and (7)), and when including it (columns (2),(4),(6), and (8)). We also present the estimators for the restricted sample of rural areas (columns (3),(4), (7), and (8)). Again, we observe a significant reduction of 12-month mortality rates in column (2), i.e., for the overall sample while controlling for migrants, and find no results otherwise. This absence of results suggests that using proximity as a proxy for exposure to mining activity averages contradictory effects, including both positive and negative externalities, and shows the importance of our main estimation strategy which relies on topographical position.

Figure 39 plots the linear trends of the 12- and 24-month mortality rates for the geographical treatment, including our overall mine and DHS sample. We see that the linear trends assumption seems to be validated for the 24-month mortality rates, but not for the 12-month mortality rates.

Table 34: Geographic Treatment

	Infant mortality first 12 months							
	All		Drop spillover		Boys		Girls	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Indus. × deposit	-0.00259	-0.00823**	-0.00189	-0.00575	0.00250	-0.00302	-0.00513	-0.00807
	[0.00329]	[0.00418]	[0.00407]	[0.00537]	[0.00570]	[0.00764]	[0.00522]	[0.00674]
Deposit	0.00130	0.00374	0.00103	-0.000128	0.00632	0.0113	-0.00366	-0.0109*
	[0.00252]	[0.00317]	[0.00392]	[0.00500]	[0.00546]	[0.00708]	[0.00513]	[0.00628]
Birth order	0.00389***	0.00315***	0.00360***	0.00320***	0.00349***	0.00304***	0.00382***	0.00356***
	[0.000345]	[0.000428]	[0.000423]	[0.000518]	[0.000606]	[0.000742]	[0.000549]	[0.000671]
Mother's age	-0.0105***	-0.0107***	-0.0102***	-0.0110***	-0.0116***	-0.0128***	-0.00884***	-0.00924***
	[0.000541]	[0.000668]	[0.000669]	[0.000824]	[0.000953]	[0.00119]	[0.000903]	[0.00111]
agesquare	0.000147***	0.000151***	0.000142***	0.000156***	0.000163***	0.000183***	0.000121***	0.000127***
	[0.00000853]	[0.0000106]	[0.0000106]	[0.0000131]	[0.0000150]	[0.0000187]	[0.0000142]	[0.0000175]
Years edu.	-0.000877***	-0.00103***	-0.000874***	-0.00101***	-0.000881***	-0.00103***	-0.000873***	-0.000968***
	[0.000135]	[0.000167]	[0.000164]	[0.000200]	[0.000238]	[0.000290]	[0.000216]	[0.000265]
Urban	-0.00610***	-0.00725***	-0.00708***	-0.00906***	-0.00825***	-0.0111***	-0.00563**	-0.00622**
	[0.00135]	[0.00172]	[0.00169]	[0.00214]	[0.00235]	[0.00297]	[0.00227]	[0.00289]
migrant		0.00543***		0.00509***		0.00255		0.00754***
		[0.00120]		[0.00145]		[0.00208]		[0.00196]
Constant	0.229***	0.232***	0.226***	0.240***	0.251***	0.273***	0.201***	0.206***
	[0.00826]	[0.0101]	[0.0103]	[0.0126]	[0.0146]	[0.0181]	[0.0138]	[0.0169]
Birth-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ctry-bthyr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dist-bthyr trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drop10-30 km	No	No	Yes	Yes	No	No	No	No
Drop t-2	No	No	Yes	Yes	No	No	No	No
N	359219	243645	236573	165202	119860	83570	116696	81601

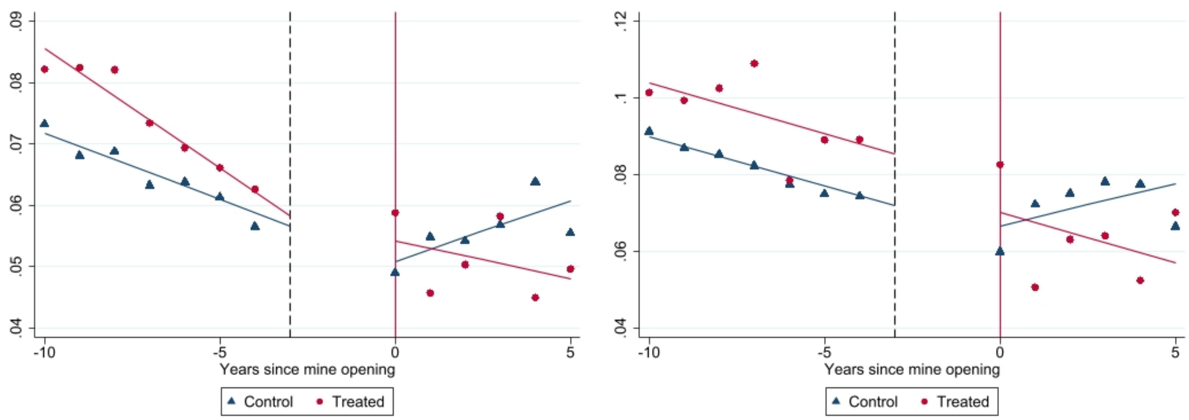
Notes: $^{*}p < 0.1, ^{**} p < 0.05, ^{***} p < 0.01$. Standard errors clustered a DHS cluster level. The variables Mine deposit [0;10km] and Industrial × mine deposit (at birth) are a replication from Benshaul-Tolonen [2018] and indicate whether the child is born within 10km of at least one industrial mining site and whether this site was active at the time of the birth. All regressions control for mother's age, age square, mother's education and whether the household is urban, for district, birth month and country-birth year. The main outcome is infant mortality in the 12 months since birth. dummies which indicate whether the individual lives in a village within at least one mining site, and whether the site opened before the birth year of the child. Each village DHS is paired to the closest opened mining site, so that each individual appears only once in the regression. Other variables are control variables.

Table 35: Effects of industrial mining activity on under 12, 24 mortality - Geographic Treatment - All Households

	Death <12m				Death < 24m			
	All	All	Rural	Rural	All	All	Rural	Rural
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
indus.×deposit	-0.00259	-0.00823**	-0.00259	-0.00627	0.000248	-0.00264	0.000248	-0.00248
	[0.00329]	[0.00418]	[0.00329]	[0.00509]	[0.00431]	[0.00535]	[0.00431]	[0.00657]
Deposit	0.00130	0.00374	0.00130	0.00313	0.000627	0.00121	0.000627	0.000859
	[0.00252]	[0.00317]	[0.00252]	[0.00368]	[0.00321]	[0.00411]	[0.00321]	[0.00477]
Indus.	0.00131	0.00222	0.00131	0.00340	0.00116	0.00122	0.00116	0.00190
	[0.00155]	[0.00200]	[0.00155]	[0.00230]	[0.00201]	[0.00259]	[0.00201]	[0.00297]
Birth order	0.00389***	0.00315***	0.00389***	0.00353***	0.00512***	0.00401***	0.00512***	0.00447***
	[0.000345]	[0.000428]	[0.000345]	[0.000500]	[0.000440]	[0.000549]	[0.000440]	[0.000642]
Mother's age	-0.0105***	-0.0107***	-0.0105***	-0.0116***	-0.0115***	-0.0124***	-0.0115***	-0.0140***
	[0.000541]	[0.000668]	[0.000541]	[0.000787]	[0.000704]	[0.000873]	[0.000704]	[0.00103]
Age square	0.000147***	0.000151***	0.000147***	0.000161***	0.000151***	0.000167***	0.000151***	0.000187***
	[0.00000853]	[0.0000106]	[0.00000853]	[0.0000122]	[0.0000110]	[0.0000136]	[0.0000110]	[0.0000159]
Years edu.	-0.000877***	-0.00103***	-0.000877***	-0.000792***	-0.00145***	-0.00157***	-0.00145***	-0.00132***
	[0.000135]	[0.000167]	[0.000135]	[0.000219]	[0.000173]	[0.000215]	[0.000173]	[0.000283]
Urban	-0.00610***	-0.00725***	-0.00610***		-0.00940***	-0.00995***	-0.00940***	
	[0.00135]	[0.00172]	[0.00135]		[0.00175]	[0.00222]	[0.00175]	
migrant		0.00543***		0.00514***		0.00727***		0.00630***
		[0.00120]		[0.00144]		[0.00155]		[0.00186]
Constant	0.229***	0.232***	0.229***	0.247***	0.273***	0.286***	0.273***	0.315***
	[0.00826]	[0.0101]	[0.00826]	[0.0120]	[0.0109]	[0.0134]	[0.0109]	[0.0159]
Birthmonth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cty-Bthyr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine Bthyr trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	359,219	243,645	359,219	179,155	265,735	179,729	265,735	132,398
R2	0.0195	0.0235	0.0195	0.0281	0.0289	0.0337	0.0289	0.0393
Mean	0.0630	0.0653	0.0630	0.0688	0.0816	0.0851	0.0816	0.0903

Notes: Standard errors clustered at the village level, $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$. The variables Proximity and Opened are dummies which indicate whether the individual lives in a DHS village within 10 km of at least one mining site, and whether the site opened before the birth year of the child. Each village DHS is paired to only one mining site, so that each individual appears only once in the regression. Other variables are control variables. The sample focuses on heavy metal mines.

Figure 39: Linear trends dropping investment phase - geographical treatment



(a) Infant mortality rate, 12 months (b) Infant mortality rate, 24 months

Sources: Authors' elaboration on DHS and SNL data.