5304 Econometrics @ SSE, Fall 2023

Group 15[1] report on Gittard (2023)

Summary

What are the effects of mining-induced water contamination on infant mortality in Africa? Industrial mining activity has ambiguous effects on local population's health: on one hand local industrial growth, job creation, consumption, access to services and infrastructure (Kotsadam and Tolonen (2016); Mamo et al. (2019)), while on the other pollution, conflicts, corruption, and migration. (Corno and de Walque (2012); Aragón and Rud (2016)). This paper contributes to the literature by building a dataset of industrial mining activity and by taking into account the villages-industrial mining sites topography to capture the effects of mining-induced water pollution. The authors accessed data from the Demographic Health Survey covering 26 countries over the 1986-2018 period and the SNL Mining and Metals data set, resulting in a merged sample of 2,016 mines. The authors then used a two-way fixed effects methodology with topographic position as a proxy for exposure to a mine and a matching strategy to reduce bias when using distance as proxy to compare 12 and 24 months mortality rates of villages upstream vs. downstream, before vs. after the opening of a mine. Results presented indicate that being downstream of a mine opening increases by 27% the 24 months mortality rate and that these effects are persistent up to 3 years after the mine opens, and start during the investment phase.

Assessment

The study focuses on three main aspects; the effects of mining opening on child mortality, child's health outcomes and women's outcomes. The health and socio-economic data used for the analysis is from the Demographic Health Surveys (DHS) and the authors have decided to restrict the analysis to the 26 African countries, out of the 36 included in the survey, that have at least two survey waves and have had a mine opening within at least one of these waves. The reason for the selection is that the study uses a difference-in-differences analysis and hence requires data of sufficient time length both before and after opening of a mine. It would have been possible to conduct a difference-in-differences analysis with just one survey wave but according to the authors, that would avert the possibility of studying long term-effects, which is why countries with one wave have been excluded. The drop of data with only one survey wave for instance led to South Africa, which the authors claim have an intense mining activity, being left out of the analysis. Thus the analysis includes 240 431 children under 5 years of age. Collection of mining data has been carried out through the SNL Metals and Mining database, from which mainly data of the dates at which mines have been opened, how far individuals live from respective mines and whether individuals live upstream or downstream of a mine, has been extracted. To determine individuals' upstream or downstream position relative to a mine, the authors have also used the HydroBASINS sub-basins geographic information.A major concern raised by the authors is the difficulty of pairing a particular data cluster of individuals from the DHS with a particular mining site, since there is a risk of violating the Stable Unit Treatment Value Assumption: other than the paired mines affecting health outcomes in the data cluster.

[1] 42624@student.hhs.se, 42613@student.hhs.se, 42632@student.hhs.se, 25164@student.hhs.se

The way the authors have tried to handle this problem is by combining each DHS cluster with its nearest mine to which the cluster is downstream, if the degree of downstream is up to the third sub-basin level. Otherwise the cluster has been considered to be upstream of its nearest mine. Upstream clusters more than 45 km from a mine have been excluded from the analysis as a way to ensure parallel trends between upstream and downstream clusters.

Empirical Strategy

Two way fixed effects estimator is applied to deal with the problem of heterogeneous treatment effects:

$$Death_{i,t,v,c,SB} = \alpha_0 + \alpha_1 Opened_{t,i,v} + \alpha_2 Downstream_{v,SB}$$

$$+\alpha_3 Opened_{t,i,v} x Downstream_{v,SB} + \alpha_4 X_i + \gamma_{SB} + \gamma_{SB} xt + \gamma_{c,t} + \epsilon_v,$$

where the authors account for the mine sub-basin fixed effect (γ_{SB}), the country-birth-year fixed effect ($\gamma_{c,t}$) and the mine sub-basin linear birth-year trend ($\gamma_{SB} xt$). The motivation behind the two way fixed effects is as follows. The country-birth-year should account for a portion of the cultural practices of the villages and how they evolve over time either exogenously or in response to water contamination. Furthermore, it accounts for local governments or non-governmental organizations responding to the water pollution problem. The mine sub-basin fixed effects account for the types of metal being mined (which affects the precise composition of contaminants discharged into the waterways) and the quality of the mine's protocol in waste management. To understand the two-way fixed effect estimator, we turn to the Bacon decomposition theorem which reduces the probability limit of the estimator to the sum of the variance weighted average treatment effect and the variance weighted common trends, minus any change in average treatment effects (i.e. if there is heterogeneity in the treatment effects within the group over time). A problem of the standard two-way fixed effect estimator is that it places negative weights on the effect of some cells, specifically comparisons between groups that are treated more and groups that are treated less. Thus, the researchers applied the heterogeneity-robust DID estimators proposed by de Chaisemartin and D'Haultfoeuille. It is robust to heterogeneous treatment effects since it restricts the focus to control groups that are untreated throughout. Using the balanced sample, which focuses on the villages driving the treatment effects, the researchers claimed it to be consistent with the findings from the dCDH estimators. Since these avant-garde estimators are all subject to existing scrutiny from econometricians, it is difficult to comment further on the validity of such estimators.

Parallel Trends Assumption

Without data on counterfactual outcomes, the parallel trend assumption is tested using pre-treatment data to show parallel pre-treatment trends. The researchers found that both treated and untreated villages followed similar mortality trends. As emphasized, the major identification assumption is that mine openings cause differences upstream and downstream of the mine only through decreased water quality. Given that historical data has been unable to predict black swans, for instance, the credibility of these tests for pre-treatment trends depend on our belief in historical continuity. For DID, this means reliably accounting for all possible sources of endogeneity. Gittard and Hu implicitly employ rely on geographical determinism, arguing that a mine opening is an exogenous outcome since the initial location of valuable raw materials are "as good as randomly placed". Firstly, even if the initial distribution of valuable raw materials are exogenous, the decision for some local governments to carry out geological surveys and approve the mine opening while other local governments refuse to do so is endogenous.

Human decisions are involved, which means we may have to consider the endogenous variables of local governance and the presence of corruption. In the ideal scenario, we have to consider the political influence of local villages in influencing policy outcomes. Hence, if we then believe that mines are more likely to open when there are more corrupt governance practices, then it may already be the case that villages downstream already have worse prospects. Nevertheless, we believe that the existing controls employed by the researchers sufficiently mitigate the issue.

Identification Risks

Other forms of pollution

In Section 11, the researchers acknowledged the role of other transmission mechanisms of pollution since their results likely underestimate the overall health effect of mining activity. Their reasonable hypothesis is that air pollution is less important since for heavier particles the effect is localized, and wind patterns are less likely to be correlated with the topographical position of being upstream or downstream.

Migration: In and Out

There is difficulty in being fully convinced by the researcher's claim that migration is not a source of selection bias. From the aggregate data, 59.4 % (Table 25) of women are migrants (i.e. they have migrated prior to settling in their current place of residence), which implies the significance of migration in their socioeconomic lives. Moreover, the focus on in-migration by the researchers is not matched by a similar focus on outward migration. While in-migration is argued to matter because of the search for economic opportunities, outward migration can be driven not just by the search for economic opportunities elsewhere, but instead as a response to the actual or anticipated adverse health effects of mining pollution. Since this relies on aggregated, anonymized data which inherently limits our ability to understand individual observations, the study will only sample mothers who choose to stay. Theorizing that migration depends on effective demand (i.e. willingness and ability to migrate), then it is possible that women who remain may be less willing to migrate. A plausible narrative is that the risk attitude of mothers, which is not captured by existing controls, may influence their decision to stay in downstream villages. Thus, the selection may weaken the parallel trend assumption. Within the data, only 60 % of the respondents (Table 25) responded to the question on migrant status, which means approximately 80,000 of the 240,000 data points. Without any clear justification for the missing entries in the sampled data, we are concerned by potential non-randomness in the missing data which may undermine the representative nature of the data used to rule out migration effects. A plausible explanation is that the DHS survey is tedious such that the question on migrant status is placed at the back. Since responses to all questions are not compulsory, then many may choose to skip it. This poses a concern that the decision to skip is a result of the women's indifference and lowered risk aversion, which as mentioned before may not be representative of the entire women population in these villages. Combined, this poses an identification risk since it undermines our confidence in the parallel trends assumption underlying the DID approach. In Section 11.2.2, the researchers acknowledge the deliberate focus on controlling for inward migration while noting that outward migration effects are left unanswered due to data limitations.

Agglomeration

Gittard and Hu further acknowledge the role of agglomeration in restricting our ability to isolate the effect of mining activity from related economic activity that is caused by mining. For instance, the construction of processing plants, power plants or transport hubs to support mining activity will inevitably contribute to water pollution. Although we can understand agglomeration activity as a possible mediating factor linking mining activity and water pollution, the pragmatic need to isolate out the real impact of mining activity on water pollution is to apply pressure on mining companies to reform their practices instead of scapegoating nearby economic activity. It remains difficult to disentangle mining activity from other forms of economic activity without data on water samples. Ideally, these water samples will measure levels of cyanide or chemicals used exclusively by mining companies and are far less likely to be used by other local economic activity.

Resource and Water Conflict

Although conflict is acknowledged as a factor uncontrolled for in the study, we find no likely cause to expect a differential impact on downstream villages relative to upstream villages. However, a plausible narrative is that the act of mining and the creation of mining-induced water contamination both exacerbate water conflict. We hypothesize this would disproportionately affect downstream villages. Although this causal mechanism can be investigated in future research, for the purposes of this study it is a source of omitted variable bias. Conflict affects mining intensity and vice versa, while also increasing child mortality from the pursuant violence. If we assume that conflict affects all African villages equally, we can say that conflict has a positive bias on child mortality rates and lay the issue to rest. Given the differing security landscape across the continent and the way in which only a portion are represented within the researchers' selected sample, the interpretation of the outcomes should be made with a caveat.

Credibility of Overall Results

The effects on child mortality are split up between 12- and 24-month mortality rate. According to the results, the 24-month mortality rate increase by 2.18 percentage points for children downstream of a mine compared to upstream, from 8.7% to 10.9%, with a stronger effect in rural areas. For the 12-month mortality rate on the other hand, there are no statistically significant differences between living upstream or downstream, which the authors explain by a higher probability of being breastfed for children under 12 months. The results control for birth order, mother's age, mother's age square, the mother's years of education, urban, and the intensity of the river network. The researchers have employed randomization inference which re-allocates treatment randomly to construct a distribution of possible outcomes, and find that their findings are outside the 95 % confidence interval. Altogether, their results are credible.

Policy Significance

The researchers estimate that up to 1.1 % of newborns die due to mining-induced water pollution. Furthermore, in Table 21 they demonstrate that membership of the Extractive Industries Transparency Initiative (EITI) does not improve the outcomes on child mortality. They expose the conundrum faced by developing countries that have a duty of care to address child mortality, but are concurrently dependent on mining for economic development.

The study implies broader mining reforms are needed to address the inadequate and potentially lip-service nature of EITI. What the paper fails to achieve with a foolproof approach, it more than compensates with its success and ambition to use topography as a basis for DiD on the African sub-continent.

Ways to improve

Developments using Existing Data

The researchers claimed that breastfeeding acts as a natural buffer for newborns against the detriments of mining-induced water pollution. An obvious way to extend analysis is to examine how the timing decision for mothers to wean newborns (i.e. before or after six months) is affected by mining-induced water pollution. From our own preliminary research, the DHS household survey does include, in areas with high malaria incidence, a question on the use of household nets. We can restrict the dataset to the subset which are asked the question (i.e. villages with high malaria incidence), to examine whether mining-induced water pollution has a more, or less pronounced effect on child mortality. The existing placebo tests only account for sexually transmitted disease and tuberculosis, but do not account for waterborne diseases. A hypothesis, although challenged by recent epidemiological studies, is that water pollution will have a short-term effect of decreasing the incidence of malaria by making it harder for malaria vectors to breed. To further examine whether or not such a hypothesized effect will fade out over time, we would need to further shrink the data set to include observations with many more rounds of DHS responses than the current two set by the researchers.

Additional Robustness Checks

Apart from conducting placebo tests on conflict, the limitations of the aggregated data mean that many of doubts cannot be fully addressed. Ideally, with panel data allowing the researcher to track individual households, outward migration can be examined through parametric models estimating attrition in panel data. Hence, the researchers have commendably ruled out other significant causal channels using the given data. Given the advancements in staggered DID, we note Rambachan and Roth (Working Paper) criticism on the over-reliance on statistical significance in interpreting the pre-treatment trends. These tests have lower power and may result in biased point estimates. As recommended, the researchers can construct "confidence sets" for the treatment effect by assuming that the counterfactual difference in trends is somewhat approximated by the difference in trends in the pre-treatment phase. Nevertheless, the researchers use of contextual knowledge on the overall declining trend in child mortality rates to show the absence of pre-trends between upstream and downstream villages within the descriptive data will likely prove to be more effective than sophisticated estimators.

Concluding Remarks

Overall, the paper had good flow and was a pleasure to read. The researchers took great care in visualized the data and in building towards a convincing narrative. The reasoning used behind the event studies to justify the absence of glaring pre-treatment trends was clear, with most remaining doubts clarified or mentioned in the relevant Appendices.

References

ARAGÓN, F. M. AND J. P. RUD (2016): "Polluting Industries and Agricultural Productivity: Evidence from Mining in Ghana," *The Economic journal (London)*, 126, 1980–2011.

CORNO, L. AND D. DE WALQUE (2012): "Mines, Migration and HIV/AIDS in Southern Africa," *Journal of African economies*, 21, 465–498.

GITTARD, M. (2023): "Climate Change, Droughts, and Water Pollution in Sub-Saharan Africa," Ph.D. thesis, thèse de doctorat dirigée par Cogneau, Denis et Quirion, Philippe Sciences économiques Marne-la-vallée, ENPC 2023.

KOTSADAM, A. AND A. TOLONEN (2016): "African Mining, Gender, and Local Employment," *World development*, 83, 325–339.

MAMO, N., S. BHATTACHARYYA, AND A. MORADI (2019): "Intensive and extensive margins of mining and development: Evidence from Sub-Saharan Africa," *Journal of development economics*, 139, 28–49.

RAMBACHAN, A. AND J. ROTH (Working Paper): "An Honest Approach to Parallel Trends," .