Columbia University

Department of Economics
Discussion Paper Series

Mines: The Local Welfare Effects of Mineral Mining in Developing Countries

Jan von der Goltz, Prabhat Barnwal

Discussion Paper No.: 1314-19

Department of Economics
Columbia University
New York, NY 10027

February 2014

Mines

The Local Welfare Effects of Mineral Mining in Developing Countries*

Jan von der Goltz †‡ Prabhat Barnwal†

February 23, 2014

Abstract

Do residents of mining communities face health-wealth trade-offs? We conduct the first extensive investigation assessing this question using micro-data from communities near about 800 mineral mines in 44 developing countries. Mining communities enjoy a substantial medium-term rise in asset wealth (0.3σ), but experience a nine percentage point increase in anemia among adult women, and a five percentage point increase in the prevalence of stunting in young children. Both of these health impacts have previously been linked to metal toxicity - and in particular, exposure to high levels of environmental lead. Benefits and costs are strongly concentrated in the immediate vicinity ($\leq$ 5km) of the nearest mine. We find no systematic evidence of general ill health, and we observe health impacts only near mines of a type where lead pollution is to be expected. Identification is based on a mine-level and mother-level panel, and in the cross-section, on group effects. A novel instrumental variable serves as a cross-sectional robustness check. To make plausible that the observed health impacts are due to pollution, we develop difference-in-difference tests based on the known association of certain mine types with lead pollution, and based on the pathophysiology of lead toxicity. Our results represent the first comprehensive assessment of the local welfare impacts of mining in developing countries, and add to the evidence suggesting that communities near industrial centers in developing countries face information or cost constraints that limit their choice sets.

Key Words: *Mines, Health, Development*

*The authors particularly thank Douglas Almond, Joseph Graziano, and Eric Verhoogen for their guidance and advice. We are indebted to Jesse Anttila-Hughes for his generous support and helpful discussions in the early stages of the project. We thank Ram Fishman, Kyle Meng, Suresh Naidu, Wolfram Schlenker, and seminar participants at Columbia University and the International Post-Graduate Workshop for Sustainable Development 2013 for comments. Nicole Dussault and Dheeraj Sharma provided excellent research assistance. We gratefully acknowledge financial support from the Centre on Global Economic Governance and the Center for International Business Education and Research at Columbia University. All errors are ours. We invite comments and feedback on the working paper.

†PhD program, School of International and Public Affairs, Columbia University. *email:* Jan von der Goltz (jcv2104@columbia.edu), Prabhat Barnwal (pb2442@columbia.edu).

‡Corresponding author.

1 Introduction

This paper studies the local welfare impacts of mineral mining in 44 developing countries. We show that, while residents of mining communities enjoy greater wealth than those living in control communities, there is a trade-off: life near mines exacts a price in terms of a specific health burden.

In any country and at any time, the decision to live near centers of industrial activity involves weighing the promise of economic opportunity against the risk of disamenity caused by pollution. Nowhere is this choice starker than in developing countries. More often than not, opportunities for making a good living are precious and few. At the same time, pollution tends to be poorly regulated, and information on health risks and on countermeasures scarce. Poor infrastructure and inflexible housing markets commonly make commuting to avoid pollution impracticable. Yet, while "the literature on the health effects of pollution has advanced greatly in the last two decades, almost all of this research has been conducted in developed country settings." (Greenstone and Jack (2013))

In the following, we present the first large-scale empirical assessment of the micro-level health-wealth trade-offs posed by industrial operations in developing countries. We use the case of mining and mineral processing to study the benefits and cost of industry to local communities across a broad range of countries. The mining industry in poor countries poses particularly sharp trade-offs, and hence, is an attractive test case to analyze. Single plants generate very high value. The location of ore deposits dictates where mines open, and because of transport cost, often also where smelters locate. Therefore, large operations are found in remote areas where they dwarf any other enterprises - and the economic opportunities generated by the latter. Cities grow around mines, and in some instances, other commercial activity may eventually reduce the importance of the mine. But other mining towns - large and small - remain centered around extractive industry. Mines and smelters therefore tend to play a conspicuous economic role. At the same time, however, mines and smelters are very large polluters, and precisely because they are important sources of revenue, foreign exchange, and employment, they are at risk of weak environmental regulation and enforcement. What is more, some of the pollution near mines is hard to see, hard to predict, and associated with health outcomes that are not widely known to be linked with pollution. Thus, in choosing where to live, local residents may be acting on less than full information.

The importance of mining to development is reflected in a long tradition of research on the macroeconomic implications of mining and the optimal management of mineral resources. However, although mining is among the oldest human economic activities, very little is known about the local economic impact of mining, and about its impacts on other dimensions of well-being.

This includes implications for the health of local communities: although there is an important body of knowledge on pollution near mines and the toxic properties of common pollutants, there is little systematic evidence on the actual clinical consequences of exposure to a mining environment. The paucity of empirical evidence on the local welfare effects of mining is in stark contrast to the strong passions that mining projects habitually evoke among the communities affected. Projects have been supported vociferously, and people have fought over the right to work in mines. Yet, in other places, mining has been desperately opposed, as citizens feared damage to their health and environment. Our work shows that these political passions are grounded in a real trade-off. Across a broad range of settings, the local benefits of mining are real, but so are the costs.

We analyze the effect of mining activity on asset wealth in local communities, on general health, and on two specific health outcomes known to be linked to pollution encountered near mines and smelters, namely anemia in adults and children, and growth in young children. We compile a very large pooled dataset of 104 household surveys from 44 countries to enable us to study the interplay of health and wealth effects across the developing world. Prior evidence suggests that pollution and its health effects are concentrated in the direct vicinity of mines, and our large pooled dataset has the added virtue of allowing us to study impacts at an appropriate level of spatial disaggregation. We then use data from multiple sources on the location, output, and type of mines to estimate the effect of closeness to mines in the cross-section, and the effect of closeness and operational status in pseudo-panels. Cross-sectional identification rests on group effects for each mine, with supporting evidence from an instrumental variables (IV) approach in which we use the location of mineral deposits to instrument for the location of mines. We identify the pseudo-panels with fixed effects at the level of individual mines, or compare only among siblings born to the same mother. In addition to these standard cross-sectional and pseudo-panel models, we develop additional difference-in-difference (DiD) tests based on prior knowledge on the toxicological properties of mining pollution.

Our results show that, at the global mean, long-run asset wealth in mining communities rises by about 0.1σ of an asset index computed for the country where the community is located and the year in which the survey was taken. The medium-term wealth of households living in the vicinity of an operating mine rises by about 0.3σ. Wealth effects are strongly concentrated in the direct vicinity of the mine, and fall to zero beyond a distance of some 15-20km. There are few signs of general ill health in mining communities. By way of contrast, we find clear evidence of two health impacts that are known consequences of exposure to environmental lead and other metals. Thus, women in mining communities show depressed blood hemoglobin values, and increases in the incidence of anemia of three to nine percentage points. They also recover more slowly from blood loss during pregnancy and delivery, a pattern consistent with prior toxicological research.

Children in mining communities suffer some adverse growth outcomes from in utero exposure, with a five percentage point increase in the incidence of stunting - although there is no conclusive evidence of lower birth weight. Growth impacts weaken among older children, perhaps because of the greater wealth enjoyed by households in mining communities. We note particularly that, while our data contains no good measure of cognitive ability, lead exposure has previously been shown to cause cognitive deficits in children at exposure levels below those associated with growth retardation, and far below those associated with overt anemia.

Throughout this paper, we show well-identified reduced-form estimates of the impacts of mining. Because they are reduced form, our health results should be interpreted as the compensated impact of mining, including any positive effects through greater wealth in mining communities. We also provide extensive evidence to make plausible that the observed health effects are likely due to pollution, not other mechanisms. Firstly, we show that the cost mining communities pay in terms of adverse health goes hand in hand with economic benefits, so that there is at least no indication that ill health is caused by deprivation. Secondly, we leverage knowledge on the association of specific mine types with particular toxicants - and by extension, specific health impacts - to conduct extensive falsification tests. We show both that we only observe those health impacts that are expected from exposure to mining pollution, and that we only observe them near mine types strongly associated with the release of contaminants. Finally, we describe a pattern of impaired ability to recover from blood loss after pregnancy, and argue that this effect is consistent with a known pathophysiological pattern of lead toxicity in adults, but not easily consistent with other mechanisms.

The present paper makes three contributions to the economics and public health literature. Firstly, we offer the first comprehensive and methodologically rigorous assessment of the local welfare impacts of mining. Secondly, we add to the very limited knowledge on the trade-off between economic opportunity and health in industrial centers in developing countries; we show that in the important case of mining and smelting, such a trade-off exists, with real benefits and real costs to individuals. Finally, we complement the toxicological and epidemiological literature by demonstrating that the health effects expected from exposure to mining pollution are salient in a well-identified study of the general population near a very large number of mines.

The remainder of the paper is organized as follows. Section 2 discusses prior work on welfare in mining communities, and reviews results from environmental science and toxicology that guide the way we develop hypotheses, measure impacts, and interpret results. Section 3 discusses data, and Section 4 summarizes econometric methods. Section 5 presents results, and Section 6 shows how observed treatment effects vary with covariates describing the economic and geographic environment. Section 7 concludes.

2 Background

This section reviews prior studies of mining and wealth (Section 2.1), and research on health in mining communities (Section 2.2). We then summarize the state of knowledge on environmental pollution near mines, and the relationship between pollution and on the body burden of toxicants (Section 2.3) . We establish that metals, and in particular lead, are of most concern as pollutants in our sample sites, and discuss the toxic effects and toxicokinetics of lead (Section 2.4). The section closes with a brief description of how the health consequences of lead poisoning - anemia, poor growth in children, and cognitive deficits in children - affect those exposed, and what the economic cost of those health conditions might be (Section 2.5).

2.1 Mining and wealth

Economics has traditionally studied mineral mining in the context of optimal resource management, or in a macroeconomic context of growth and public finance.[1]

The welfare impacts of mining at the local level have only recently received some attention. As of the time of writing, we are aware of only two published papers that study mining at the kind of disaggregated scale we consider. In a pioneering paper, Aragón and Rud (2013b) leverage a change in local hiring and procurement policies in a single very large gold mine in Peru to identify local economic impacts. Incomes in communities within 100km of the mine showed an elasticity of 0.17 to production at the mine, alongside significant increases in the price of housing and of locally produced agricultural output, and higher local public spending. Wilson (2012) shows that asset ownership increased among residents of copper mining communities in Zambia during a boom in the 2000s. A working paper by Aragón and Rud (2013a) investigates the impacts of gold mining in twelve operations in Ghana on agricultural productivity. It finds stark decreases in productivity (40%) in the general vicinity (less than 20km) of mines, relative to control areas farther away. These are accompanied by large increases in the poverty headcount (18 percentage points), and decreases in consumption, all driven by dire developments for rural households. The latter two papers and another working paper by Kotsadam and Tolonen (2013) use sub-sets of the micro data from the Demographic and Health Surveys also used for the present study. Kotsadam and Tolonen (2013) argue that mining activity in a comprehensive sample of African mines fosters sectoral shifts in employment out of agriculture (among women, into services, and among men,

[1] For a textbook-level overview of the former, see, e.g., Hartwick et al. (1986); for a survey of the latter, Frankel (2010).

into skilled manual labor) and increases cash employment among women, but is also associated with women leaving the labor force altogether.

Long-term welfare in mining communities was also brought to the attention of the research community by Dell's work on the *mita* forced labor policy in Peru, although the focus of the paper is on institutions and development, rather than the direct welfare impacts of mining per se. (Dell (2010)) In other related work, Acemoglu et al. (2013), Dube and Vargas (2013), and Monteiro and Ferraz (2009) have recently leveraged resource revenue at a disaggregated scale as an instrument in the study of other objects of interest (health expenditure, conflict, and corruption, respectively).

2.2 Health effects of mining

Our paper asks how significant are the ultimate health effects of exposure to pollution from every-day mining and mineral processing operations. Few studies have attempted this before, and to the best of our knowledge, none considers the possible trade-off between wealth and health effects, and none has studied the issue in a well-identified manner across many mine sites.

Prior work in economics on the issue is very limited. Aragón and Rud (2013b) find a significant decrease in general health problems among adults with an expansion of production in the Yanacocha mine, Peru, and no effect among children. In a recent working paper, the same authors find evidence of an adverse effect of mining activity on weight-for-height ratios and the prevalence of cough in children living in the general vicinity of gold mines in Ghana, but no impact on stunting and diarrhea. (Aragón and Rud (2013a)) Both of these results are incidental to the main focus of the respective papers, and neither one of the studies argues that the observed health impacts are likely to result from pollution. Some attention has been given to behavioral correlates of mining activity. Wilson (2012) finds that sexual risk-taking tended to decrease in Zambian copper towns during a boom (a finding we confirm, below). Corno and De Walque (2012) argue that in mining communities in southern Africa, there was increased risk taking and HIV infection among migrant miners, but no such effect among non-migrants.

In the field of public health, some case studies directly analyze health impacts in communities near smelters. Factor-Litvak et al. (1999) find impacts on "intelligence, physical growth, preschool behavior problems, renal function, blood pressure and hematopoiesis," among children of up to 7.5 years of age living in a smelter town in Kosovo. (p. 14) Roels et al. (1976) show that among school age children living near a lead smelter in Belgium, "undue lead absorption was accompanied by early biochemical indications of disturbed heme biosynthesis," but not by overt anemia. (Quotation from Roels et al. (1980), p. 82.) Both papers show comparisons to a matched control group

in addition to dose-response relationships. Dose-response relationships between blood lead (PbB) and lower blood hemoglobin (Hgb)[2] as well as reduced nerve conductivity, have also been reported among children living near a lead smelter in Idaho, U.S. (Landrigan and Baker (1981), Schwartz et al. (1990)) Baghurst et al. (1992) show a dose-response relationship between PbB and IQ in children living near a lead smelter in Port Pirie, Australia. A range of papers by Hendryx and various co-authors (see for instance Hendryx and Ahern (2008)) shows cross-sectional correlations between county-level health outcomes and Appalachian coal mining. We review below the much more extensive literature on individual links in the causal chain from mining to ultimate health impacts, namely studies of (i) pollution near mines, (ii) the body burden of pollutants in residents of mining communities, and (iii) the toxic impacts of substances released near mines.

Pollution due to mining is a special case of industrial pollution, and the latter has been analyzed in large and well-identified studies. (See Currie et al. (2013) for a major recent contribution.) Yet, most of these investigate developed countries; studies of developing countries - certainly with many sites - are rare. Chen et al. (2013) study reduced life expectancy from air pollution due to power generation in China; Ebenstein (2012) assesses the effect of water pollution on gastrointestinal cancer rates in China; and Rau et al. (2013) show cognitive losses from lead exposure near an abandoned toxic waste site in Chile. (Studies of air pollution from urban traffic - e.g., Gallego et al. (2013) - are less immediately related.) At the same time, mining in developing countries is distinct from other industrial activity in important ways. As argued, mining operations are large polluters, but tend to contribute a large share to the local economy; health-wealth trade-offs can therefore be expected to be particularly stark. Conversely, although regulation and enforcement of environmental safeguards is weak in many developing countries, as very large point sources of pollution, mines in principle lend themselves well to environmental regulation and remediation. Furthermore, it is worth recalling that mining quite simply affects very many people - 3% (11%) of our sample population reside within 5km (20km) of a mine we observe. Establishing whether mining causes health effects is therefore of immediate use in flagging a potentially important leverage point for policy.

2.3 Environmental pollution due to mining and its relationship to the body burden of toxicants

A voluminous literature in environmental science has catalogued the pollutants emitted in the course of normal operations near mines of different types. We base the following discussion on Alloway (2013), Ripley et al. (1996), and Wright and Welbourn (2002).

[2]The papers report hematocrit, not hemoglobin levels, but the two measures are closely correlated, and are both used to define anemia.

Local communities can be exposed to pollution from mining and smelting through a multitude of channels. These include dust from mining, handling, and processing; mine waste water; direct exposure to abandoned mine tailings; metals from tailings leached into soil and water; and particulate and gaseous emissions from roasting and smelting. Sometimes, the material extracted is itself of concern, such as in lead, uranium, or asbestos mining. At other times, the concern is with toxicants co-located with the mineral mined and released in processing, such as in the case of heavy metals co-located with gold or silver, and released either in processing or weathering of mine spoils. Finally, sometimes pollutants are used in processing, such as in the case of cyanide leaching of gold, or gold and silver extraction by amalgamation.

We highlight two stylized facts on pollution near mines that are essential to the way we analyze the health impacts of mining in this study.

(i) The kinds of pollutants near a given mine can be predicted well from the ore mined.
Table 1 summarizes in a highly stylized way which pollutants are associated with common (and non-exclusively defined) mine types in our sample. The mapping is far from exact, but serves as a useful first-order approximation. We leverage the association between target minerals and toxicants to compare health effects across mine types, and to show that we find predicted health impacts only near mine types where pollutants specific to the health impact in question are found.

Of particular interest to us is the association of most 'polymetallic' mines where any combination of copper, gold, lead, silver, and zinc are extracted with a characteristic suite of highly toxic pollutants that includes most prominently lead, but also arsenic, cadmium, and chromium. (We will refer to these metals and metalloids as 'heavy metals' - a term that is imprecise in that it does not refer to a well-defined group of chemical elements, but has the advantage of being in everyday semantics associated with the pollutants we have in mind. See Section 3.2.1 for coding notes.) Pollution near polymetallic mines is of particular concern both because these elements are important toxicants, and because the minerals mined are often nested in sulfide rock. When exposed to air and water, the latter will tend to generate sulphuric acid, which in turn leaches metals from the mine's tailings; the resulting acid mine drainage can pose severe health and environmental concerns. (Wright and Welbourn, 2002, p. 439)

(ii) The area in which highly polluted sites are found is typically small, and extends to at most a few kilometers around the mine.
Thus, for lead and in the case of smelters, high exposure ranges have variously been associated with distances from the point source of emissions of 0.5-1.8km (Fontúrbel et al. (2011)); 1-2.5km

(Roels et al. (1980)); 2km (Recio-Vega et al. (2012)); perhaps 2-4km (within the town of Kosova Mitrovica, Factor-Litvak et al. (1999)); about 3km, with an exponential decline of exposure up to this distance (Benin et al. (1999)); and up to 4km (Landrigan and Baker (1981)). Mean PbB in the high exposure groups was for Roels et al. (1980), 13-30 μg/dL; for Recio-Vega et al. (2012), 14-19 μg/dL; for Factor-Litvak et al. (1998), 28-39 μg/dL; and for Benin et al. (1999), 20-40 μg/dL. Landrigan and Baker (1981) found PbB of no less than 40 μg/dL in 87% of children living within 4km of the source, as opposed to 19% of children living at a distance of 4km to 32km.[3] All of these values far exceed the reference value of 5 μg/dL (the 97.5th percentile of blood lead levels found in the U.S.) set by the Centers for Disease Control to "trigger lead education, environmental investigations, and additional medical monitoring," (CDC (2012)) as well as the laxer and more dated 'level of concern' of 10 μg/dL. (Roper et al. (1991))

In the present paper, we do not observe directly environmental pollution or the body burden of toxicants. Rather, we use distance to the nearest mine as a proxy. The choice of a distance cutoff to define the treated group is therefore crucial. Because pollution is known to generally be localized, we look for health effects in a tightly defined treatment group, and consider only households within no more than five kilometers of a mine to have been exposed. This choice is in line with the empirical evidence summarized above on the extent of exposure in the case of lead - the toxicant of most interest to us. It is a highly restrictive choice - and considerably tighter than in other current studies of mining in economics, as is appropriate for our focus on health impacts.[4]

A key benefit of working with a large multi-country dataset is that it allows us to limit our treatment group to residents of very limited areas around mines - while retaining enough statistical power. With perfect data, we might define closeness even more restrictively. In the context of available data, tighter definitions would risk making our distance proxy of exposure noisier both because of the practice of jittering cluster geolocations in our socio-economic data (up to a distance of about 5km - see below), and because of the fact that we work with (imperfectly recorded) mine point locations, while mining operations can measure several kilometers across.[5]

[3]Ranges indicate means for populations groups that differ in age, gender, and other characteristics. Incidence for Landrigan and Baker summarized by the authors. In the case of Benin et al, PbB was predicted from observed environmental pollution; in all other studies, PbB was measured directly.

[4]Wilson (2012) uses a cutoff of 10km, while Aragón and Rud (2013b), Aragón and Rud (2013a) and Kotsadam and Tolonen (2013) use a baseline cutoff of 20km, with sensitivity analysis for other choices.

[5]We also note that across a broad range of studies, the spatial distribution of pollution near mines is both uneven and hard to predict: measurements taken at neighboring sampling points commonly show very different levels of pollution. This implies that even when residents of mining communities have a general sense that they are at risk of being exposed to pollution, they are unlikely to be able to assess very well how high their risk is. In consequence, their incentives to take countermeasures may be weakened. It also suggests that our large population study will be particularly useful in arriving at good general population impacts in the presence of large sampling variability. In terms of modeling pollution, we further conclude that could at best hope to modestly improve our proxy of exposure to contamination if we were to use geographic information in addition to distance.

2.4 Pathophysiological and clinical effects of lead and other metal exposure

As noted, the mines in our sample are associated with characteristic sets of pollutants. Because the latter are known to cause specific health effects, we can develop predictions for expected health impacts that are well-grounded in scientific knowledge. To the degree that we find expected health impacts, but not others, we strengthen our case that impacts are likely due to environmental pollution, rather than any other mechanism.

In our baseline investigation of health impacts, we do not distinguish between different types of mines. Yet, our main concern is with the health consequences of environmental contamination with heavy metals, and in particular, with lead. We focus on heavy metal contamination, first, because the health impacts of exposure are well-known, important, and observable in our data, and second, because a large number of mines in our sample are associated with this type of pollution (337 out of 838 mines other than quarries within the cross-sectional sample). Since mines associated with heavy metal pollution are well-represented in the data, our estimates of average health impacts near all mines will tend to reflect impacts near these particular mines. It is also more feasible to study health effects exclusively near this type of mine than near other types. Among heavy metal pollutants, lead takes a central role, because it is known that the lead body burdens previously measured near mines are high enough to cause health problems. We therefore discuss lead toxicity and health impacts in the following.

The health effects of lead exposure have long been studied and are very well understood. (See ATSDR (2007) for a general discussion.) The wide-ranging effects on adults include reduced blood Hgb and overt anemia, cognitive defects, hypertension, and impaired renal function. In our data, we are only able to observe one of these conditions, namely low blood Hgb/anemia. We adduce two additional unspecific health outcomes as falsification tests, namely miscarriage and general grave illness. For children under five years of age, we analyze two health outcomes that have previously been linked with lead exposure *in utero* and among young children: anemia and growth retardation. We also observe some health outcomes that have not been linked to lead (cough, fever), or linked only weakly or at very high exposure (all-cause mortality, gastro-intestinal problems), and use these as a falsification test. Regrettably, we do not have a good proxy of what is by far the most feared consequence of lead exposure in children, namely impaired cognitive performance and behavioral problems due to neurological damage.

2.4.1 Hematotoxic effects of lead

Lead depresses blood Hgb levels both by shortening red blood cell life spans, and by interfering with enzymes (ALAD and ferrochelatase) essential to the synthesis of the heme component in red

blood cells. Enzyme activity begins to be disrupted "at very low [PbB], with no threshold yet apparent," whether in adults, children, or newborns. (ATSDR, 2007, p. 69) Yet, we do not observe ALAD or other sensitive biomarkers of lead hemotoxicity. The only biomarker we do observe, Hgb, is among the least sensitive ones. Changes of Hgb have been reported at high PbB levels, in excess of 40μg/dL in children, and 50μg/dL in adults. (ATSDR, 2007, p. 69,71f) That is, we expect the hurdle to finding impacts on Hgb to be quite high.

Because Hgb levels are likely affected only by high lead exposure, we devise an additional, more sensitive test of hematotoxic effects. We build upon the insight in Grandjean et al. (1989) that, even when lead exposure is too low to reduce Hgb *levels* in adults, "increased demand on the formation of blood following blood loss could result in a delayed blood *regeneration* in individuals exposed to lead." (p. 1385 - our emphasis) Grandjean et al. (1989) compare recovery from blood donation in workers in a battery factory (with high lead exposure) to workers in a tin can factory (with low exposure). While PbB in lead workers was high (40μg/dL at the median, compared to about 7μg/dL among workers in the tin factory), there was virtually no difference between the two groups of workers in initial Hgb levels. (p. 1386) Yet, 15 days after donating blood, "the lower hemoglobin and erythrocyte count in the lead-exposed workers suggested delayed blood regeneration." (p. 1387) In our study, we show that Hgb recovery is similarly impaired among women in mining communities after another kind of blood loss, namely, pregnancy and delivery - a setup directly analogous to the experiment in Grandjean et al. (1989). Indeed, studying delayed Hgb recovery in women after delivery may be a particularly suitable approach. Thus, Graziano et al. (1991) point out that impaired recovery from blood loss could be mediated by depressed levels of the hormone erythropoietin (EPO) in lead-exposed adults, and show that EPO levels are lowered among pregnant women even at only modest PbB levels. That is, the population we study is known to have the 'right' characteristics for a known physiological mechanism of delayed recovery to be at work.

The effect of lead on children is of particular concern, since children are both more sensitive in their reaction to body burdens of lead, and absorb far larger portions of lead present in water or solids. In the case of anemia, however, we expect effects to be harder to demonstrate in children than in adults. Lead has the same hematotoxic effect in children as in adults. However, by contrast to adults, children are able to compensate for erythrocyte loss by increasing production of EPO, and thus boosting the production of red blood cells. Thus, in a cohort of children aged 4.5-6.5 years, "children with elevated [PbB] maintain normal Hgb, but require hyperproduction of EPO to do so." (Factor-Litvak et al., 1998, p. 361) This ability fades with age: in the study cited, it had decreased significantly in children who had reached 9.5 years of age, and, as noted, it is lost completely in adults, where lead exposure is associated with *lower* EPO, rather than compensatory over-production.

In summary, we would expect Hgb in residents of mining communities to be measurably affected only if there is substantial exposure to environmental lead. An effect should be detected most easily in the recovery of Hgb after blood loss, followed by Hgb levels in adult women, and least readily in Hgb in children.

2.4.2 Neurodevelopmental effects of lead exposure in children

Perhaps the chief concern with lead exposure in children is damage to the developing brain, resulting in behavior problems and deficits in cognitive performance that may not be reversible with conventional therapy. The presence of cognitive effects is well-established epidemiologically and experimentally, as well as through an understanding of the mechanism of neutoxicity. (Lanphear et al. (2005), Toscano and Guilarte (2005))

In the present study, we do not observe a good measure of cognitive performance. Our data does record schooling, but we leave this outcome aside, since it is both a notoriously poor proxy of cognitive ability, and likely to be strongly affected by any effects of mining on wealth. Instead, we argue that cognitive impacts are very nearly implied by our findings. In particular, we note that lead exposure levels previously reported to have been associated with cognitive impacts are far lower than the high levels associated with decreased Hgb. Indeed, the main thrust of the recent literature has been to demonstrate that there is *no apparent threshold* for cognitive effects, and that brain development is affected from very low levels of exposure. (CDC (2012)) Thus, Lanphear et al. (2005) found "evidence of lead-related intellectual deficits among children who had maximal blood lead levels $\leq 7.5\mu$g/dL," with no evidence of a threshold. As reported above, this is far below PbB levels of 40μg/dL or more previously associated with overt anemia. Hence, demonstrating overt anemia implies a strong likelihood that the affected individuals - and presumably many others with lower PbB - also suffer some cognitive and behavioral impairment.

2.4.3 Effects of lead on child growth

While there is a clear epidemiological link between lead and anemia, epidemiological studies are in less agreement on the nature of the effect of lead on growth in children - and indeed on whether there is any effect. In addition, while there are several known mechanisms of the hematotoxicity of lead, "the mechanism by which lead may reduce a newborn's size is unknown." (Hernandez-Avila et al., 2002, p. 486)

Correlations have been observed - including at moderate PbB on the order of 10μg/dL - between maternal or child blood lead and a wide range of measures of physical growth. These include: low birth weight, BMI at birth, birth length, gestational age, and weight for gestational age; low weight gain in one month-olds; low head circumference at 12 months and in children aged 18-36 months; and low height and weight in children of less than three years of age and children of less than seven years of age. (ATSDR (2007), Bellinger et al. (1991), Hernandez-Avila et al. (2002), Sanın et al. (2001), Zhu et al. (2010)) Interestingly for our discussion of results, Shukla et al. (1989) and Shukla et al. (1991) find growth impacts during the first three years of life only among children consistently exposed to lead, but no persistent effects of exposure *in utero* or in infancy alone. In children whose lead exposure decreased eventually, higher growth was observed, which "allowed them to catch up and even overshoot" the cohort mean. (Shukla et al., 1991, p. 890)

Conversely, however, other studies - as well as some of the same papers that found associations with some index of growth - have variously failed to show correlations with gestational age, birth weight, weight for gestational age, head circumference, and length at birth; and height (at age four, in children of less than 58 months of age, in children aged 8-18 years who had been treated for lead poisoning during childhood, and in a cohort of children followed over 13 years). (ATSDR (2007), Bellinger et al. (1991), Factor-Litvak et al. (1999), Zhu et al. (2010))

In the present paper, we seek to exclude both endogeneity and small-sample bias as potential sources of the ambiguous evidence in epidemiological studies. We are able to show some conclusive evidence of effects of *in utero* exposure on one dimension of growth (height for age) among very young children. However, our results mirror the existing evidence in that we find no clear effects on another key measure of growth (birth weight).

2.4.4 Exposure to lead in utero and in infancy

Our results suggest that in our study sites, health effects are concentrated among the younger children in our sample - infants, and perhaps children in their second year of life. For context, we review the literature on lead exposure at different ages.

As a stylized fact, "blood lead levels [peak] in the age range of 1 to 3 years [...], probably because this period encompasses both the onset of independent ambulation and the time when a child's oral exploration of the environment is greatest." (Bellinger, 2004, p. 1017) However, there is an important earlier path of exposure, through transfer of lead from the mother's body through cord blood and breast milk. Lead crosses the placental barrier with ease: ratios of lead in umbilical

cord blood to lead in the mothers blood can exceed 90% (ATSDR, 2007, p. 172), and correlations as high as 0.8 have been found between maternal and infant PbB. (Lauwerys et al., 1978, p. 280) Indeed, "infants are born with a lead body burden that reflects the burden of the mother." (ATSDR, 2007, p. 223) Exposure of children to lead in their mother's body further continues through breast feeding. Studies have found ratios of lead in breast milk to lead in maternal blood of up to 10%, and one study found that lead from breast milk contributed between 40-80% of infant blood lead. (ATSDR, 2007, p. 172)

Finding health impacts among infants is therefore particularly plausible if there is evidence of maternal lead burdens. An absence of apparent impacts among older children in mining communities could be due to cessation of exposure, or to countervailing beneficial wealth effects.

2.5 Functional consequences of anemia, early childhood stunting, and cognitive effects of high lead exposure

The principal functional consequence of anemia in adults is listlessness and a reduced ability to perform physical work. The productivity loss due to (iron-deficiency) anemia in manual workers has been estimated to be on the order of 5-17%, depending on how much the activities performed rely on heavy physical labor. (Horton and Ross (2003)) In children, a range of randomized control trials of iron supplements to cure anemia has shown that supplementation increases performance on standardized tests (see Horton and Ross (2003) for a review), as well as primary school attendance. (Bobonis et al. (2006))[6] In addition, a possible association of iron deficiency anemia with deficits in early cognitive development has received some attention. A meta-analysis found that a one gram per deciliter increase in Hgb was associated with a 1.73 point increase on the IQ scale, but causality has not been compellingly shown. (Balarajan et al. (2012), McCann and Ames (2007))

To the degree that our results conclusively establish that residents of mining communities are overtly anemic or growth impaired due to lead exposure, direct cognitive damage due to lead is to be expected, and would likely dominate any indirect effect from anemia. As we have noted above, reduced Hgb has previously been observed in respondents with strongly elevated PbB, likely above 40 μg/dL. The threshold for growth effects is not clearly known, though some effects have been demonstrated at levels above 10 μg/dL. Exposure to levels of lead in this range is expected to cause considerable cognitive deficits. Using the dose-response from a reanalysis of seven longitudinal studies by Lanphear et al. (2005), a move from a background lead level of perhaps 2 μg/dL to 20 μg/dL is associated with a loss of about six IQ points (0.29σ, or slightly more than the reported

[6]The intervention in Bobonis et al. (2006) combined iron supplementation with deworming.

inter-generational IQ difference). The dose-response estimate in a recent well-identified study by Rau et al. (2013) yields much larger estimates, with a 2.7 standard deviation expected decrement from the same increase in PbB.[7] Following Horton and Ross (2003), and assuming a correlation of 0.62 between childhood and adult IQ, and a wage decrease of about eight percent associated with a one-standard deviation decrement in adult IQ, we could estimate that children with a PbB of 20 μg/dL could be expected to experience a lasting productivity loss of about 1.6% using the Lanphear et al. (2005) relationship, and 13% using the Rau et al. (2013) relationship.

It is harder to predict the functional consequences of lead toxicity-induced stunting in our study setting. Stunting in infancy and early childhood is a serious condition, associated with "short adult stature, reduced lean body mass, less schooling, diminished intellectual functioning, reduced earnings, and lower birth weight of infants born to women who themselves had been stunted as children" - with often large adverse effects. (Dewey and Begum, 2011, p. 8) However, it is not clear to what extent these findings apply to our context. This is, first, because our results show unambiguous evidence of stunting in infants, but less crisp evidence of stunting in older children in mining areas. Secondly, the impacts of stunting have principally been studied in the context of malnutrition and repeated infection. (Fullerton et al. (2008)) By way of contrast, in our setting, affected households are wealthier on average than the controls, and we find no evidence of more frequent episodes of diarrhea and fever, or general ill health. That said, stunting is a disconcerting indicator of poor development, and even a transitory effect is worrisome, given that "risk factors that interfere with cognitive function are especially important during infancy." (Berkman et al., 2002, p. 564)

3 Data

3.1 Socio-economic and health data

We obtain socio-economic and health data from the Demographic and Health Surveys (DHS) series of surveys, and pool all 104 available geo-coded DHS data from countries for which we have mining data. This yields a dataset of repeated cross-sections covering 44 countries, with a total of 1.2m households, and several million individual records. About 170,000 households are within no more than 20km of a mine recorded in our data, and enter our analysis. (Table 2) Their location

[7] Lanphear et al. (2005) obtain an empirical dose-response function of IQ to PbB that is close to semi-log linear, with a predicted loss of 2.7 IQ points with a log-unit increase in PbB. Rau et al. (2013) find very large effects, with an estimated linear decrease of 0.15-0.21 standard deviations in ability scores with a PbB increase of one μg/dL.

is shown in Figure 1. There is a small but growing literature working with multiple DHS rounds, although we are not aware of any published work that uses all available survey rounds. We extend the use of the data along a number of dimensions, summarized in Appendix A.

The DHS data has some notable strengths: it covers a very large range of developing countries; surveys have been conducted for nearly 30 years; individual surveys are fairly comparable; sampling cluster geocodes are available for many survey rounds; and there is strong data on maternal and under-five health, including anthropometrics and specifically, Hgb. These features currently make DHS virtually the only choice to study global health and development at the micro level.[8]

However, the data also has some important limitations with implications for our work. (i) There is relatively little data on socio-economic status, no information on wages, and little information on employment. We therefore work with an asset index rather than more direct measures of wealth or of income, and discuss employment outcomes only in passing. (ii) Because the surveys have kept changing and improving, very few indicators of interest to us were collected in all surveys. Indeed, working with the largest set of observations for which all indicators are available is impractical, because the number of observations is very small. On the other hand, estimating results on pair-wise common sets would lead to tedious repetition. In our discussion of results, we seek to strike a balance. We present side-by-side comparisons for core results where the sample changes meaningfully. Where there is a strong regional pattern in the sub-sample that moves us away from the ambition of a study of effects across a diverse set of developing countries, we seek to point it out. (iii) Finally, we stress again that the data is cross-sectional. We therefore construct from the pooled cross-section pseudo-panels at a level of commonality in the data (generally, at the level of individual mines). In addition, we deploy a battery of other tests that help bolster causal claims.

3.1.1 Socio-economic data

Our core measure of wealth is a standard asset index computed over household durables and housing characteristics. (Filmer and Pritchett (2001); see Appendix C for details) We do not include slow-moving or immutable traits of the household head, such gender, marital status, or education. An SES index including these characteristics might be of interest for the long-run cross-sectional analysis, but our key interest is in dimensions of wealth that are reasonably sensitive at least to medium-term effects. We compute the wealth index based on the largest set of assets and housing characteristics available within each survey round.

[8]Other data with high coverage that include both health and socio-economics are either much less rich (IPUMS), or much less harmonized (LSMS).

The DHS surveys contain no information on wages, and limited information on employment. In some surveys, women are asked about their employment status and occupation; in a smaller number of surveys, we observe the same indicators for men. Information collected includes questions on unemployment, some questions on seasonality, and occupation. We also collect a dummy variable for whether a household owned any land.

3.1.2 Health data

We obtain from the DHS detailed data on health among children below five years of age, and among women aged 15-49 years. There is little information on older children, men of any age, and women aged 50 years and over. In addition, we extract proxies of perinatal health care recorded for up to five recent births. Finally, we construct infant and under-five mortality data for all children whose births were recorded in any survey module.[9]

Our core health indicators are blood Hgb levels and an age-adjusted height index. Hgb is adjusted for altitude, and expressed either as a continuous measure in standard units of grams of hemoglobin per deciliter of blood (g/dL), or as a binary indicator for the clinical condition of anemia, associated with blood Hgb below 12 g/dL in non-pregnant women and 11 g/dL in pregnant women and children. (World Health Organization (2011))

Following standard practice, height is expressed as the difference between a respondent's height and the age-group median, normalized to standard deviations. We normalize using the median and standard deviation provided by DHS (alternative normalizations make no empirical difference). We consider the continuous height measure, as well as the clinical outcome of stunting (severe stunting), defined as a height of at least two (three) standard deviations below the median.

We also extract a range of variables on other adult and child health outcomes, maternal health care and the use of health care among all adult women, sexual risk taking, and nutrition. Definitions are briefly listed in Appendix B.

[9]Because we construct these variables from birth records of all children ever born to the women in sample, the mortality variables must be interpreted as being conditional on the mother's survival until the time the survey was taken.

3.2 Mining data

We obtain data on the location and characteristics of mines and mineral deposits from four data sources. These include a very large cross-sectional dataset that allows us to make meaningful claims about the mean effect of mining across many developing countries; two panel datasets of mine output that permit us to estimate mine-level panels and strengthen identification; and an additional dataset of mine locations that serves to ensure robustness of our findings to measurement error in the location of mines. In total, we observe communities near 838 mines in the cross-section, and 515 mines in the panel.[10] (Table 3)

3.2.1 Cross-sectional data on mine location

In the cross-section, we work with the United States Geological Survey's Mineral Resource Database. (United States Geological Survey (2005)) It contains the location of a very large set of mines, legacies, deposits, and smelters (about 25,000 locations in total) across developing countries. Coverage is high globally, but especially dense for Latin America. The data contains geological information and some basic description of the nature of the mine for a substantial subset of entries. However, there is no data on production, and start dates and status of operation are only available for very few mines.

In our baseline cross-sectional model, we include among treatment locations all active mines, legacies (that is, former mines that are now dormant), and smelters. We include smelters, because in many of the locations we study, they are an integral part of the mining industrial complex. We include legacies, because the cross-sectional data gives us little guidance in defining whether a mine was operational during a given survey round. The resulting treatment definition should be thought of as yielding 'the effect of living in a location ever exposed to mineral mining or processing'.

We extensively parse information on the types of mineral mined (or present in lesser quantities) in a given location, first, to generate coherent groups of similar mines, and in a second processing step, to sort these into larger (and not mutually exclusive) groups that share the same expected pollutants and hence, the same health effects. We remove from our baseline sample all quarries, gravel pits and sand mines. We do so because we seek to study the welfare impacts of mining as an industry that generates very high value added, but is potentially severely polluting. Quarries differ

[10] Nearly all of those mines enter into our model when we use state-level effects (see Part 4). The number of mines near which we observe at least one community within 5km (treatment) and one within 5-20km (control) is lower, with 226 mines in the cross-section, and 175 in the panel. These are the mines that enter into our mine-effects models.

from mineral mines in both respects, at least as a matter of degrees. As we have argued above, we are particularly interested in polymetallic mines near which we expect pollution with heavy metals, and particularly with lead. For the purposes of the present paper, we define a mine to be in this category if (i) lead is being mined or (if not targeted for extraction) known to be present in significant amounts, or (ii) any two of the metals copper, gold, silver, and zinc are being mined. This definition is necessarily imprecise, but gives due recognition to the special role of lead, and seeks to exclude metal mines with different pollutant characteristics, such as alluvial gold deposits.

3.2.2 Mine-level production data

Since the USGS data provides virtually no time variation, we draw additional information from two business intelligence firms; IntierraRMG (IntierraRMG (2013) - for whose product we henceforth write 'RMD', for 'Raw Materials Data') and Infomine (Infomine (2013)).Both sources record dates of operation and production information, alongside diverse additional characteristics of the mines. Most mines included in the Infomine data are also available in the RMD data, but not vice versa. To conserve consistency as far as possible, we therefore work with RMD as our basic data, and add those Infomine entries that are not also contained in the RMD data. RMD mines are more homogenous than those in the USGS sample: most of them are large mines, and most of those close to DHS clusters are metal mines. While the set of mines included is far smaller than for the USGS data, coverage of large mines is very comprehensive, and the mines recorded in the dataset account for a very large share of global production of most metals since 1984. (For instance, they account for around 80% of global gold production and 80-90% of global iron ore production in the most recent decade for which data is available.)

Because there is some question as to the precision of geolocations recorded in the RMD data, we gather mine geolocations from an additional dataset, Mining Atlas, for all RMD mines recorded in this additional data source. In the current version of this paper, we use information from Mining Atlas for three purposes: (i) to add geolocations for RMD mines wherever location is missing in the original data; (ii) to identify and discard a small number of mines where location is plainly not recorded with any precision in either dataset; and (iii) to check robustness of our results to measurement error in geolocation. (See details in Appendix K.)

3.2.3 Other data

In the final section of this paper, we study policy correlates of treatment effects. We obtain country-level data on GDP, governance, and conflict from the World Development Indicators;

data on the efforts a given country made toward compliance with the Extractive Industries Transparency Initiative from the Initiative's website[11]; and state-level data on governance, geography, and macroeconomic performance from Gennaioli et al. (2013).

4 Econometric Specification

4.1 Baseline treatment definition

In studying both economic and health impacts of mining, we define exposure to mining as being geographically close to a mine in the cross-section, and as closeness interacted with the mine being active in the panel. This choice is immediate for the study of economic impacts: with transport and search cost, distance is the treatment of interest. For the purpose of studying health impacts, distance acts as a proxy for the actual treatment of interest, pollution - which we do not observe. We measure distance as the geographic distance between a cluster and the mine that is nearest to it. In our baseline model, we define a cluster as being 'close', and hence, 'treated', when it is within five kilometers of the nearest mine. We will also refer to this as the 'direct vicinity' of the mine. We define a cluster as being in the control group when it is within 5-20km of the nearest mine. We will refer to this as the 'general vicinity' of the mine. We do not use data from clusters more distant than 20km from the nearest mine.

As we noted above, we bound our treatment group tightly, to enable us to detect health impacts within the region in which pollution is likely to occur. Bounding our control group conservatively helps identification. One need only consult maps of mining areas to confirm that over distances as substantial as 40km or even 200km - as used elsewhere - many things other than closeness to mines change, whether in the natural and the built environment, or in institutions. Because we nearly exclusively study health impacts on women and children, we argue that impacts even within the tight geographic bound of our definition are likely general population impacts, rather than the effects of occupational exposure. We show below that, empirically, wealth effects dissipate within 20km at most, and far more rapidly in the panel. For the purpose of studying health impacts, this is welcome news: we are comparing the treatment group to a control group that is in a meaningful sense within the more general vicinity of a mine.[12]

[11] See www.eiti.org. The Extractive Industries Transparency Initiative (EITI) describes itself as "a global coalition of governments, companies and civil society working together to improve openness and accountable management of revenues from natural resources.'

[12] For the study of wealth benefits alone, one might have decided to compare those living within 20km of a mine to those living farther afield. A loser definition of closeness would help balance the panel, but also weaken identification.

In the panel, we define mining activity as a dummy variable taking value one when the mine had non-zero output, and value zero when the mine was known to have had zero output. (We take a conservative approach in that we never impute activity, not even indirectly by simply contrasting observations before and after an opening date. - See Appendix A.) That is, in the present version of the paper, we consider only extensive margin impacts of production. We do so because intensive margin year-on-year variation in output is likely more weakly associated with health outcomes. In this, mines differ from sources of pollution studied elsewhere. Extracting minerals from the ground, breaking them up, and processing them generates a flow of pollution. At the same time however, the stock of tailings dumped after processing will in many cases continue to generate a pollution flow that correlates more with cumulative production (and due to weathering, continues even as production ceases). The exact time pattern of pollution is thus hard to predict, but is bound to lie somewhere between a pure flow and a pure stock problem. We therefore argue that an extensive margin definition is more appropriate than either an intensive measure of activity, cumulative production, or the cross-sectional 'once-on-always-on' measure.

4.2 Cross-sectional models

Our cross-sectional results reflect the long-run impact of mining. We include restrictive group effects allow for unobserved local characteristics. As is well known, preferential sorting is likely to undermine the identification of individual effects. In the cross-section, we are not able to decisively address this issue, although we do explore its importance to the degree possible by studying the population of never-movers. Because of the possibility of sorting, the correct way to read our cross-sectional results is therefore to view them as the long-run effect of mining on 'mining *communities*' (much as a district or county-level study estimates effects on those units). As such, the effect is identified; and to the degree that regional disparities matter, it is of policy interest. Our difference-in-differences setup will then allow us to ascertain that impacts are not driven by sorting, and to make stronger claims about the well-being of 'people living in mining communities'.

4.2.1 Group effects

Our baseline cross-sectional model bases identification on demanding group effects. In the preferred specification, we consider outcomes y for individuals or households i in sampling cluster j within no more than 20km a mine, conditional on whether the cluster is *close* (within 5km)

Perhaps the most natural alternative, therefore, would be to move away from defining treatment as distance to nearest mine, and instead study effects of mine density in (hopefully quite balanced) panels of administrative units. This would, however, vitiate the purpose of studying health effects.

to a mine, and conditional on other covariates X. Because distance is measured between mines and sampling clusters, the treatment varies at the cluster level, not the individual level. Covariates always include an indicator for whether the cluster is in an urban or rural setting, and some appropriate measure of the age of the respondent, the respondent's mother, or the household head. Because DHS conducts repeated cross-sections, our model allows for repeated measurements of effects near the same mine, while accounting for year-specific effects in each round of measurements. We therefore use common effects γ for all observations near the same mine surveyed in the same year (mine-year effects), and account for residual correlations by clustering error terms at the mine level (not the mine-year level). We compare results to those from a model with less demanding state-year effects (with the obvious modification in Equation 1).[13] As is evident, wherever the outcome of interest is binary, we model it using a linear probability model.

$$y_i = \beta_1 close_j + \beta_2 X_i + \gamma_{mine-year} + \epsilon_i \tag{1}$$

Threats to identification would arise if mining towns differed from neighboring communities in geography, institutions or other characteristics in ways that correlate with potential outcomes. However, for identification to be undermined, they would have to do so even compared to locations very close by, because we choose distance cut-offs conservatively. Identification is also only affected by such differences if they are not in some way due to the presence of the mine in long-run equilibrium. For instance, infrastructure is likely to be affected by mining. Companies may build railroads and roads; mining consumes water, and might make mining communities water-scarce. Similarly, mining might have impacts - whether good or bad - on local governance and public goods. In turn, infrastructure, governance, and public goods likely affect outcomes. Yet, this does not affect the validity of our reduced-form estimates. Rather, it requires us to show further evidence before we make claims about the mechanisms - crucially, pollution - by which mines affect local communities.

4.2.2 IV-group effects

To address any concerns over endogenous choice of mining locations, we use the location and characteristics of mineral *deposits* to instrument for the presence of mines.[14] Work is ongoing to improve efficiency of the IV estimates; in the present version of the paper, we therefore discuss preliminary results as a robustness check on our cross-sectional group effects estimates. We use

[13] As we discuss in Appendix A, our estimates are always only based on observations within 20km of a mine, including in the state-year model. The latter simply allows for less restrictive group effects than the mine-year model.

[14] This is similar in spirit to the geographic instruments pioneered by Duflo and Pande (2007).

deposit information in two simple specifications to instrument for the treatment *close*, recording whether cluster j is within 5km of a mine. (Equation 2) First, we deploy a dummy Wald instrument that simply indicates whether there is a mineral deposit within 5km of a given cluster. In this case, *deposit* is scalar, despite the vector notation in Equation 2. In our second specification, *deposit* is a vector of dummy variables recording whether there is a deposit with a certain geological characteristic within 5km of the cluster. In each case, the sample is restricted to clusters within no more than 20km of any deposit. (See Appendix A for details on implementation.)

$$\begin{aligned} y_i &= \beta_1 close_j + \beta_2 X_i + \gamma_{state-year} + \epsilon_i \\ close_j &= \phi' deposit_j + \delta_{state-year} + \eta_j \end{aligned} \tag{2}$$

Because there can be no mine without a mineral deposit, there are neither 'defiers' nor 'always-takers' in our setup, and we can interpret IV estimates as the effect of treatment on the treated (ETTs). Unsurprisingly, the instrument is exceedingly strong. (See Appendix D.) Since the true global distribution of mineral deposits is clearly exogenous, the instrument is also exogenous, as long as the assumption holds that there is no preferential *prospecting* for minerals. We believe this is very likely the case, given (i) that most of the countries in our sample were surveyed long ago, in colonial times, and given (ii) that there is little indication that either geographic or political obstacles keep mining companies from pursuing promising deposits. We also believe that the instrument is excludable. The most likely violations would be due to topographical features such as land quality, gradient, or water availability. But because we so strongly restrict our analysis in space, the exclusion restriction is only violated if these vary systematically between locations within 5km of a deposit and those within 5-20km of a deposit. For the group instrument, the restriction is even easier to satisfy: it requires only that the richness of deposits, as defined by geological characteristics, not be correlated with outcomes other than through mining.

The group instrument is hence attractive in that it requires weaker identifying assumptions. Regrettably, however, rich deposit information (e.g. more detailed information on geologic characteristics) is only available in the USGS data for a relatively small sub-set of mines. In the present version of the paper, we present only results obtained from the most widely available information, namely, data on which minerals are present at a deposit. This does not mean that the group Wald instrument compares "gold deposits to copper deposits", etc. Rather, the data records complex combinations of metals that tend to correlate with other geologic characteristics. Our group Wald estimates pertain to this subsample only. By way of contrast, coverage of deposit *locations* (without further details on characteristics) in the data is comprehensive. For the dummy Wald instrument, we therefore have good reason to believe that our IV yields a sound estimate of the general population effect. The Wald instrument also has the advantage of supreme tractability:

since implementation requires literally no judgment calls that might be sensitive, it is highly robust.

We note that our cross-sectional instrumental variables strategy can very naturally be extended to the panel setting by interacting the presence of mineral deposits with world minerals prices. However, we empirically find that the interaction of deposits with price works well for instrumenting for the intensive margin, but is not clearly a strong instrument for the extensive margin. We therefore leave use of the time-varying instrument for further work.

4.3 Difference in difference estimators

4.3.1 Pseudo-panels

Because of the highly conservative way in which we define the cross-sectional control group and impose group effects, and because of the attractive characteristics of our instrument, we believe our cross-sectional setup offers well-identified estimates of the long-run impact of mining on communities. Still, because of the possibility of sorting, it does not allow us to make claims about the impact of mining on individuals. Because of the long-run nature of estimates, it also says less than is desirable about the mechanisms through which mining affects well-being.

An immediate way of addressing both challenges is to construct pseudo-panels from the repeated cross-sectional DHS surveys. We construct these in two ways. Firstly, we compare observations from households surveyed at different times, but near the same mine. Secondly, we compare children born to the same mother at different times. Plainly, since the data consists of repeated cross-sections, comparisons in each case are across different individuals.

We then analyze outcomes for individuals i in cluster j at time $t - \sigma$. We model outcomes as being conditional on whether the respondent lived in a community *close* to a mine at time $t - \tau$, and whether the mine was *operating* at time $t - \tau$.[15] The survey year is t, and the indices τ and σ are chosen, respectively, to refer to the time periods of interest for treatment and for outcomes. Both depend on the outcome being investigated, and are either the interview year (τ or σ = 0), the birth year (τ or σ = age), or the year a child was *in utero* (τ or σ = age + 1). (Obviously, $\sigma \leq \tau$) We allow for time-invariant effects γ either for each mine, or for each mother (in which case

[15]For each respondent in our sample, we only observe current residence, and how long the household has been resident. We have no information on previous residence. Therefore, the panel is inherently restricted to respondents who have lived in the location where they were surveyed at least for τ years. (Although they may have moved to their present location at a time before $t - \tau$.)

γ_{mom} replaces γ_{mine} in Equation 3), and for time-specific effects $f(t)$. (Notice that in the mother-level panel, while the DiD effect of interest is well identified, the coefficient on *close* cannot be estimated, because we do not observe households moving.) We believe country-year dummies and country linear trends are sufficiently flexible and appropriate for sample size in the mine-level and mom-level model, respectively. We use these as our baseline models, and show robustness to using different time effects.

$$y_{i(t-\sigma)} = \beta_1 close_j + \beta_2 operating_{j(t-\tau)} + \beta_3 close_j * operating_{j(t-\tau)} + \beta_4 X_i + \gamma_{mine} + f(t) + \epsilon_{i(t-\sigma)} \tag{3}$$

The mother-level pseudo-panel has the virtue of controlling more tightly for unobservable characteristics than the mine-level panel. Perhaps most crucially, it is virtually immune to sorting, barring complex sorting behavior on changes in expected outcomes over time. However, the number of mothers for whom we observe birth outcomes both when the mine is operational and when it is not operational is very small for many indicators of interest. While mom-level estimates are thus most impervious to identification challenges, they are vulnerable to small-sample issues, and representative only of a sub-sample of our population. Mine-level estimates rest on larger samples, but require somewhat stronger sorting assumptions, namely, that sorting is sufficiently slow relative to the frequency at which the outcomes of interest are measured.

We react to this trade-off in two ways. Firstly, we focus throughout on estimates where the mom-level results are at least consistent with mine-level results. Secondly, we look for DiD tests other than the pseudo-panel that are offer similar reassurance against sorting and other violations of the identification assumptions as the mom-level panel, while conserving more sample size. We describe these in the following section.

4.3.2 Other DiD tests

For some indicators, our sample is relatively small near mines where there is production information, so that pseudo-panel estimates encounter power problems, in particular when using mother-level fixed effects. We therefore leverage the scientific understanding of the health conditions of interest to our study to construct additional difference in differences tests that can be implemented without the use of time-varying production data. All of these tests are similar in spirit, but build on different insights into the likely nature of exposure and the organism's reaction to it.

Mine types: Firstly, we make use of the fact that - as discussed above - distinct mine types are associated with specific pollutants and health effects. This allows us to contrast differences across distance groups near mines where an effect is expected, and near mines where no effect is expected. (The effect of *mine type* alone is collinear with mine-year effects.)

$$y_i = \beta_1 close_j + \beta_2 minetype_j + \beta_3 close_j * minetype_j + \beta_4 X_i + \gamma_{mine-year} + \epsilon_i \quad (4)$$

Identification rests on the assumption that potential health outcomes vary among those close and not close to the mine in similar ways near mines of different types. With respect to sorting, the assumption would be violated only if respondents were aware of how mine types differ in health outcomes, and sort accordingly. This is of course conceivable. We address the problem in three ways. Firstly, we define mine types narrowly, in ways that are plausibly hard to observe. Secondly, we compare DiD results on health to those on wealth, and show that differences arise for health outcomes, but not wealth. Thirdly, we show that there are DiD results only for specific expected health outcomes, not general health.

Timing of birth relative to migration: Secondly, we use the fact that respondents who have long been resident in mining communities are likely to have been exposed to more pollution. Specifically, we leverage differences in exposure *in utero*, and compare among children i born to migrants before their family moved to mining communities j, and those *conceived after move* (and contrast this with the same statistic observed among those who migrated to locations slightly farther away from the mine). (See Appendix A.)

$$\begin{aligned} y_i = \beta_1 close_j + \beta_2 conceived\ after\ move_i + \beta_3 close_j * conceived\ after\ move_i \\ + \beta_4 X_i + \gamma_{mine-year} + \epsilon_i \end{aligned} \quad (5)$$

The identifying assumption is hence that potential health outcomes do not vary systematically with the exact timing of birth relative to the move, comparing those who moved to mining communities to those who moved to communities farther away. It would be violated by sorting if, for instance, pregnant women are more likely to hold off on moving when they are about to move to mining communities, perhaps because they are weary of pollution, or conversely, if pregnant women are more likely to speed up relocation, perhaps because they hope for good economic opportunities. We estimate the model both with mine-year and mother fixed effects.

Maternal Hgb recovery: Finally, we develop a DiD test based on the observation that in lead-exposed adults, the recovery of Hgb after blood loss is even more readily affected than the steady-state level of Hgb. As discussed above, this result was previously proven by studying Hgb recovery after donating blood. Of course, we cannot identify blood donors in our sample. We do, however, observe one population group that experiences dramatic drops in Hgb: women who are pregnant, or have recently given birth. This allows us to formulate a test that asks whether differences in Hgb between women i in mining and control communities j are particularly affected during pregnancy and postpartum. In our preferred specification, we estimate the model with state-year effects, since the number of women we observe within the time period of interest is borderline too small for allowing for mine-year effects. (We discuss identifying assumptions and extensive robustness checks below, in Section 5.3.)

$$y_i = \beta_1 close_j + \beta_2 pregnant\ or\ postpartum_i + \beta_3 close_j * pregnant\ or\ postpartum_i + \beta_4 X_i + \gamma_{state-year} + \epsilon_i \tag{6}$$

5 Results

5.1 Effects on wealth

In the long run, mining towns are wealthier than neighboring communities
Households in mining communities are at the mean considerably wealthier in terms of asset ownership than those living farther afield, allowing for unobserved common wealth effects specific to each mine in a given year.

The magnitude of the effect at the global average is on the order of 0.11 standard deviations of the asset index. (Table 4) This is a considerable effect, given that in the countries in our sample, there is generally great within-country variation in asset ownership. In the linear index, the magnitude of the effect is comparable to that of owning a car or motorbike in the case of Peru in the year 2000, and to the effect of owning a radio or a watch in the case of Burkina Faso in the year 2010. (See Appendix C for a description of the index and for examples of factor loadings.) We argue below that, because of the spatial pattern of wealth effects, the estimate should be interpreted as a lower bound on the long-run effect of mining on community wealth. In Appendix E, we show that our unweighted baseline estimates are considerably smaller than estimates obtained by (i) weighting each mine equally, or (ii) weighting by estimates of the mine-year population.

Further, as a robustness check, IV estimates of the wealth effect give little indication that endogeneity might be driving results in the group-effects model. (Appendix D) The dummy IV estimate

is not significant (t = 1.35), but the estimated effect is close to the FE benchmark, with a reasonably inflated IV standard error. (It is also not statistically different from the effect obtained from our baseline specification.) The group-Wald result shows a significant and positive effect estimate above the FE benchmark, though again not significantly different.[16]

We have argued that, if the object of interest is the effect of mining on household welfare, rather than on communities, the most salient concern about identification in the cross-section is preferential sorting. Our panel setup will allow us to decisively deal with this issue. To gather preliminary information, we analyze cross-sectional results separately for households that report never having moved from their current location. (Our data does not record previous place of residence, so we cannot assess out-migration.) We find little difference between never-movers and the population when we allow for state-year effects, but find considerably (if not significantly) smaller and weaker wealth effects for never-movers when allowing for mine-year effects. We interpret this as weak evidence of sorting of migrants with better potential socio-economic outcomes into mining communities, or sorting of previous residents with better potential outcomes out of mining communities.[17]

Wealth results in the panel confirm the cross-sectional pattern

Pseudo-panel results confirm that mining activity is associated with higher wealth in communities in the vicinity. (Table 5) The DiD coefficient on the effect of living close to a mine in a year when it is operating is 0.26 standard deviations of the asset index in our preferred specification. While the estimates vary with different models, they are always large, positive, and significant, whether we work with mine-level or state-level pseudo-panels, and are very robust to different non-parametric time controls. In terms of the linear asset index, the effect size is comparable to the impact of having an electricity connection or living in a dwelling with finished flooring in the case of Peru in the year 2000, and to the effect of owning a motorbike or mobile phone in the case of Burkina Faso, in the year 2010. Panel results on asset wealth can be presumed to be more robust to sorting than our cross-sectional results. However, because survey rounds are generally about five years apart, there is still the possibility that rapid sorting could influence results.

By way of contrast to the DiD result, the single difference coefficients on whether a mine is operating are highly sensitive to the non-parametric controls used. That is, there is no stable effect

[16] As is evident from both FE and IV estimates, the wealth effect is more dramatic for the sub-sample for which the kind of deposit information used in the group-Wald instrument is available.

[17] For background, we note that there is only weakly more migration in mining communities than in neighboring communities. However, in both mining and neighboring communities, the share of migrant households is very high (around 60% of households migrated at some time, and about 23% migrated within the five years preceding the survey). Sorting could therefore easily explain cross-sectional differences, if the characteristics of migrants (including those unobserved households who left the communities) are sufficiently different.

of mine operations beyond the immediate vicinity of the mine. The single difference coefficients comparing areas close to mines to areas somewhat farther away tend to be negative, and not significant. Thus, communities within the direct vicinity of mines were wealthier than those in the more general vicinity only when the mine was operating. This is of course reassuring for the validity of our cross-sectional analysis. We also note that the combined effect of living close to a mine during a year of operation is in the range of 0.1-0.15 standard deviations of the asset wealth index. This is of the same order of magnitude as the cross-sectional effect; a very encouraging finding for the external validity of our results, given that the two estimates come from entirely independent mine data sets. At the same time, the panel effect is slightly larger, as is intuitive for a medium term effect on the spatial distribution of wealth (recall that panel observations tend to be about five years apart), as compared to the long-run effect.

Spatial extent of the wealth effect

Wealth effects decay steeply with distance to the nearest mine; communities beyond a distance of about 15-20km show no effects in either the cross-section or the panel. In the panel, effects are limited essentially to those living with 5km or at most 10km. (Figures 2 and 3) The difference in spatial patterns between the cross-section and the panel allows for a number of possible explanations. If both patterns are well-identified, one would argue that the discrepancy reflects the contrast between short-run and long-run impacts, with further diffusion of wealth effects over time. If we were not convinced of identification in the cross-section, we might feel that the pattern suggests that mines tend to locate in places that are already wealthier than their surroundings - but this is not consistent with the absence of a panel effect on the un-interacted dummy for closeness. Of course, panel results come from a smaller sample, and simply might be more attenuated.

We have argued that, by comparing those living within 5km of a mine to those within 5-20km, we estimate the effect of living within the direct vicinity of a mine, compared to living within the more general vicinity. Figure 3 shows that this is the only first-order treatment effect in the panel. Yet, patterns in the cross-section are more subtle. As Figure 2 illustrates, the treatment effect in our definition is smaller than the long-run wealth effect on the direct vicinity of mines, as compared to those living more than 20km away from the mine. The latter is a multiple of our baseline estimate, on the order of 0.4 standard deviations of the asset index. The baseline estimate is, however, larger than the average effect of living within the general vicinity of the mine, as opposed to living more than 20km from the mine (0.05σ - results not shown).

Even in the cross-section, the estimated spatial extent of treatment effects is smaller than in the case study analyzed in Aragón and Rud (2013b), who find "positive and significant [income effects] for households located within 100km of Cajamarca city," the community closest to the

mine studied. The discrepancy could be due to the fact that the paper studies a policy change that can be presumed to be very favorable for local welfare; or the fact that it considers the case of a very large mine in a region with reportedly high transport cost. In addition, Aragón and Rud (2013b) have income data available; presumably, a more sensitive measure of well-being than our asset index. For Ghana, Aragón and Rud (2013a) find (adverse) effects "within 20km of mines" only, and Kotsadam and Tolonen (2013) find sectoral employment shifts in a 20km region; both are roughly in keeping with our definition of the general vicinity of a mine, though not with our treatment definition.

Despite evidence of cross-sectional effects beyond 5km, we maintain as our basic specification our dummy model using 5km as the cutoff for closeness to a mine. We make this choice because (i) it emphasizes tight identification, especially given the steep decay of the panel effect, and because (ii) it serves best our main purpose of analyzing the trade-off between wealth and geographically limited health impacts. For the purpose of the joint investigation of health and wealth, we read the spatial pattern in wealth impacts to confirm that our choice of communities within 5-20km of a mine as a control group corresponds reasonably well to a meaningful definition of being in the general vicinity of a mine, where communities are economically affected to some degree.

Other dimensions of welfare - health care investment and employment

Beyond asset wealth, we analyze impacts of living in the vicinity of a mine on two additional dimensions of welfare, namely health care and employment. We briefly summarize prominent patterns here; detailed results are shown in Appendix F.

There is strong evidence to suggest that in the long run, higher asset wealth in mining communities goes hand-in-hand with higher insurance coverage, and with a larger share of women giving birth with some level of skilled assistance. Panel results suggest that benefits may extend beyond the immediate vicinity of the mine. This is an intuitive finding, and suggests that access to public goods generated by the mine extends further than direct wealth benefits.

To the degree that our interest is in contrasting wealth and health impacts, we are reassured that residents of mining communities are no worse off in terms of health care than those living farther afield, and perhaps weakly better off. We have therefore no reason to suspect that residents of mining communities should experience adverse health outcomes because they are less well cared for. The one potential exception to this pattern is that in our mother-level panel, we find significant decreases in the share of women who gave birth in an improved setting in mining communities when the mine was operational. The cross-sectional and mine-level panel evidence contradicts this finding. However, we mention it here because it is at odds with our otherwise consistent evidence

on wealth. We return to it explicitly when we discuss the recovery of maternal Hgb levels, below, and show that there is no evidence that our results might be contaminated by differences in maternal health care.

Regrettably, as mentioned, the DHS surveys contain no information on wages, and have limited coverage of employment. This prevents us from assessing whether the positive income effects demonstrated by Aragón and Rud (2013b) for the context of high transport cost in the Peruvian Andes hold more generally. In the cross-section, unemployment among men is virtually unaffected, as is consistent with long-run general equilibrium. As is intuitive, the sectoral share of agriculture decreases alongside ownership of agricultural land (five percentage points); there is weak evidence of a counterbalancing increase in manual wage work outside of the agriculture sector. Unfortunately, due to the small sample of men for whom employment data was collected (nearly exclusively limited to mining towns in Ghana and Zimbabwe), it is hard to validate the cross-sectional results in the panel. We note that employment effects are nearly always negative in sign - consistent with cueing - but caution that the estimates are noisy and not stable.[18]

5.2 General health

In the following sections, we will demonstrate that residents of mining communities suffer from specific health impacts known to be associated with heavy metal pollution. Before we discuss these patterns, we show here that there is no systematic evidence of poor general health in mining communities. (We also show in Appendix F that we find no indication of greater sexual risk taking or substance abuse, consistent with Wilson (2012).)

For the cross-section, Table 6 shows that only in one instance (diarrhea in children under five years of age) is there a significant adverse health effect on children or adults among the general population living near any mine.[19] Appendix G shows similarly sparse patterns among never-movers, and for differential effects near heavy metal mines. We also observe that there is never a differential general health impact of living near any mine or a heavy metal mine on infants - an important falsification finding to which we return below. On the contrary, in the cross-section, infant and under-five mortality rates decrease weakly among all households, and significantly and strongly among never-movers. This result is of course consistent with greater wealth in mining

[18] We refer the reader to Kotsadam and Tolonen (2013) for a detailed discussion of effects on women and sectoral shifts in sub-Saharan Africa.

[19] We test for differences in all available health indicators - as noted, the DHS surveys contain much less data on adult health than on child health, and the selection of adult health outcomes is slim. We show no results for men's health outcomes, since the samples are very small and the results correspondingly uninstructive.

communities, and indeed, spatial patterns in mortality rates virtually mirror patterns in wealth. (Figure 4)

Panel results show no significant adverse health impacts at all, whether on adults or children. (Table 8. See Appendix H for additional model specifications.) The estimated effects are beneficial in the one instance in which coefficients are significant (reduced diarrhea among children in the mom-level model), and the few other instances when they are borderline significant. Panel results do not show a significant decrease in mortality, although the sign on the DiD coefficient is negative.[20]

General health results are of some interest in themselves - it is perhaps instructive to know that residents of mining communities at the global average are not indiscriminately less healthy than their peers. However, they must be read with some caution. In the context of our quasi-global study, the evidence of specific health impacts we show below is compelling because it confirms predictions from theory. Where there is no clear theory, the absence of mean effects must not be read to have any predictive power. Rather, we present these results here first and foremost as a falsification test. The absence of such effects will serve as prima facie evidence to make plausible that specific health impacts are due to pollution, not other mechanisms.

5.3 Evidence of hematotoxic effects

We have argued above that exposure to lead among residents of mining communities may affect the hematopoietic system and reduce red blood cell survival. In the DHS data, we observe only a single indicator of potential hematotoxicity - blood Hgb concentrations. As argued in Section 2.4, we would expect most strongly to see a reduced ability to recover from blood loss in adults, alongside perhaps depressed Hgb levels. In children, we might expect to see reduced blood Hgb levels, though in the age group we observe, children are likely able to compensate for lead exposure. Our results confirm this expectation: we find strong evidence of lower Hgb levels and slower Hgb recovery after blood loss in adult women, and limited weaker evidence of lower Hgb levels in children.

[20] We do not further investigate whether the difference between the cross-section and the panel is due to sorting or to the pace at which effects unfold; for our purposes, it is sufficient to establish here that there is certainly no evidence that mining communities are worse off in terms of mortality than those farther afield, and some evidence to the contrary.

Hemoglobin levels in adult women are strongly depressed in mining communities
In the cross-section, blood Hgb levels are depressed among women living in mining communities by about 0.09 g/dL. The effect among never-movers is larger (0.13 g/dL), consistent with longer exposure to environmental lead, although (on this smaller sub-sample) it is just below significance (t = 1.56). Considering directly the clinical outcome of anemia, we find that prevalence is significantly elevated by three percentage points among all households, and by five percentage points among never-movers. Estimates in the state-year model are attenuated.(Table 8)

Panel results confirm these patterns. Point estimates are larger, with DiD coefficients of a 0.33 g/dL decrease in blood Hgb, and nine percentage point increase in the incidence of anemia in our preferred specification. (Table 9) A number of causes could account for the larger point estimate in the panel; notably, the share of metal mines associated with lead pollution is higher in the panel sample (and, as we show below, the treatment effect is concentrated near such mines). In the long-run, there might also be more adaptation to avoid pollution.

The size of the effect can be compared, for instance, to changes in Hgb on the order of 1g/dL associated with treating anemic pregnant women with a course of iron supplementation. (Sloan et al. (2002)) That is, we obtain a general population effect estimate on the order of one-tenth to one-third of the effect of a targeted intervention in a highly susceptible population. Another point of comparison is the drop in blood Hgb during pregnancy and the first year post-partum, estimated in our sample to be on the order of 0.44 g/dL (compared to women who gave birth two or three years ago, and among women living at least 20km away from a mine). The increase in the incidence of anemia is a large effect in absolute terms, but must be seen in the context of a baseline proportion of anemic women of 35-37% in control locations. That is, the cross-sectional effect amounts to an 8-9% relative increase in incidence, and the panel effect, to a 24-26% relative increase.

We note that the single difference coefficient in distance suggests that when the mine is not operational, residents of mining communities have higher Hgb levels than the control group. This is perhaps a surprising finding, given that our wealth results showed a zero or weak negative effect in mining communities when the mine is not operational. (Table 5) However, it further reassures us against any concerns that geographic features, for instance altitude, might be driving the results. It might be an additional factor in the difference between the cross-sectional and panel effect sizes.

We adduce two additional tests, both to further bolster identification, and to help establish that pollution, rather than other possible causes, is the likely cause of depressed blood hemoglobin. (i) Firstly, we show that Hgb effects are only observed near mines where the combination of minerals mined suggests that lead contamination is likely to be present, and show that these mines do not

differ from others in the wealth and general health of residents. (ii) Secondly, we provide direct evidence of reduced ability to recover Hgb after blood loss, with less of pronounced differences in levels - an effect that is hard to reconcile with any cause other than lead toxicity.

We observe effects on hemoglobin levels only near mines where we expect heavy metal pollution

We first show cross-sectional difference-in-difference results comparing the effect of living close to a mine of a type where contamination with hematotoxic metals can be expected to living close to other mines.[21] As noted (in Section 3.2.1), our definition of such 'heavy metal' mines includes lead mines, and polymetallic mines that are often associated with the release of lead and other toxic metals. We also include all metals processing sites, because the pollutants emitted are similar (with an additional burden of SO_2 emissions). We expect this definition to be a meaningful but far from perfect proxy of the probability that heavy metals are released into the environment, and assume that DiD estimates will be attenuated from the true difference between heavy metal and non-heavy metal mines.

DiD results show that effects on Hgb levels are statistically zero (and mildly negative) in women living near mines where there is less reason to expect heavy metal contamination. (Table 10) However, in mines where there is a high likelihood of such contamination, Hgb levels are strongly and significantly depressed - by about 0.22 g/dL relative to women living farther away from the same mines, and by 0.19 g/dL compared to women living near non-heavy metal mines. Correspondingly, the incidence of anemia is five percentage points higher compared to women living near non-heavy metal mines (compared to women living further away from the same mines, it is six percentage points higher). The size of the cross-sectional effect near heavy metal mines is far closer to the panel effect than the average effect in the cross-section (as alluded to above).

The DiD effect is robust to including interactions of the treatment dummy with region indicators (hence allaying any concerns over geographical clustering of heavy metal mines), as well as to including an interaction of the treatment with a pregnancy dummy. We note that there is a significant negative effect of living near *any* mine in Latin America (the base category for the region interaction), perhaps due to the imperfect nature of our definition of heavy metal mines. The effect near any mine is statistically zero for the other regions. We then estimate the same model for the asset index, and confirm that there is no differential wealth impact of living close to a heavy metal mine, as opposed to any mine. We also show that we do not observe similar differential effects of living near a mine associated with heavy metal contamination on two additional indicators of ill

[21] A similar test is hard to construct for the panel, since mines that are potentially associated with heavy metal contamination make up a very large part of the sample.

health among women, namely miscarriage, and grave sickness (as discussed above, there are also no differential effects on children's health).

The trajectory of maternal Hgb recovery after birth in mining communities corresponds with known pathophysiological patterns

The left panel in Figure 5 shows the pattern of recovery from blood loss during pregnancy and delivery among women living close to heavy metal mines, and those living in adjacent areas. Hgb levels conspicuously diverge during pregnancy, and stay apart during the first one and one-half years of the child's life. However, thereafter, they converge to an apparent noise pattern about a common mean. (The right panel shows the same data, with effects smoothed out for the nine months from conception to birth, and each year of the newborn's life, thereafter.) The pattern is characteristic of a pollution-induced decreased ability to recover Hgb after blood loss, as described in Grandjean et al. (1989) and discussed above, in Section 2.4.1, but not of other causes of anemia.

While the pattern is visually striking, given limited sample size, it is too strong a test to assess the difference between coefficients for the two distance groups in each individual trimester. Instead, we test for the difference in differences between the groups across two time periods: pregnancy and the first year of the infant's life (when there is the clear impression of divergence), and the second and third years of the child's life (when there is not). The results presented in Table 11 show that the DiD coefficient is negative, large (0.26 g/dL), and significant. That is, the difference in Hgb levels among women exposed to mining and other women is far greater during and after blood loss due to pregnancy and delivery, than after some time has passed since delivery. The single difference in distance is negative, consistent with the cross-sectional effect estimate. It varies in between one-half the size and the full size of the cross-sectional estimate, but is far from significant on the smaller sub-sample of women who are pregnant or have given birth recently. As expected, Hgb is dramatically lower in all women during pregnancy and in the first year post-partum

The pattern is similar when we estimate the model with mine-level group effects, as shown in Column (2); because of the small sample size and strong identification from the DiD setup, we prefer the state-level model. In our baseline model, we consider a postpartum period of three years. This seems more appropriate than shorter periods because the detailed time pattern of Hgb recovery shown in Figure 5 suggests that differences even out only in the second year of the child's life. It seems more appropriate than longer periods because extending the window too far weakens identification if birth spacing relates to wealth and potential health outcomes. Columns (3-5) show that the effects are robust to extending the post-partum control period to four or five years; they are directionally consistent but insignificant when we shorten it to just two years.

Alternative explanations for the pattern of Hgb recovery after delivery are harder to come by than those for single-difference difference in Hgb levels. Because the test uses as a counterfactual women whose most recent birth lies somewhat more in the past, identification requires only that the precise timing of pregnancies is ignorable within a limited time window. However, somewhat complex behavior patterns could generate the observed effect. Perhaps most simply, wealth could be associated with different child bearing choices in mining communities and control locations. For instance, it might be that wealthier women (with higher baseline Hgb levels) tend to have fewer children or space out births more in mining communities than in communities farther afield - perhaps because of better earnings opportunities. The DiD effect could then be due to comparing (relatively) poorer women in mining towns to richer controls in the pregnancy and post-partum group, and (relatively) wealthier women in mining towns to poorer controls for the following years.

To conclusively assess this concern, we first (i) refer to the falsification finding discussed above (in 4.2) that indicated that in terms of general health, infants are not more strongly affected than older children by closeness to any mine or to a heavy metal mine. This is not what one would expect if women with poorer health outcomes tended to have more children in mining communities than in other communities. Furthermore, we (ii) test for DiD effects in asset wealth. Column (9) shows that there are no significant effects. However, since the sign of the coefficient is negative, we (iii) also show results for Hgb recovery when we control directly for the household's asset wealth, the woman's height, or whether she gave birth in an 'improved' setting. (As we have shown, asset wealth is clearly endogenous to mining. Height is a wealth proxy that is plausibly endogenous only in a sub-sample, namely among women who lived in their current location when they were growing up, and in communities with a long history of mining.) As is evident from Columns (6-8), the DiD effect is not sensitive to these controls. Finally, we (iv) show a placebo regression, in which we select mothers observed in the same states and years as those in our DiD sample, but in locations at least 20km away from any mine, and define as a placebo treatment the fact of a household being in the bottom wealth quintile, as opposed to the top quintile. As expected, Columns (10-12) show that women in poor households always have lower Hgb levels than those in wealthy households - but there is no indication of a particular time pattern around pregnancy and postpartum, with placebo DiD coefficients near zero.

In summary, we obtain two DiD tests by disaggregating effects, first among mine types, and then with respect to recent pregnancy. The results are instructive both with respect to mechanisms of treatment transmission and with respect to identification. In terms of mechanisms, they offer strong evidence that the observed health effect is caused by pollution, not other facets of life near mines. For instance, if the observed effect on Hgb were due to iron deficiency or malaria infection,

nutritional behavior and infection rates would have to vary in systematically different ways near metal and non-metal mines, and among pregnant and non-pregnant women across distance groups, despite the fact that socio-economic outcomes do not vary in such ways. The results provide reassurance on identification, most importantly because they are very hard to explain with preferential sorting. Because mine types differ in health impacts, but not in wealth impacts, one would have to hypothesize that in their migration decisions, people not only take mine type into account, but also differentially sort on their potential health and wealth outcomes. (We have discussed above the corollary for Hgb recovery.) This would require an extraordinary level of sophistication and information.

Patterns of anemia among children mirror those among women, but are less conclusive

Our data shows patterns of anemia among children in mining communities that resemble those found among adult women. However, significant results are hard to come by. This may be because the true treatment effect is weaker - we have noted above (Section 2.4.1) that young children are known to be able to effectively compensate for the hematotoxic effects of lead. It may also be due to small sample size (for children, we only have about half the number of observations in the women's sample), and the fact that the anemia effect appears to depend on age, with the youngest children the most affected. Because the evidence is inconclusive, we discuss results relatively briefly.

In the cross-section, we observe decreases in blood Hgb in children of about 0.07-0.09 g/dL (significant only with state-year group effects, rather than mine-year effects), with insignificant increases in anemia on the order of two percentage points. (Table 12) The panel shows statistically insignificant losses of 0.32 g/dL from current exposure, but insignificant gains from consistent exposure. (Table 13, Columns 4-5) When disaggregating by mine type, the effect is strongly concentrated near heavy metal mines, but the DiD coefficient is again not significant. (Table 14, Column 3)

Because we have shown that in mining communities, women (and pregnant women in particular) tend to be anemic, and because it is known that children are born with a lead burden mirroring that of their mothers, we explore whether infants might be more strongly affected by pollution than older children. Results are listed in Table 14. When we consider impacts on infants only, we find a larger but insignificant effect on Hgb (0.14 g/dL), with no differential impact on the incidence of anemia, relative to infants living farther away from the mine. (Columns 5 and 10) However, the differential impact on infants near heavy metal mines (Columns 7 and 11) is both significant and large. The triple-difference coefficient shows a 0.56 g/dL difference in Hgb levels, with an increase in anemia incidence of 15 percentage points, and a nearly identical difference in differences

between the effect on infants near heavy metal mines and other mines. However, we caution that infants born in the direct vicinity of heavy metal mines tend to live in poorer households. (Column 8) As shown above, we did not find such a correlation between mine type and wealth in our analysis of hematotoxic effects among women living near heavy metal mines. The fact that we do find it here makes it less compelling to interpret the difference among mine types as evidence that the health impacts are due to pollution.

5.4 Evidence of adverse growth outcomes

As noted, exposure to environmental lead has previously been linked to decreased growth early in life. However, the evidence is mixed. In the following, we consider impacts on height for age, and the incidence of stunting (height more than two standard deviations below the age-appropriate median), as well as on absolute birth weight and the incidence of low birth weight (less than 2,500g). We find strong evidence of lower height among infants exposed to a mining environment, and evidence of lower height in older children, but little indication of reduced birth weight.

In the long-run, children in mining communities grow taller than their peers; they are born with equal birth weight, but are likely to be born less tall

In the cross-section, we observe *better* outcomes for height among children of less than five years of age in mining communities (including near heavy metal mines) than in the controls, and no differential impact on birth weight. (Table 15) This may not be surprising: growth is strongly linked with nutrition, and with greater wealth in mining communities, there may also be better diets.

However, the evidence is somewhat more subtle. Firstly, there is no indication of a positive effect among never-movers. This is certainly consistent with a higher likelihood of exposure to pollution, both directly and through the maternal body burden of lead - although it could of course also be due to sorting or the somewhat lower though not statistically different economic benefits among never-movers as compared to the general population. More intriguingly, although we find no differences in birth weight, height is affected in infants born near heavy metal mines. We do not find a differential impact on stunting in all children under five years of age among mines where contamination with heavy metals is to be expected. However, for *infants*, the differential effect of closeness to a heavy metal mine matters: the triple-difference effects are adverse, and significant in the case of stunting and severe stunting. The DiD comparing the treatment effect of closeness on infants near metal mines and other mines amounts to a loss of 0.1 standard deviations

in the height measure, and a four and two percentage point increase in the incidence of stunting and severe stunting, respectively. (Table 16) There is no difference between mine types in the economic status of families with infants. (Columns 7 and 8) As a falsification test, we show that there is no differential impact of closeness to a heavy metal mine on birth weight among infants, as expected. (Columns 9-12)

While the cross-sectional evidence is mixed, it thus points to an adverse effect of *in utero* exposure on height (though not birth weight) that then either dissipates with age, or is reversed by some countervailing positive impact.

Panel evidence shows that in utero exposure to mining increases the incidence of stunting

Results from the mine-level pseudo-panel confirm that there is no effect of mining activity on birth weight, that there is an effect on height, and that the latter is chiefly due to exposure *in utero*, and attenuates with age. It also allows us to at least suggest that there are genuinely positive effects of life in mining communities on growth in older children, so that children do not simply 'out-grow' the effects of *in utero* exposure without further exposure, as earlier reported by Shukla et al. (1991).

The DiD effect of *in utero* exposure among all children under five years of age shows a loss of 0.14 standard deviations in the height index, and a five percentage point increase in the incidence of stunting and severe stunting (over a baseline incidence of 23% and 8%, respectively). (Table 17, Column 1) The effect on the discrete outcomes is significant and robust; the one on the continuous measure not significant (t = 1.39), but stable. In the case of the continuous index and of stunting, the effect is larger and stronger when we estimate it for infants only. (Column 2 and 12)

The estimated effect of exposure during the first year of life is centered near zero for the continuous index, and adverse for the discrete outcomes. (Column 3) Results when estimating in utero and birth-year effects jointly are more instructive. (Column 4) We find robust, stable, and large adverse effects of *in utero* exposure on the continuous index (0.5 standard deviations) and the discrete outcomes (an increase in incidence of eight and seven percentage points, respectively), alongside stable (and in one case, significant) beneficial effects of birth-year exposure. This is at least consistent with exposure to maternal lead loads *in utero*, alongside positive effects from the socio-economic benefits of mining, once the child is born. (Since mine operation *in utero* and during the first year of life is highly correlated, this finding also mechanically explains the relatively subdued effects from the single-year regressions.)

While it is attractive to allow *in utero* and birth-year effects to jointly enter into the model, the sample of children born just before and just after a mine opened or closed is small.[22] To

[22]The DHS surveys record only health data from children born no more than five years before the survey time. This helps identification, but limits sample size, in particular where we use mother-level effects.

further solidify the result, we therefore show that a similar pattern emerges when we first estimate separately the effect of the mine operating during the survey year (Column 5), and then compare this estimate to the one obtained when we include also the effect of the mine operating during gestation. (Column 6) Similarly, for height-for-age and stunting, there is an attenuated effect of continuous exposure from conception to the survey time, contrasted with children who have never been exposed. (Column 7) Both results conform with a push-pull between pollution and socio-economic effects, and are less consistent with a mere attenuation of impacts as in Shukla et al. (1991).

In Table 17, the last three columns for each outcome variable (e.g. column 8,9 and 10) show that, when we estimate the effects of in utero and birth-year exposure with mother-level effects, the results match the pattern found in the mine-level panel, but are not significant. This is regrettable, but perhaps to be expected: although we observe more than 2,000 women near mines in our sample for whom our data records child growth outcomes for at least two children born within five years of each other, the sub-sample of mothers with births both while the mine was operational and while it was not operational is now very small.

Patterns among children born to migrants provide further evidence of adverse effects on growth

Because the panel evidence suggests that it is *in utero* exposure to mining that matters for growth outcomes, we are able to leverage differences in the timing of exposure among children born to migrants for an additional test. More precisely, we compare measures of growth in children born to migrants - before and after migrating (in our preferred specification, within four years at most), and born to mothers who moved to locations close to a mine and those who moved to slightly more remote locations. (For details on how exposure groups are defined, refer to Appendix A.) This strategy uses the cross-sectional mining data, and therefore helps generate a larger sample.

The DiD estimate of the effect on height-for-age of moving to a mining community before conception is a 0.36 standard deviation adverse effect when controlling for mother-fixed effects, and a 0.11 standard deviation insignificant adverse effect when allowing only for mine-year effects. (Table 18) When we discard any observation for which treatment status is ambiguous, results in the mother-level model are consistent in sign and approximate magnitude, but strongly attenuated. This is as expected, given that the 'donut hole' approach strongly restricts and already small sample. In the mine-year model, estimated effects are larger and stronger than in the single cutoff model. This is again intuitive, since the 'donuthole' definition likely reduces measurement error in the treatment variable, but in the mine-year model, achieves this without reducing sample size

too drastically. The results are reasonably stable when we change time window around the time of migration. Additional robustness checks are provided in Appendix I.

For birth weight, we observe a large and significant adverse effect on the order of 160g-190g when we compare births to the same mother, but no impact when we allow only for mine-level effects. We can think of no obvious explanation for the divergence between the mine and mom-level models - in particular since sorting is unlikely to be an issue.

6 Correlates of Treatment Effects

We conclude our analysis of the local welfare effects of mining by showing how treatment effects vary with some characteristics of the local economies. To this end, we regress mine-level estimates of the effect of closeness to the mine on variables describing the economic, geographic and policy features either of the countries or the sub-national administrative units where the mine is located. We describe in Appendix E the process of obtaining mine-level estimates; estimates can only be obtained for a sub-set of mines in our sample. We also note that, while our aggregate estimates are well-identified, the correlates of treatment effects we show here have no causal interpretation.

With these caveats in mind, we describe three noteworthy patterns. First, we show that in a regional disaggregation, residents of communities near mines in sub-Saharan Africa benefit more in terms of wealth, and suffer less in terms of health impacts, than residents of mining towns elsewhere. Secondly, the data suggests that the economic benefits to mining communities are greatest where the overall economic environment is poor. (This result has a weaker correlate in health effects.) Finally, we note that, while we find no compelling evidence that effects vary with general measures of the quality of governance, participation in the Extractive Industries Transparency Initiative (EITI) does correlate with greater wealth benefits.

Figure 6 shows the distribution of mine-level coefficients for the effect of closeness on wealth and women's hemoglobin. We note, first, that, while there is a wide range of effects, there is no indication that outliers are driving results. Median effects go in the same direction as the global mean effect. A regional disaggregation of treatment effects shows economic gains in all regions, including the two most comprehensively represented in the sample - Latin America and the Caribbean and sub-Saharan Africa. (Figure 7) Economic benefits are substantially higher in sub-Saharan Africa, and (perhaps in consequence, given our reduced-form estimates), mean health effects near African mines are statistically zero, while they are strongly negative in the other regions. (Table

19, Columns 3 and 13) The distribution of coefficients by country is displayed in Figure 8; as is evident, there is great variation in treatment effects even within world regions.

We find evidence to suggest that wealth effects in mining communities are strongest where the economic environment is weak. This can be interpreted to suggest that the local economic effect of mining is driven not by the interaction of mining with other economic activity, but by the opportunities mining provides in areas where there is a paucity of other options. Thus, the local wealth benefits of mining matter less in country-years with high GDP. Differences in GDP largely absorb the regional disparities reported above, although the sample is too small to yield conclusive evidence. Similarly, the expected economic benefits are lower (if never negative) in states with high average levels of schooling; indeed, poor human capital dominates GDP as a predictor. Yet again similarly, mining has less influence on the economic prospects of local communities in areas with good access to the sea, and in areas with good access to electricity (although the effect correlates only weakly with problems with access to land, and does not correlate with average travel time to the nearest city). The relationship of effect estimates with the aggregate measure of institutional quality defined in Gennaioli et al. (2013) is weak but consistent with the other evidence (better institutions are associated with weaker effects). We find no correlation with dimensions of the World Bank's CPIA (not shown).

In light of the absence of strong correlations with general governance measures, it is noteworthy that we do observe stronger wealth effects in surveys conducted in countries at a time when they had at least completed an EITI report - whether they were in compliance or not. (We caution again against a causal interpretation of this correlation.) The EITI initiative is recent, and hence, few mines are observed after completion of a report. Since engagement with the EITI predates completion of reports, we also show the correlation between effects and a dummy recording whether a country participated in the EITI at any point in time. When we use the latter definition, we find weak but consistent positive effects.

In the smaller sample of mine-level effects on hemoglobin, the estimated impact is more adverse in country-years with higher GDP. (There is no conclusive evidence of correlates of treatment effects other than GDP; we omit results for conciseness.) Stronger adverse effects in wealthy countries are consistent, of course, with weaker socio-economic benefits, and hence, potentially weaker compensation for pollution effects. However, whereas there was no indication of a curvature in the relationship between GDP and local wealth effects, the relationship between Hgb effects and GDP exhibits a distinct inverse-U shape. That is, adverse health effects are strongest in the poorest and the wealthiest countries in the sample. It is noteworthy that the data shows more adverse effects in the poorest countries, despite the fact that it is here that wealth effects are strongest. Possible explanations include a lower wealth elasticity of health spending in poorer countries, or an independent effect of poor regulation; we are not able to assess these pathways conclusively.

7 Conclusion

Summary

We present the first comprehensive empirical assessment of the health-wealth tradeoff in communities near mines, using micro-data from 44 developing countries. In communities in the vicinity of mines, we observe important economic benefits, alongside serious health impacts, namely increases in the incidence of anemia in adult women, and of stunting in young children. These health impacts have previously been observed at levels of exposure to lead pollution that are known to also cause cognitive deficits in children.

We obtain estimates of short to medium-term effects from mine-level and mother-level pseudo-panels. Long-run cross-sectional estimates come from a group-effects model; we confirm them with IV results that use deposit location and characteristics to instrument for mine locations. To make plausible that the observed health impacts are due to pollution, we develop additional difference-in-difference tests that leverage (i) the association of certain mine types with lead pollution, and (ii) known pathological patterns of Hgb recovery in adults exposed to lead.

The economic benefits to mining communities in the long run are on the order of 0.1 standard deviations of a country and year-specific asset index. Short to medium-term benefits to households in communities near operating mines are larger, on the order of 0.3 standard deviations. Benefits are strongly concentrated within the immediate vicinity (5km) of mines, and there are no asset wealth effects at all beyond some 15-20km. Benefits in terms of healthcare extend beyond the most direct vicinity of mines, although mining communities benefit at least as much as communities farther afield.

The minerals mined in a given location are tightly associated with a range of pollutants likely to be present in the environment, and the latter, with specific health conditions. We particularly leverage the association of polymetallic (Copper, Gold, Lead, Silver, and Zinc) mines with lead and other heavy metal pollution. We focus on analyzing impacts in mining communities on two known consequences of lead exposure - namely, anemia in children and adults, and growth deficits in children. The evidence conclusively reveals that the real economic benefits generated in mining communities go hand in hand with increases in the incidence of anemia, by three to nine percentage points in adult women. There is weaker but consistent evidence of similar effects in children. Children in mining communities are not disadvantaged in all aspects of physical growth. Yet, young children in mining communities are far more likely to be stunted or severely stunted than those born in control groups, with an increase in incidence of five percentage points. There is very limited evidence of reduced birth weight, and increases in stunting are clearly strongest among

infants, and may persist less among older children. We show that by contrast to previously reported results, our data suggests that it is *in utero* exposure to mining that matters for growth impacts in children, and that such impacts do not exclusively result from prolonged exposure - at least when we consider reduced-form impacts among a population where there are significant economic benefits from mining operations.

This paper has shown reduced-form results throughout. To make plausible that the observed health impacts are due to pollution, we demonstrate that impacts on known consequences of lead exposure are by far the strongest near mines where pollution with lead and other heavy metals is particularly likely. We also show that women who live close to mines recover more slowly from blood loss during pregnancy and delivery, a pattern characteristic of lead exposure. By way of contrast to these specific health impacts, there is no general pattern of ill health across all mining communities, nor is there such a pattern in communities close to mines where heavy metal contamination is to be expected.

Contribution to the literature

We hope to contribute to the literature in three distinct ways: by presenting the first comprehensive assessment of the local welfare effects of mineral mining; by highlighting the presence of a wealth-health trade-off in a major industry in developing countries; and by complementing the case study-based public health literature with a broad and well-identified assessment of overt health effects in the general population exposed to every-day mining operations, across a wide range of developing countries.

The fact that we find adverse health effects even in the presence of economic benefits poses an important puzzle. The standard framework of utility optimization over health and consumption does not easily yield an optimal choice of worse health alongside higher wealth among residents of mining towns. Possible explanations include limited information - we have argued above that contamination near mines may not always be easily observed, and that its health impacts may not be widely known. The fact that we find strongly raised wealth levels, but only weakly better health care among households in the direct vicinity of mines at least suggests that residents of mining communities are not making very decisive health investments to compensate for exposure to pollution. An alternative explanation might suggest that the cost of avoiding exposure to pollution is high, perhaps because decision on whether to live in mining towns in developing countries is less like choosing an optimal distance along a continuum, and more like a discrete choice between two stark options - namely living either in relatively unpolluted communities outside of a reasonable commuting distance of the mine, or in a highly polluted but economically active community adjacent to the mine. The fact that in our analysis of correlates of treatment effects, we observe the

greatest wealth effects in areas that are economically less active is certainly consistent with such a situation.

Policy implications

We conclude by highlighting some policy implications of our results.

Firstly, we note that, while we must leave open the exact magnitude of aggregate health damages,[23] the presence of an observable health externality due to normal operations at mines in our sample suggests that containment and remediation of mining pollution need to be revisited. The cost to affected individuals is clearly very significant, with estimates of productivity losses due to anemia in adults have been estimated to be on the order of 5-17%. Yet, the case for added scrutiny is even more urgent: we work throughout with reduced-form estimates, so that our findings should be read as *compensated* health impacts. As Graff Zivin and Neidell (2013) note, "optimal regulation will occur at the point where the marginal costs of regulations ... are equal to the averted health, avoidance, and medical costs associated with that marginal reduction in pollution". (p. 12) The policy-relevant uncompensated cost of health effects may therefore be considerably larger than the reduced-form results reflect.

Secondly, we have noted that the presence of adverse health impacts in a wealthier population may suggest either that there is insufficient information on the health risks associated with mining pollution, or that the cost of avoiding exposure to pollution is very high - perhaps because of where housing is available, and the quality of public transport. Information deficits would be more easily addressed than infrastructure deficits. However, since we also find that health impacts are limited to the immediate vicinity of mines (consistent with prior evidence on the extent of pollution), there may be some scope for policy to protect residents from the worst pollution.

Finally, empirical evidence on a high but localized wealth impact of mining suggests that mining will tend to generate spatial inequalities. These need not necessarily be of concern, but could be worrisome where they undermine political stability, or lead to dynamic inefficiencies in access to schooling and economic opportunity.

[23] Estimating the aggregate cost of health impacts is complicated by data limitations; we show illustrative calculations in Appendix J.

References

Acemoglu, D., Finkelstein, A., and Notowidigdo, M. J. (2013). Income and health spending: evidence from oil price shocks. *Review of Economics and Statistics*, 95(4):1079–1095.

Alloway, B. J. (2013). Sources of heavy metals and metalloids in soils. In *Heavy metals in soils*, pages 11–50. Springer.

Aragón, F. M. and Rud, J. P. (2013a). Modern industries, pollution and agricultural productivity: Evidence from Ghana. Working paper.

Aragón, F. M. and Rud, J. P. (2013b). Natural resources and local communities: evidence from a Peruvian gold mine. *American Economic Journal: Economic Policy*, 5(2):1–25.

ATSDR (2007). Toxicological profile for lead. Technical report, US Department of Health and Human Services. Public Health Service, Atlanta, GA, USA.

Baghurst, P. A., McMichael, A. J., Wigg, N. R., Vimpani, G. V., Robertson, E. F., Roberts, R. J., and Tong, S.-L. (1992). Environmental exposure to lead and children's intelligence at the age of seven years: the Port Pirie Cohort Study. *New England Journal of Medicine*, 327(18):1279–1284.

Balarajan, Y., Ramakrishnan, U., Özaltin, E., Shankar, A. H., and Subramanian, S. (2012). Anaemia in low-income and middle-income countries. *The Lancet*, 378(9809):2123–2135.

Bellinger, D., Leviton, A., Rabinowitz, M., Allred, E., Needleman, H., and Schoenbaum, S. (1991). Weight gain and maturity in fetuses exposed to low levels of lead. *Environmental research*, 54(2):151–158.

Bellinger, D. C. (2004). Lead. *Pediatrics*, 113(4 Suppl):1016–1022.

Benin, A. L., Sargent, J. D., Dalton, M., and Roda, S. (1999). High concentrations of heavy metals in neighborhoods near ore smelters in northern Mexico. *Environmental Health Perspectives*, 107(4):279.

Berkman, D. S., Lescano, A. G., Gilman, R. H., Lopez, S. L., and Black, M. M. (2002). Effects of stunting, diarrhoeal disease, and parasitic infection during infancy on cognition in late childhood: a follow-up study. *The Lancet*, 359(9306):564–571.

Bobonis, G. J., Miguel, E., and Puri-Sharma, C. (2006). Anemia and school participation. *Journal of Human Resources*, 41(4):692–721.

CDC (2012). Low level lead exposure harms children: A renewed call of primary prevention. *Atlanta, GA: Centers for Disease Control and Prevention*.

Chen, Y., Ebenstein, A., Greenstone, M., and Li, H. (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. *Proceedings of the National Academy of Sciences*, 110(32):12936–12941.

Corno, L. and De Walque, D. (2012). Mines, migration and HIV/AIDS in southern africa. *Journal of African Economies*, 21(3):465–498.

Currie, J., Davis, L., Greenstone, M., and Walker, R. (2013). Do housing prices reflect environmental health risks? evidence from more than 1600 toxic plant openings and closings. Working Paper 18700, National Bureau of Economic Research.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica*, 78(6):1863–1903.

Dewey, K. G. and Begum, K. (2011). Long-term consequences of stunting in early life. *Maternal & child nutrition*, 7(s3):5–18.

Dube, O. and Vargas, J. (2013). Commodity price shocks and civil conflict: Evidence from Colombia. *The Review of Economic Studies*.

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

Ebenstein, A. (2012). The consequences of industrialization: Evidence from water pollution and digestive cancers in China. *Review of Economics and Statistics*, 94(1):186–201.

Factor-Litvak, P., Slavkovich, V., Liu, X., Popovac, D., Preteni, E., Capuni-Paracka, S., Hadzialjevic, S., Lekic, V., LoIacono, N., Kline, J., et al. (1998). Hyperproduction of erythropoietin in nonanemic lead-exposed children. *Environmental Health Perspectives*, 106(6):361.

Factor-Litvak, P., Wasserman, G., Kline, J. K., and Graziano, J. (1999). The yugoslavia prospective study of environmental lead exposure. *Environmental Health Perspectives*, 107(1):9.

Filmer, D. and Pritchett, L. H. (2001). Estimating wealth effects without expenditure dataor tears: An application to educational enrollments in states of India. *Demography*, 38(1):115–132.

Fontúrbel, F. E., Barbieri, E., Herbas, C., Barbieri, F. L., and Gardon, J. (2011). Indoor metallic pollution related to mining activity in the Bolivian Altiplano. *Environmental Pollution*, 159(10):2870–2875.

Frankel, J. A. (2010). The natural resource curse: a survey. Working Paper 15836, National Bureau of Economic Research.

Fullerton, D. G., Bruce, N., and Gordon, S. B. (2008). Indoor air pollution from biomass fuel smoke is a major health concern in the developing world. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 102(9):843–851.

Gallego, F., Montero, J.-P., and Salas, C. (2013). The effect of transport policies on car use: Evidence from Latin American cities. *Journal of Public Economics*, 107(0):4762.

Gennaioli, N., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2013). Human capital and regional development. *The Quarterly Journal of Economics*, 128(1):105–164.

Graff Zivin, J. and Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, 51:689–730.

Grandjean, P., Jensen, B. M., Sandø, S., Jøgensen, P., and Antonsen, S. (1989). Delayed blood regeneration in lead exposure: an effect on reserve capacity. *American journal of public health*, 79(10):1385–1388.

Graziano, J. H., Slavkovic, V., Factor-Litvak, P., Popovac, D., Ahmedi, X., and Mehmeti, A. (1991). Depressed serum erythropoietin in pregnant women with elevated blood lead. *Archives of Environmental Health: An International Journal*, 46(6):347–350.

Greenstone, M. and Jack, B. K. (2013). Envirodevonomics: A research agenda for a young field. Working Paper 19426, National Bureau of Economic Research.

Hartwick, J. M., Olewiler, N. D., et al. (1986). *The economics of natural resource use*. Harper & Row New York.

Hendryx, M. and Ahern, M. M. (2008). Relations between health indicators and residential proximity to coal mining in West Virginia. *American Journal of Public Health*, 98(4):669–671.

Hernandez-Avila, M., Peterson, K. E., Gonzalez-Cossio, T., Sanin, L. H., Aro, A., Schnaas, L., and Hu, H. (2002). Effect of maternal bone lead on length and head circumference of newborns and 1-month-old infants. *Archives of Environmental Health: An International Journal*, 57(5):482–488.

Horton, S. and Ross, J. (2003). The economics of iron deficiency. *Food policy*, 28(1):51–75.

Infomine (2013). Company and property mining intelligence.

IntierraRMG (2013). Raw materials data.

Kotsadam, A. and Tolonen, A. (2013). Mineral mining and female employment. *Oxcarre Research Papers*, 114.

Landrigan, P. J. and Baker, E. L. (1981). Exposure of children to heavy metals from smelters: epidemiology and toxic consequences. *Environmental Research*, 25(1):204–224.

Lanphear, B. P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D. C., Canfield, R. L., Dietrich, K. N., Bornschein, R., Greene, T., et al. (2005). Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. *Environmental Health Perspectives*, 113(7):894.

Lauwerys, R., Buchet, J., Roels, H., and Hubermont, G. (1978). Placental transfer of lead, mercury, cadmium, and carbon monoxide in women: I. comparison of the frequency distributions of the biological indices in maternal and umbilical cord blood. *Environmental research*, 15(2):278–289.

McCann, J. C. and Ames, B. N. (2007). An overview of evidence for a causal relation between iron deficiency during development and deficits in cognitive or behavioral function. *The American journal of clinical nutrition*, 85(4):931–945.

Monteiro, J. and Ferraz, C. (2009). Does oil make leaders unaccountable? Technical report, Mimeo. Pontificia Universidade Catolica do Rio de Janeiro.

Rau, T., Reyes, L., and Urzúa, S. S. (2013). The long-term effects of early lead exposure: Evidence from a case of environmental negligence. Working Paper 18915, National Bureau of Economic Research.

Recio-Vega, R., Valdez-Abrego, C., Adame-Lopez, B., and Gurrola-Mendez, A. (2012). Surveillance of elevated blood lead levels in children in Torreon, Coahuila, Mexico, 1998–2010. *International Journal of Hygiene and Environmental Health*, 215(5):507–513.

Ripley, E. A., Redmann, R. E., Crowder, A. A., et al. (1996). *Environmental effects of mining*. St. Lucie Press Delray Beach, FL.

Roels, H., Bucket, J.-P., Lauwerys, R., Hubermont, G., Bruaux, P., Claeys-ThOreau, F., Lafontaine, A., and Overschelde, J. V. (1976). Impact of air pollution by lead on the heme biosynthetic pathway in school-age children. *Archives of Environmental Health: An International Journal*, 31(6):310–316.

Roels, H. A., Buchet, J.-P., Lauwerys, R. R., Bruaux, P., Claeys-Thoreau, F., Lafontaine, A., and Verduyn, G. (1980). Exposure to Lead by the oral and the pulmonary routes of children living in the vicinity of a primary lead smelter. *Environmental research*, 22(1):81–94.

Roper, W. L., Houk, V., Falk, H., and Binder, S. (1991). Preventing lead poisoning in young children: A statement by the Centers for Disease Control, october 1991. Technical report, Centers for Disease Control, Atlanta, GA (United States).

Sanın, L. H., González-Cossıo, T., Romieu, I., Peterson, K. E., Ruız, S., Palazuelos, E., Hernández-Avila, M., and Hu, H. (2001). Effect of maternal lead burden on infant weight and weight gain at one month of age among breastfed infants. *Pediatrics*, 107(5):1016–1023.

Schwartz, J., Landrigan, P. J., Baker Jr, E. L., Orenstein, W. A., and Von Lindern, I. (1990). Lead-induced anemia: dose-response relationships and evidence for a threshold. *American journal of public health*, 80(2):165–168.

Shukla, R., Bornschein, R. L., Dietrich, K. N., Buncher, C., Berger, O. G., Hammond, P. B., and Succop, P. A. (1989). Fetal and infant lead exposure: effects on growth in stature. *Pediatrics*, 84(4):604–612.

Shukla, R., Dietrich, K. N., Bornschein, R. L., Berger, O., and Hammond, P. B. (1991). Lead exposure and growth in the early preschool child: a follow-up report from the Cincinnati lead study. *Pediatrics*, 88(5):886–892.

Sloan, N. L., Jordan, E., and Winikoff, B. (2002). Effects of iron supplementation on maternal hematologic status in pregnancy. *American Journal of Public Health*, 92(2):288–293.

Toscano, C. D. and Guilarte, T. R. (2005). Lead neurotoxicity: from exposure to molecular effects. *Brain Research Reviews*, 49(3):529–554.

United States Geological Survey (2005). Mineral Resources Data System.

Wilson, N. (2012). Economic booms and risky sexual behavior: Evidence from Zambian copper mining cities. *Journal of Health Economics*.

World Health Organization (2011). Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity.

Wright, D. A. and Welbourn, P. (2002). *Environmental toxicology*, volume 11. Cambridge University Press.

Zhu, M., Fitzgerald, E. F., Gelberg, K. H., Lin, S., and Druschel, C. M. (2010). Maternal low-level lead exposure and fetal growth. *Environmental health perspectives*, 118(10):1471.

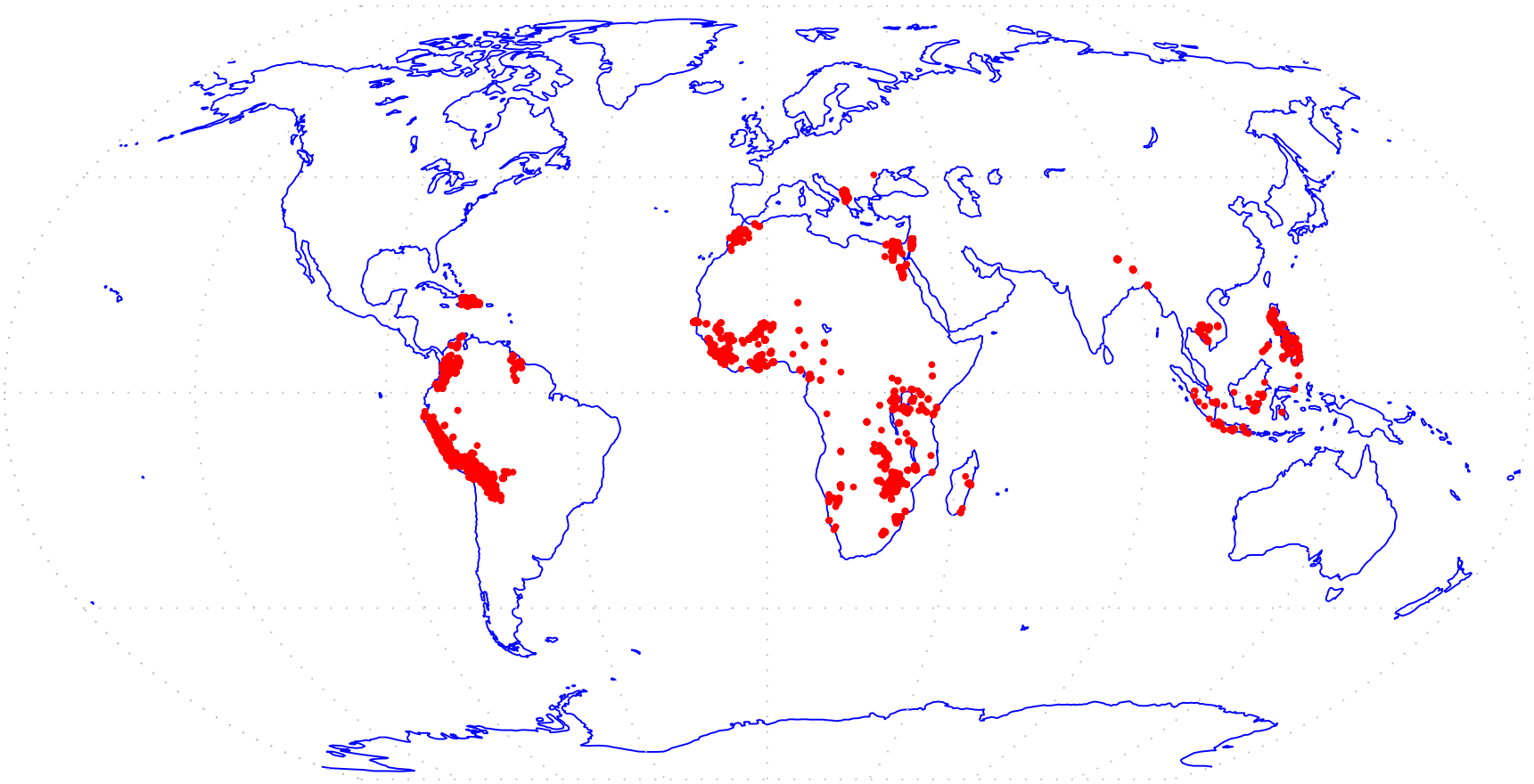

Figure 1: DHS clusters within no more than 20km of a mine in the sample

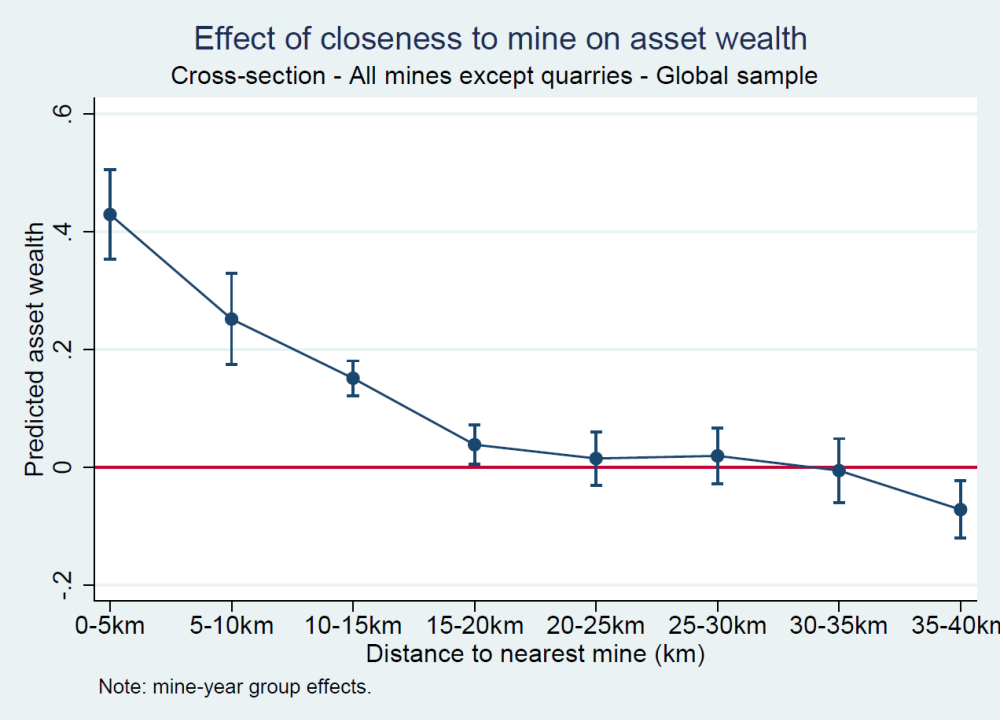

Figure 2: Effect of closeness to mine on asset wealth

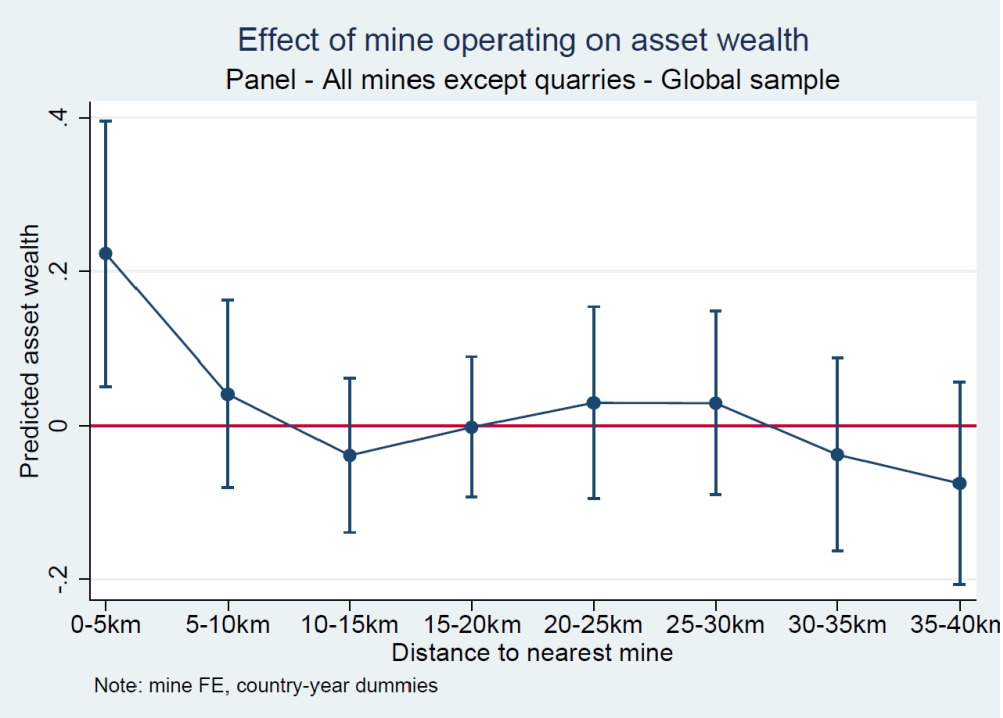

Figure 3: Effect of mine operating on asset wealth

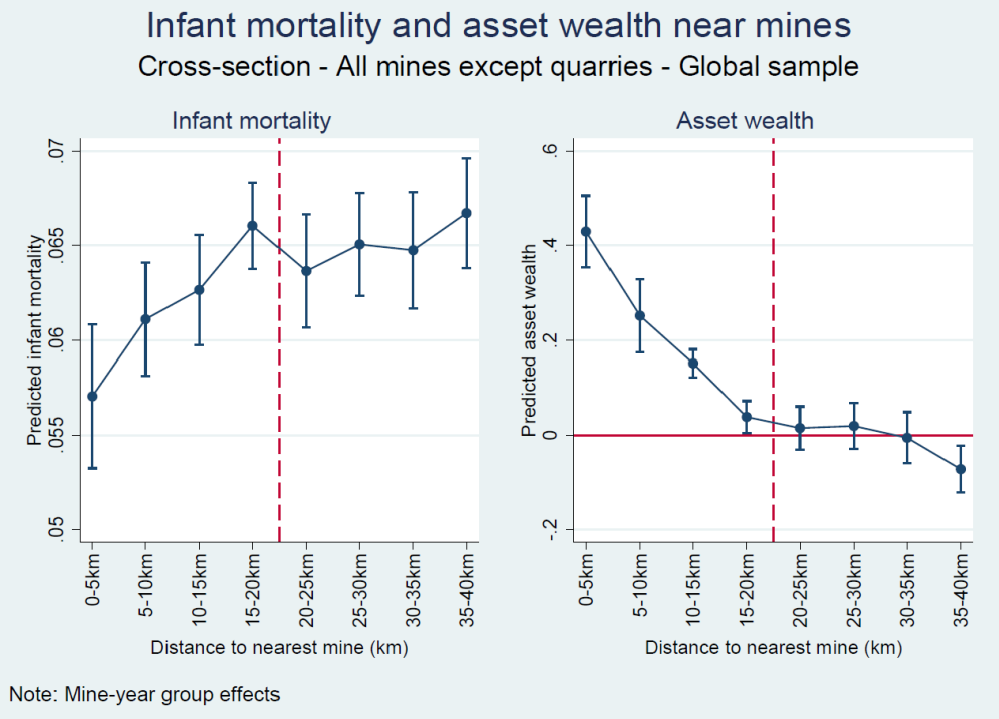

Figure 4: Infant mortality and assest wealth near mines

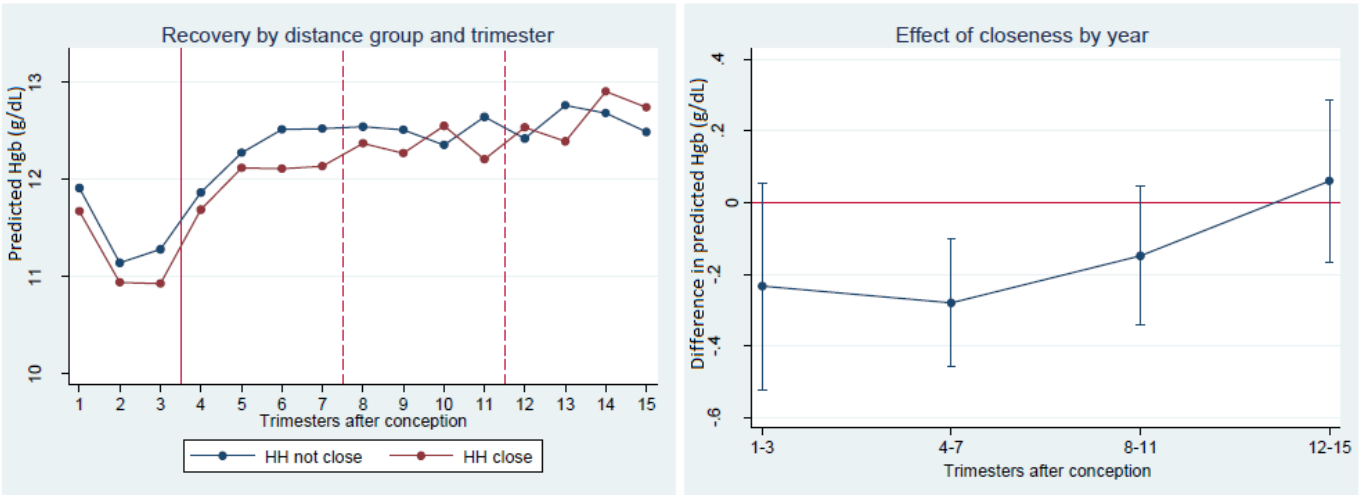

Figure 5: Effect of closeness to mine on maternal Hgb recovery

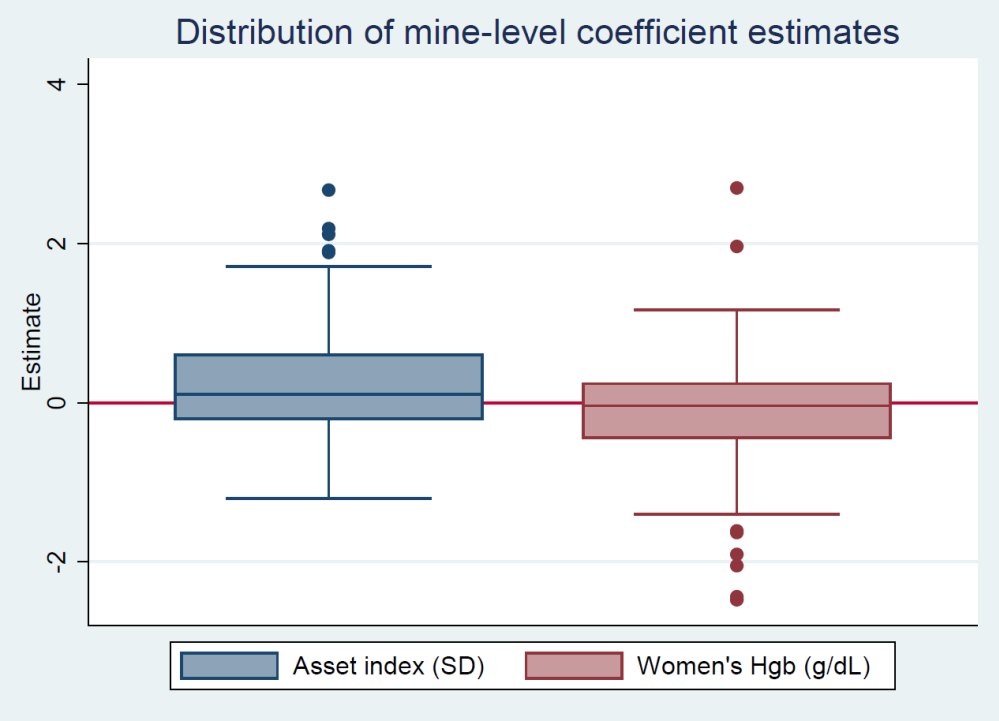

Figure 6: Distribution of mine-level coefficient estimates

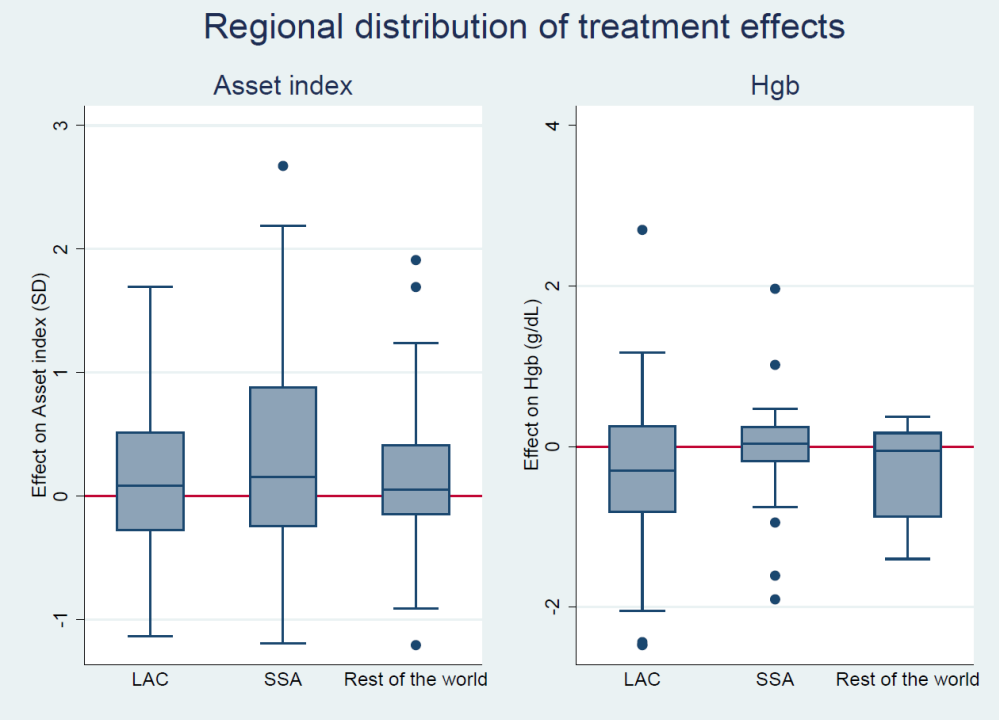

Figure 7: Regional distribution of treatment effects

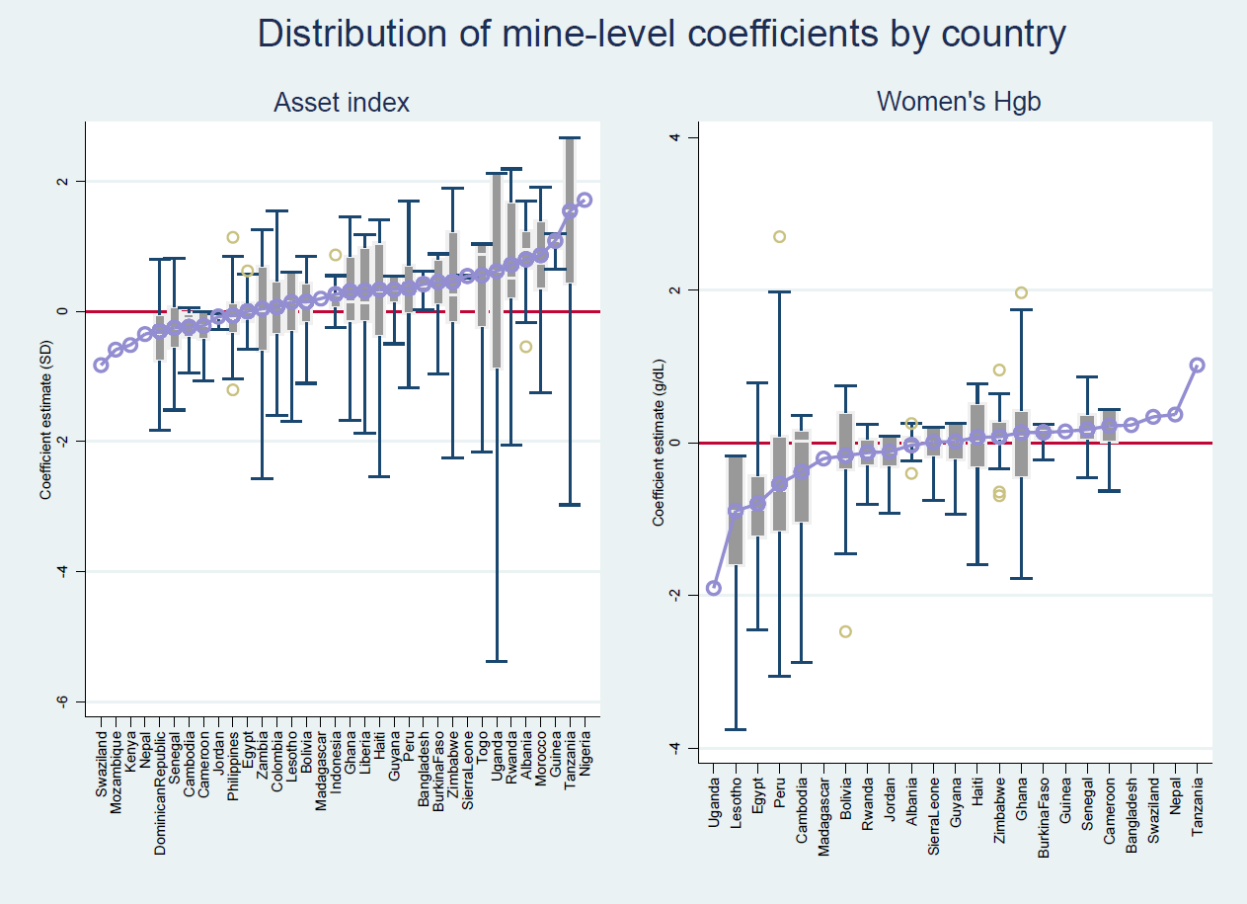

Figure 8: Distribution of mine-level coefficients by country

Table 1 - Mine types and pollutants

Mine type	Pollutants of concern	Health effects of particular concern from plausible exposure levels
Polymetallic mines	Heavy metals, especially lead	Neurodevelopmental damage, anemia, growth deficits (from Lead)
Small-scale gold and silver mining	Mercury	Renal problems, neurological conditions
Large-scale gold mining	Cyanide	Heart irregularities, thyroid problems
Bulk metal mines	Particulates	Respiratory problems, GI problems from turbid water
Coal	Particulates, radionuclides	Respiratory problems, GI problems, lung cancer, non-cancer respiratory disease
Phosphate rock	Radionuclides	Lung cancer and non-cancer respiratory disease
Quarries, including diamond mines	Particulates	Respiratory problems, GI problems from turbid water
Smelters	*Heavy metals, SO2*	*As above, and respiratory problems*

Note: based on ATSDR Toxicological Profiles, Alloway 2013, Ripley 1996, Wright and Welbourn 2002. Categories are not exclusive. Health effects as reported from chronic low-level environmental exposure.

Table 2 - Sample size

	Overall sample size		
	Pooled DHS	Within 5km of a mine	Within 5-20km of a mine
Households	1,192,492	37,608	132,797
% of total		*3.2%*	*11.1%*
Children under five years of age	1,364,156	31,964	121,519
Women aged 15 and over	2,877,024	87,234	310,096
Men aged 15 and over	2,717,928	82,973	294,723

	With observations within 20km of a mine
Countries	44
Interview years	25
Survey rounds	104

Note: sample size based on all mines excluding quarries.

Table 3 - Sample of mines

	Mines and mineral deposits				
		Mines			
	All developing-country locations	With DHS cluster within 20km	With DHS cluster within 5-20km	With DHS cluster within 0-5km	With DHS cluster in both distance categories
USGS					
All entries	25,068				
All entries excluding quarries	21,190	838	687	339	226
Of which:					
Active mines	*3,839*				
Legacies	*7,526*				
Unmined deposits	*9,525*				
Smelters	*300*				
RMD					
All entries excluding quarries		508	455	225	172
Infomine					
All entries excluding quarries		7	6	4	3

Table 4 - Cross-sectional effects on asset wealth

	Asset index			
	All HHs		Never-movers	
	(1)	(2)	(3)	(4)
HH close to mine	0.105***	0.105***	0.0784*	0.102***
	(0.035)	(0.0314)	(0.0423)	(0.0349)
N	90,319	90,319	31,079	31,079
R-squared	0.094	0.152	0.081	0.203
Number of groups	1,562	554	1,371	441
FE	Mine*year	State*year	Mine*year	State*year
Other controls	Quadratic in household head's age; urban/rural dummy.			

Standard errors clustered at the mine or state level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5 - Asset wealth in the panel

	Asset factor index		
	All HHs		
	(1)	(2)	(3)
HH close to mine	-0.0555	**-0.113**	-0.0874
	(0.0878)	**(0.089)**	(0.0954)
Mine operating	0.0651	**-0.0296**	0.0639
	(0.0566)	**(0.0348)**	(0.133)
Mine operating * HH close (DiD)	0.159*	**0.262*** **	0.229**
	(0.0821)	**(0.0958)**	(0.105)
N	22,579	**22,579**	22,579
R-squared	0.153	**0.13**	0.152
States	141		
Mines		**218**	218
FE	State	Mine	Mine
Time effects	Country*year	Country*year	State*year
	Never-movers		
	(4)	(5)	(6)
HH close to mine	-0.0381	**-0.0352**	0.0128
	(0.0758)	**(0.0892)**	(0.0863)
Mine operating	0.0822	**-0.0585***	0.261
	(0.0663)	**(0.0354)**	(0.161)
Mine operating * HH close (DiD)	0.126*	**0.173***	0.106
	(0.0697)	**(0.0897)**	(0.0879)
N	9,459	**9,459**	9,459
R-squared	0.171	**0.141**	0.167
States	136		
Mines		**205**	205
FE	State	Mine	Mine
Time effects	Country*year	Country*year	State*year

Standard errors clustered at the state or mine level. Baseline specification highlighted in bold type. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 6 - General health outcomes in mining communities in the cross-section

	Child health outcomes				
	Infant mortality	Under-five mortality	Diarrhea	Cough	Fever
	(1)	(2)	(3)	(4)	(5)
HHs within 5 km	-0.00234	-0.00288	0.0115**	0.00514	0.00227
	(0.00229)	(0.00277)	(0.00579)	(0.00969)	(0.00791)
Observations	301,895	301,895	61,671	60,406	59,592
R-squared	0.003	0.005	0.028	0.007	0.009
Number of mine*years	1,566	1,566	1,510	1,503	1,384

	Adult health outcomes			
	Ever miscarried	Female respondent very sick	Night blindness during pregnancy	Male respondent very sick
	(6)	(7)	(8)	(9)
HHs within 5 km	0.00263	0.00328	0.00254	0.0120
	(0.00460)	(0.00527)	(0.0104)	(0.00977)
Observations	117,118	11,022	29,317	9,808
R-squared	0.061	0.011	0.001	0.011
Number of mine*years	1,469	151	1,185	151

Standard errors clustered at the mine level. Sample: all households. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7 - General health outcomes in mining communities in the panel

	Child health outcomes					Adult health outcomes	
	Infant mortality	Under-five mortality	Cough	Diarrhea	Fever	Ever miscarried	Respondent very sick
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HH close to mine	0.00132	0.000568	-0.0265	0.0123	-0.00241	-0.00838	0.00329
	(0.00664)	(0.00768)	(0.0256)	(0.0243)	(0.0203)	(0.0143)	(0.00407)
Mine operating in exposure period	0.00192	0.00133	0.00668	0.00662	0.00791	-0.0191	0.0731***
	(0.00563)	(0.00652)	(0.0219)	(0.0154)	(0.0224)	(0.0134)	(0.0183)
Mine operating in exposure period * HH close	-0.00137	-0.00399	0.00921	-0.00382	-0.0221	-0.00236	-0.00845
	(0.00761)	(0.00881)	(0.0303)	(0.0289)	(0.0262)	(0.0152)	(0.0119)
Observations	43,057	43,057	15,342	15,467	15,594	29,666	4,111
R-squared	0.027	0.031	0.048	0.056	0.046	0.065	0.005
Number of fixed effects	259	259	236	237	230	202	63
Exposure period	In utero	In utero	Survey year	Survey year	Survey year	Survey year	Survey year
Fixed effects	Mine					Mine	
Time effects	Country*year dummies					Country*year dummies	
Other controls	Quadratic in mom age at birth; child age dummies, urban/rural dummy.					Quadratic in respondent's age; urban/rural dummy.	

Standard errors clustered at the mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 8 - Cross-sectional effects on women's Hgb

	Hematotoxic effects on women							
	Altitude-adjusted hemoglobin (g/dL)				Anemia			
	All HHs		Never-movers		All HHs		Never-movers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HHs within 5km	-0.0863**	-0.0494	-0.131	-0.0493	0.0262**	0.00957	0.0495*	0.0217
	(0.0438)	(0.0386)	(0.0838)	(0.0553)	(0.0126)	(0.0101)	(0.0268)	(0.0157)
N	38,217	38,217	13,506	13,506	36,225	36,225	13,204	13,204
R-squared	0.000	0.001	0.001	0.002	0.000	0.001	0.001	0.001
Number of groups	934	292	785	216	934	292	784	216
FE	Mine*year	State*year	Mine*year	State*year	Mine*year	State*year	Mine*year	State*year
Other controls	Quadratic in respondent's age; urban/rural dummy.				Quadratic in respondent's age; urban/rural dummy.			

Standard errors clustered at the state or mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 9 - Panel effects on women's Hgb

	Hematotoxic effects on women					
	Altitude-adjusted hemoglobin (g/dL)			Anemia		
	(1)	(2)	(3)	(4)	(5)	(6)
HH close to mine	0.261*	0.368***	**0.396*** **	-0.0572	-0.0856**	**-0.0949*** **
	(0.151)	(0.137)	**(0.146)**	(0.0391)	(0.0340)	**(0.0347)**
Mine operating in survey year	-0.0234	0.0217	**0.0852**	-0.00637	-0.0221	**-0.0298**
	(0.117)	(0.121)	**(0.136)**	(0.0243)	(0.0315)	**(0.0310)**
Mine operating * HH close (DiD)	-0.280	-0.298*	**-0.330***	0.0634	0.0779*	**0.0864***
	(0.189)	(0.170)	**(0.173)**	(0.0496)	(0.0454)	**(0.0453)**
Observations	9,845	9,845	**9,845**	9,865	9,865	**9,865**
R-squared	0.008	0.011	**0.007**	0.005	0.006	**0.005**
States	69			69		
Mines		122	**122**		122	**122**
FE	State	Mine	Mine	State	Mine	Mine
Time effects	Country*year	Year	Country*year	Country*year	Year	Country*year
Other controls	Quadratic in respondent's age; urban/rural dummy			Quadratic in respondent's age; urban/rural dummy		

Standard errors clustered at the state or mine level. Baseline specification highlighted in bold type. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 10 - Cross-sectional hematotoxic effects on women near different mine types

Hematotoxic effects near different mine types

	Women's hemoglobin (g/dL) and asset index								Anemia and asset index		Falsification tests	
	Hgb (g/dL) (1)	Asset index (2)	Hgb (g/dL) (3)	Asset index (4)	Hgb (g/dL) (5)	Asset index (6)	Hgb (g/dL) (7)	Asset index (8)	Anemia (9)	Asset index (10)	Miscarriage (11)	Grave illness (12)
HH close to mine	-0.0863**	0.145**	-0.0317	0.140*	-0.161*	0.0558	-0.0285	0.144*	0.0125	0.141*	0.00347	0.00366
	(0.0438)	(0.0576)	(0.0533)	(0.0756)	(0.0886)	(0.0573)	(0.0507)	(0.0763)	(0.0154)	(0.0758)	(0.00519)	(0.00595)
HHs close to a 'heavy metal' mine (DiD)			-0.192**	0.0176	-0.253***	-0.0724	-0.192**	0.0162	0.0463*	0.0163	-0.00377	-0.00286
			(0.0944)	(0.101)	(0.0876)	(0.125)	(0.0902)	(0.101)	(0.0248)	(0.101)	(0.0109)	(0.00898)
Other interactions					Region dummies		Pregnancy dummy					
N	38,217	25,695	38,217	25,695	38,217	25,695	36,225	25,676	36,225	25,676	117,118	11,022
R-squared	0	0.111	0.001	0.111	0.001	0.113	0.027	0.113	0.000	0.111	0.061	0.011
Mines	934	932	934	932	934	932	934	932	934	932	1,469	151

Standard errors clustered at the mine level. Other covariates: urban-rural dummy, quadratic in respondent's age. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 11 - Recovery of maternal Hgb

Recovery of maternal Hgb after birth

	Women's hemoglobin (g/dL)								Falsification and placebo			
										Women's Hgb (placebo treatment)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	Asset index (9)	(10)	(11)	(12)
Pregnancy and infancy	-0.569***	-0.602***	-0.530***	-0.588***	-0.588***	-0.564***	-0.560***	-0.467***	-0.0441	-0.625***	-0.609***	-0.546***
	(0.0588)	(0.0416)	(0.0680)	(0.0515)	(0.0500)	(0.0584)	(0.0612)	(0.0612)	(0.0289)	(0.117)	(0.109)	(0.0934)
HH close to mine	-0.0389	0.0362	-0.146	-0.0678	-0.115	-0.0586	-0.0295	-0.0640	0.0699			
	(0.0834)	(0.103)	(0.102)	(0.0741)	(0.0714)	(0.0847)	(0.0899)	(0.0758)	(0.0490)			
Pregnancy and infancy * HH close (DiD)	-0.261**	-0.187*	-0.150	-0.257**	-0.239**	-0.241**	-0.267**	-0.238**	-0.0612			
	(0.109)	(0.107)	(0.122)	(0.110)	(0.110)	(0.113)	(0.112)	(0.117)	(0.0613)			
Placebo - HH in lowest wealth quintile										-0.0708	-0.179	-0.172
										(0.126)	(0.126)	(0.104)
Pregnancy and infancy * placebo										0.0199	0.0217	-0.0369
										(0.144)	(0.128)	(0.112)
Observations	5,094	5,094	4,012	5,911	6,618	4,976	4,797	4,089	4,976	1,950	2,582	3,975
R-squared	0.047	0.049	0.039	0.048	0.044	0.047	0.045	0.038	0.149	0.033	0.031	0.032
Number of group effects	167	557	166	169	169	165	162	155	165	88	120	158
Group effects	State*Year	Mine*Year		State*Year			State * Year		State*Year		State*Year	
Years postpartum included	Three		Two	Four	Five		Three		Three		Three	
Additional controls	No			No		SES	Height-for-age	Improved delivery	No		No	
Placebo restrictions										State-year in sample, at least 40km from mine	Country-year in sample, at least 40km from mine	Country-year in sample, at least 20km from mine

Standard errors clustered at the state level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 12 - Cross-sectional effects on children's Hgb

	Hematotoxic effects on children							
	Altitude-adjusted hemoglobin (g/dL)				Anemia			
	All HHs		Never-movers		All HHs		Never-movers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HH close to mine	-0.0688	-0.0920**	-0.108	-0.137*	0.0154	0.0205*	0.0413	0.0293
	(0.0461)	(0.0388)	(0.0885)	(0.0803)	(0.0141)	(0.0115)	(0.0270)	(0.0191)
N	18,070	18,070	5,521	5,521	18,428	18,428	5,585	5,585
R-squared	0.065	0.065	0.068	0.066	0.051	0.051	0.053	0.051
Mines*Year	907		660		908		662	
State*Year		317		220		318		221

Standard errors clustered at the state or mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 13 - Panel effects on children's Hgb

Hematotoxic effects on children - various exposure definitions

	Altitude-adjusted hemoglobin (g/dL)								Anemia							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
HH close to mine	-0.171	-0.230	-0.202	-0.410**	0.0589				0.0736*	0.0651	0.0695	0.122**	0.0264			
	(0.167)	(0.160)	(0.172)	(0.166)	(0.242)				(0.0379)	(0.0437)	(0.0423)	(0.0484)	(0.0721)			
Mine operating during pregnancy	-0.0369		0.117			-0.175		-0.297	0.0120		-0.00275			-0.0212		0.0791
	(0.124)		(0.136)			(0.211)		(0.207)	(0.0468)		(0.0467)			(0.0634)		(0.0595)
Mine operating during pregnancy * HH close	-0.0241		-0.222			-0.118		0.00322	-0.0217		-0.0343			0.176		0.112
	(0.178)		(0.205)			(0.541)		(0.603)	(0.0415)		(0.0789)			(0.152)		(0.136)
Mine operating in birth year		-0.166	-0.241				0.154	0.268		0.0275	0.0262				-0.186***	-0.216***
		(0.133)	(0.153)				(0.242)	(0.231)		(0.0430)	(0.0484)				(0.0616)	(0.0645)
Mine operating in birth year * HH close		0.0552	0.238				-0.184	-0.252		-0.0105	0.0174				0.0994	0.0918
		(0.172)	(0.219)				(0.470)	(0.528)		(0.0476)	(0.0852)				(0.176)	(0.184)
Child consistently exposed				-0.303								0.0972				
				(0.292)								(0.107)				
Child consistently exposed * HH close				0.163								-0.0464				
				(0.180)								(0.0527)				
Mine operating in survey year					-0.180								0.0537			
					(0.160)								(0.0551)			
Mine operating in survey year * HH close					-0.323								0.0474			
					(0.257)								(0.0827)			
Observations	5,015	5,015	5,015	3,888	4,695	5,015	5,015	5,015	5,092	5,092	5,092	3,944	4,772	5,092	5,092	5,092
R-squared	0.128	0.129	0.129	0.139	0.138	0.189	0.189	0.190	0.098	0.098	0.098	0.108	0.104	0.141	0.147	0.148
Number of fixed effects	141	141	141	129	132	4,088	4,088	4,088	142	142	142	130	133	4,144	4,144	4,144
Fixed effects	Mine					Mom			Mine					Mom		
Time effects	Country-birth year dummies					Country linear trends			Country-birth year dummies					Country linear trends		
Other controls	Quadratic in mom age at birth; child age dummies, urban/rural dummy.					Quadratic in mom age at birth; child age dummies.			Quadratic in mom age at birth; child age dummies, urban/rural dummy.					Quadratic in mom age at birth; child age dummies.		

Standard errors clustered at the mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 14 - Cross-sectional hematotoxic effects on children near different mine types

Hematotoxic effects on children near different mine types

	Children's hemoglobin (g/dL) and asset index								Anemia			Falsification tests		
	Hgb (g/dL)	Asset index	Hgb (g/dL)	Asset index	Hgb (g/dL)	Asset index	Hgb (g/dL)	Asset index	Anemia	Anemia	Anemia	Diarrhea	Cough	Fever
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
HH close to mine	-0.0688	0.188***	-0.0457	0.176***	-0.0516	0.201***	-0.0505	0.181***	0.0110	0.0153	0.0112	0.00747	-0.00372	0.00248
	(0.0461)	(0.0529)	(0.0516)	(0.0643)	(0.0509)	(0.0540)	(0.0574)	(0.0649)	(0.0148)	(0.0157)	(0.0154)	(0.00743)	(0.0117)	(0.00925)
HHs close to a 'heavy metal' mine (DiD)			-0.108	0.0537			-0.0126	0.0876	0.0216		-0.00967	0.000953	0.0515**	0.00477
			(0.109)	(0.110)			(0.119)	(0.116)	(0.0417)		(0.0389)	(0.0138)	(0.0220)	(0.0192)
Child in infancy					-0.419***	0.0474*	-0.417***	0.0301		0.144***	0.146***	0.0427***	0.0173***	0.0204***
					(0.0594)	(0.0272)	(0.0647)	(0.0240)		(0.0163)	(0.0187)	(0.00580)	(0.00609)	(0.00690)
HH close to mine and child in infancy					-0.0881	-0.118**	0.0503	-0.0538		-0.00749	-0.0505*	0.0116	-0.00826	-0.00819
					(0.100)	(0.0557)	(0.109)	(0.0548)		(0.0261)	(0.0293)	(0.0134)	(0.0120)	(0.0134)
Nearest mine (< 20 Km)) is a 'heavy metal' mine and child in infancy							-0.00707	0.0682			0.0114	-0.0185	0.00319	0.00618
							(0.155)	(0.0801)			(0.0345)	(0.0113)	(0.0119)	(0.0126)
HH close to a 'heavy metal' mine and child in infancy							-0.563***	-0.247*			0.146**	-0.0207	-0.0250	-0.0125
							(0.197)	(0.146)			(0.0586)	(0.0242)	(0.0261)	(0.0236)
N	18,070	12,713	18,070	12,713	18,070	12,713	18,070	12,713	18,428	18,428	20,278	61,671	60,406	59,592
R-squared	0.065	0.119	0.065	0.119	0.012	0.120	0.013	0.120	0.051	0.013	0.012	0.002	0.000	0.000
Mines	907	901	907	901	907	901	907	901	908	908	911	1,510	1,503	1,384

Standard errors clustered at the mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 15 - Cross-sectional effects on growth outcomes in children

	Growth outcomes in children					
	Height for age		Stunting		Severe stunting	
	All HHs	Never-movers	All HHs	Never-movers	All HHs	Never-movers
	(1)	(2)	(3)	(4)	(5)	(6)
HH close to mine	0.0828**	-0.00113	-0.0152*	-0.0140	-0.00768	-0.00198
	(0.0397)	(0.0498)	(0.00860)	(0.0137)	(0.00610)	(0.00769)
N	40,653	16,982	40,653	16,982	40,653	16,982
R-squared	0.060	0.050	0.033	0.030	0.013	0.013
Mines*Year	1,244	1,042	1,244	1,042	1,244	1,042

	Birth weight (g)		Low birth weight		*Height for age (adult women)*	
	All HHs	Never-movers	All HHs	Never-movers	*All HHs*	*Never-movers*
	(7)	(8)	(9)	(10)	(11)	(12)
HH close to mine	6.173	-15.43	-0.00879	0.000516	*0.0212*	*0.0325*
	(11.94)	(19.50)	(0.00832)	(0.0109)	*(0.0257)*	*(0.0301)*
N	37,032	13,274	37,032	13,274	*79,224*	*33,787*
R-squared	0.004	0.004	0.001	0.002	*0.007*	*0.013*
Mines*Year	1,428	1,156	1,428	1,156	*1,218*	*1,093*

Standard errors clustered at the mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 16 - Cross-sectional growth effects on children near different mine types

Growth effects in children near different mine types

	Height for age								Birth weight					
	Height for age		Stunting		Severe stunting		Asset index for height-for-age sample		Birth weight (g)		Low birth weight		Asset index for birth weight sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
HH close to mine	0.0760*	0.0607	-0.0180*	-0.0122	-0.0105	-0.00877	0.108***	0.118**	2.811	-3.117	-0.00619	-0.00463	0.0797*	0.0715
	(0.0425)	(0.0472)	(0.00967)	(0.0105)	(0.00667)	(0.00748)	(0.0406)	(0.0472)	(12.58)	(14.06)	(0.00846)	(0.0101)	(0.0471)	(0.0573)
HH close to a 'heavy metal' mine (DiD)		0.0679		-0.0258		-0.00690		-0.0527		26.85		-0.00757		0.0390
		(0.104)		(0.0251)		(0.0164)		(0.0930)		(32.30)		(0.0153)		(0.0906)
Child in infancy	0.759***	0.737***	-0.165***	-0.153***	-0.0699***	-0.0663***	0.00319	0.0124	-10.89	-11.37	0.000691	0.00178	-0.0206	-0.0143
	(0.0366)	(0.0434)	(0.00989)	(0.0110)	(0.00552)	(0.00658)	(0.0125)	(0.0148)	(6.917)	(7.763)	(0.00390)	(0.00437)	(0.0128)	(0.0148)
HH close to mine and child in infancy	0.0346	0.0761	0.0107	-0.00549	0.0121*	0.00500	-0.0339	-0.0397	15.10	24.21	-0.0114	-0.0131	-0.0220	-0.0270
	(0.0532)	(0.0634)	(0.0128)	(0.0135)	(0.00731)	(0.00824)	(0.0296)	(0.0354)	(15.58)	(18.86)	(0.00951)	(0.0111)	(0.0287)	(0.0344)
Nearest mine (< 20 Km) is a 'heavy metal' mine and child in infancy		0.0948		-0.0550**		-0.0156		-0.0370		2.230		-0.00557		-0.0301
		(0.0726)		(0.0216)		(0.0107)		(0.0272)		(17.08)		(0.00984)		(0.0287)
HH close to a 'heavy metal' mine and child in infancy		-0.169		0.0681**		0.0290*		0.0282		-34.08		0.00787		0.0255
		(0.111)		(0.0328)		(0.0165)		(0.0642)		(33.82)		(0.0220)		(0.0634)
N	40,653	40,653	40,653	40,653	40,653	40,653	28,585	28,585	37,032	37,032	37,032	37,032	27,466	27,466
R-squared	0.058	0.058	0.031	0.032	0.013	0.013	0.084	0.085	0.004	0.004	0.001	0.001	0.081	0.081
Mines	1,244	1,244	1,244	1,244	1,244	1,244	1,243	1,243	1,428	1,428	1,428	1,428	1,425	1,425

Standard errors clustered at the mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 17 - Panel effects on growth outcomes in children

	Growth effects in children																			
	Height-for-age										Stunting									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
HH close to mine	0.188**	0.521**	0.0894	0.136	0.0686	0.130	0.159*				-0.0483*	-0.123***	-0.0371	-0.0442	-0.0419	-0.0377	-0.0457*			
	(0.0800)	(0.206)	(0.0898)	(0.0889)	(0.106)	(0.0951)	(0.0916)				(0.0250)	(0.0471)	(0.0269)	(0.0271)	(0.0339)	(0.0309)	(0.0274)			
Mine operating during pregnancy	0.0198	0.0992		0.0968		0.0980		0.180		0.0983	0.00109	-0.0375		0.00584		-0.00894		0.0240		0.0426
	(0.0609)	(0.130)		(0.0712)		(0.0828)		(0.162)		(0.172)	(0.0174)	(0.0308)		(0.0230)		(0.0246)		(0.0662)		(0.0614)
Mine operating during pregnancy * HH close	-0.138	-0.371*		-0.517***		-0.423**		-0.473		-0.449	0.0517*	0.145***		0.0793*		0.0780		0.0214		0.0361
	(0.0990)	(0.216)		(0.137)		(0.183)		(0.352)		(0.350)	(0.0275)	(0.0507)		(0.0421)		(0.0709)		(0.125)		(0.114)
Mine operating in birth year			-0.00575	-0.0877					0.189	0.156			-0.00779	-0.00819					-0.0162	-0.0358
			(0.0594)	(0.0699)					(0.156)	(0.171)			(0.0196)	(0.0259)					(0.0587)	(0.0572)
Mine operating in birth year * HH close			-0.00437	0.446***					-0.128	-0.00917			0.0362	-0.0330					-0.0743	-0.0724
			(0.112)	(0.149)					(0.423)	(0.415)			(0.0287)	(0.0437)					(0.110)	(0.105)
Mine operating in survey year					-0.00294	-0.0966									-0.00528	-0.00179				
					(0.0528)	(0.0927)									(0.0182)	(0.0249)				
Mine operating in survey year * HH close					0.0479	0.365*									0.0343	-0.0392				
					(0.134)	(0.187)									(0.0374)	(0.0843)				
Child consistently exposed							0.0194										-0.00498			
							(0.0859)										(0.0245)			
Child consistently exposed * HH close							-0.0632										0.0448			
							(0.113)										(0.0311)			
Observations	11,654	2,429	11,654	11,654	11,176	11,344	9,828	11,654	11,654	11,654	11,654	2,429	11,654	11,654	11,176	11,344	9,828	11,654	11,654	11,654
R-squared	0.110	0.085	0.109	0.110	0.111	0.113	0.115	0.196	0.196	0.197	0.069	0.094	0.069	0.069	0.070	0.070	0.071	0.089	0.089	0.090
Number of fixed effects	200	186	200	200	188	191	189	9,427	9,427	9,427	200	186	200	200	188	191	189	9,427	9,427	9,427
Fixed effects	Mine							Mom			Mine							Mom		
Time effects	Country-birth year dummies							Country linear trends			Country-birth year dummies							Country linear trends		
Other controls	Quadratic in mom age at birth; child age dummies, urban/rural dummy.							Quadratic in mom age at birth; child age dummies.			Quadratic in mom age at birth; child age dummies, urban/rural dummy.							Quadratic in mom age at birth; child age dummies.		
Subsample	No	Infants only	No					No			No	Infants only	No					No		

Standard errors clustered at the mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 17 (ct'd)

Growth effects in children

	Severe stunting										Birth weight					
	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
HH close to mine	-0.0416***	-0.0443	-0.0325***	-0.0388***	-0.0389**	-0.0353**	-0.0495***				-13.04	-17.46	-15.99			
	(0.0109)	(0.0341)	(0.0121)	(0.0115)	(0.0169)	(0.0141)	(0.0110)				(41.86)	(40.60)	(42.45)			
Mine operating during pregnancy	-0.00125	-0.0138		-0.00663		-0.0217		-0.101**		-0.115***	60.33*		56.27	84.12		91.74
	(0.0124)	(0.0139)		(0.0152)		(0.0167)		(0.0427)		(0.0430)	(31.28)		(39.87)	(61.91)		(67.16)
Mine operating during pregnancy * HH close	0.0500***	0.0491		0.0703**		0.0866**		0.143		0.155	17.66		-8.429	-195.0		-241.7
	(0.0140)	(0.0342)		(0.0295)		(0.0361)		(0.112)		(0.110)	(44.29)		(73.02)	(133.4)		(153.4)
Mine operating in birth year			-0.000598	0.00645					-0.0213	0.0261		42.05*	6.921		18.27	-13.72
			(0.0122)	(0.0151)					(0.0374)	(0.0327)		(24.81)	(29.84)		(49.78)	(54.90)
Mine operating in birth year * HH close			0.0375**	-0.0239					0.0284	-0.0311		23.36	29.94		63.56	128.7
			(0.0154)	(0.0305)					(0.0649)	(0.0520)		(43.23)	(69.70)		(109.7)	(115.8)
Mine operating in survey year					-0.00411	0.0299										
					(0.0119)	(0.0216)										
Mine operating in survey year * HH close					0.0383**	-0.0450										
					(0.0192)	(0.0404)										
Child consistently exposed							0.00909									
							(0.0180)									
Child consistently exposed * HH close							0.0565***									
							(0.0148)									
Observations	11,654	2,429	11,654	11,654	11,176	11,344	9,828	11,654	11,654	11,654	11,313	11,313	11,313	11,313	11,313	11,313
R-squared	0.055	0.055	0.055	0.055	0.056	0.056	0.060	0.051	0.047	0.051	0.038	0.038	0.038	0.029	0.028	0.029
Number of fixed effects	200	186	200	200	188	191	189	9,427	9,427	9,427	246	246	246	9,298	9,298	9,298
Fixed effects	Mine							Mom			Mine			Mom		
Time effects	Country-birth year dummies							Country linear trends			Country-birth year dummies			Country linear trends		
Other controls	Quadratic in mom age at birth; child age dummies, urban/rural dummy.							Quadratic in mom age at birth; child age dummies.			child age dummies, urban/rural dummy.			Quadratic in mom age at birth; child age dummies.		
Subsample	No	Infants only	No					No			No			No		

Standard errors clustered at the mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 18 - Cross-sectional growth effects on children born to migrants

	Height for age				Birth weight (g)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HH moved to within 5km of a mine	0.318***	0.394***			-3.203	-15.56		
	(0.0913)	(0.100)			(38.11)	(40.51)		
Child born after move	0.0859**	0.0271	0.325***	0.171	2.981	18.26	40.14	208.3***
	(0.0367)	(0.0501)	(0.110)	(0.172)	(17.42)	(21.89)	(49.71)	(77.74)
Child born after move * HH now within 5km of mine	-0.108	-0.180*	-0.356**	-0.161	11.24	4.932	-162.5**	-193.1**
	(0.0790)	(0.0998)	(0.177)	(0.350)	(36.00)	(40.77)	(66.84)	(95.84)
Observations	9,745	7,018	9,745	7,018	6,390	4,557	6,390	4,557
R-squared	0.057	0.050	0.122	0.130	0.015	0.020	0.025	0.130
Number of fixed effects	1,021	962	8,102	6,257	876	804	5,522	4,156
Group effects	Mine*year FE		Mom FE		Mine*year FE		Mom FE	
Treatment definition	Mean assumed	Ambiguous excluded	Mean assumed	Ambiguous excluded	Mean assumed	Ambiguous excluded	Mean assumed	Ambiguous excluded

Standard erros clustered at the mine level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively. Data includes children born within no more than four years of migration. Child age is recorded in months, but residence in years. Whether a child is born after the family moved is therefore ambiguous for some children. We either assume that families who have been resident Y years moved exactly (Y*12 + 6) months ago ('Mean assumed'), or exclude all children for whom treatment status is ambiguous ('Ambiguous excluded').

Table 19 - Correlates of treatment effects

	Asset index										Women's Hgb (g/dL)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Country log GDP	-0.0928*	-0.0784		-0.0932	0.0808	0.0446	-0.0303	-0.0893*	-0.0726	-0.0606	-0.203*	-0.335**		-0.138
	(0.0493)	(0.0570)		(0.0846)	(0.0941)	(0.0878)	(0.0880)	(0.0488)	(0.0514)	(0.124)	(0.104)	(0.132)		(0.145)
Country log GDP squared		0.0296										-0.363		
		(0.0584)										(0.228)		
Sub-Saharan Africa			0.162	-0.00676									0.311*	0.150
			(0.101)	(0.177)									(0.170)	(0.251)
Rest of the World			-0.0317	-0.0862									0.0193	-0.0320
			(0.106)	(0.115)									(0.212)	(0.223)
State average years of education					-0.0794**									
					(0.0328)									
Inverse distance to coast						-1.004**								
						(0.483)								
Power line density (log)							-0.127**							
							(0.0622)							
Completed an EITI report								1.500**						
								(0.644)						
Ever participated in EITI									0.122					
									(0.0898)					
Institutional quality										-0.771				
										(0.781)				
Constant	0.184***	0.161***	0.150**	0.209**	0.604***	1.091**	0.312***	0.177***	0.119*	0.102	-0.241***	-0.0343	-0.319***	-0.277**
	(0.0430)	(0.0618)	(0.0660)	(0.0842)	(0.172)	(0.432)	(0.0770)	(0.0427)	(0.0641)	(0.0794)	(0.0806)	(0.153)	(0.110)	(0.122)
Observations	228	228	238	228	135	137	137	228	228	70	101	101	107	101
R-squared	0.015	0.017	0.015	0.019	0.043	0.031	0.030	0.039	0.023	0.023	0.037	0.062	0.034	0.042

Conventional standard errors.

Supplementary material

URL: http://www.columbia.edu/~pb2442/documents/appendix_mines.pdf

Appendix A	Technical notes
Appendix B	Variable definitions
Appendix C	Composition of asset index
Appendix D	IV results
Appendix E	Weighted cross-sectional results
Appendix F	Health investment, insurance, risk taking, and employment results
Appendix G	Additional falsification tables- Cross-section results
Appendix H	Additional falsification tables- Panel results
Appendix I	Robustness check with migrant difference in difference
Appendix J	Health economic cost calculation
Appendix K	Measurement error in mine geo-locations

Appendix D IV results

Table 1 - IV estimates of cross-sectional wealth effects

	Asset index			
	All HHs			
	FE Benchmark	IV 1	FE Benchmark	IV 2
	(1)	(2)	(3)	(4)
HH close to mine	0.0772**	0.0627	0.213**	0.333*
	(0.0388)	(0.0465)	(0.0902)	(0.172)
N	102,159	102,159	19,188	19,188
R-squared	0.158		0.122	
State*year	410	410	207	207
	Never-movers			
	FE Benchmark	IV 1	FE Benchmark	IV 2
	(1)	(2)	(3)	(4)
HH close to mine	0.052	0.0209	0.199**	0.416***
	(0.0452)	(0.052)	(0.0954)	(0.158)
N	37,065	37,065	6,800	6,800
R-squared	0.185		0.193	
State*year	315	315	170	170

Cluster bootstrap standard errors, clustered at the state level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.