Mines

The local wealth and health effects of mineral mining in developing countries *

Jan von der Goltz Prabhat Barnwal

September 12, 2015

Abstract

Do residents of mining communities face health-wealth trade-offs? We conduct the first extensive assessment of this question using micro-data from communities near about 800 mineral mines in 44 developing countries. Households in mining communities enjoy a substantial medium-term gain in asset wealth (0.3 standard deviations), but experience a ten percentage point increase in the incidence of anemia among adult women, and a five percentage point rise in the prevalence of stunting in young children. Prior evidence links both of these health impacts to metal toxicity – and in particular, exposure to high levels of lead. We observe health impacts only near mines of a type where heavy metal pollution is to be expected, and find no systematic evidence that health is affected in ways that are not specific to exposure to such pollutants. Benefits and costs are strongly concentrated in the immediate vicinity ($\leq$ 5km) of a mine. Consistent results emerge from a range of distinct identification strategies. Baseline results come from a cross-sectional fixed effects model, and mine-level and mother-level panels. An instrumental variables approach serves as a robustness check. To demonstrate that the observed health impacts are due to pollution, we develop two difference-in-difference tests tailored to the known association of certain mine types with heavy metal pollution, and to the pathophysiology of lead toxicity. Our results add to the nascent literature on health impacts near industrial operations in developing countries.

Key Words: *Mining, Health, Development* **JEL codes:** *I15, O13, Q53, Q56*

*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, Alexander van Geen, Kyle Meng, Suresh Naidu, Matthew Neidell, Cristian Pop-Eleches, Bernard Salanié, Wolfram Schlenker, Anna Tompsett, and seminar participants 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.

Online appendices: http://www.columbia.edu/%7E jcv2104/research.shtml.

1 Introduction

This paper studies the local wealth and health impacts of mineral mining in 44 developing countries. We show that, while residents living close to mines enjoy greater wealth, there is a trade-off: life near mines exacts a price in terms of specific health burdens.

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 ways to avoid them, 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, p. 12)

In the following paper, we present the first systematic empirical assessment of the micro-level trade-offs between health and wealth posed by the mining industry in developing countries. We seek to add to the very limited number of broad micro data analyses of local health impacts near *any* kind of industrial operations in developing countries. For the study of industrial pollution in poor countries, the mining and mineral processing industry is an attractive test case in that it poses particularly sharp trade-offs. Single plants generate very high value – in some instances, in the hundreds of millions or billions of dollars per year. The location of ore deposits dictates where mines open, and because of transport costs, 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. Mines and smelters therefore tend to play a conspicuous economic role. At the same time, however, they are very large polluters, and precisely because they are important sources of revenue, foreign exchange, and employment, there is a risk of weak environmental regulation and enforcement.

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, there is little empirical evidence on the local economic impact of mining, and on its effects on other dimensions of well-being. This particularly includes implications for the health of local communities: although there is an important literature on pollution near mines, and an extensive body of knowledge on the toxic properties of common pollutants, there is scant systematic evidence linking the two. No more than a handful of case studies have carefully assessed the actual clinical consequences of exposure to a mining environment. This 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. In some places, 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, on general health, and on two specific health outcomes known to be linked to pollutants that may be present around mines in our sample, namely anemia in adults and children, and growth in young children. To study the interplay of health and wealth effects across the developing world, we compile 104 waves of Demographic and Health Surveys from 44 countries. The pooled data provide us with about 1.2m household-level records and several million individual-level observations, spanning a time period from 1986 until 2012; they also record the geo-location of each cluster of households sampled. We overlay this household data with information on the location of mining and smelting operations across the world. We use a large cross-sectional dataset of mines that records the types of minerals mined, and characteristics of the local geology; as well as two business intelligence datasets that document annual production at individual mines. Guided by prior evidence from the environmental sciences on the spatial extent of pollution around mines, we define households *within 5km* of a mine to be in its direct vicinity, and consider those to be treated. We regard households *within 5-20km* of a mine

to be in its general vicinity, and rely on those in constructing control groups. Using production data, we create pseudo-panels that enable us to compare our treatment and control groups across time, namely between years when the mine was operational, and when it was dormant.

We then construct a broad range of complementary statistical tests that rely on different control groups, and offer extensive placebo tests. In our baseline models, we estimate the effect of closeness to mines and smelters in the cross-section, and the effect of closeness and operational status in the panel. (We prefer pooling mineral mining and processing facilities, but show robustness to excluding smelters from the sample.) We weaken cross-sectional identifying assumptions by defining our control group conservatively, and by allowing for a fixed effect common to all clusters observed near the same mine, and in the same survey year. Because of the possibility of residential sorting, we argue that our cross-sectional estimates are best read as the long-run effect of mining on *communities*, much like district or county-level studies assess impacts on those units of analysis. To assess the effect of exposure to mining on individuals, we create two sets of pseudo-panels: a mine-level panel compares households observed near the same mine in different years, and a mother-level panel compares among siblings born in different years. The panels allow for common effects shared by all households observed in the same country and survey round. An instrumental variables (IV) approach further reassures us that our results are not due to endogenous choice of mine location or periods of operation; to this end, we use the location of mineral deposits and world mineral prices to instrument for the location and operational status of mines.

Beyond these standard identification frameworks, we develop two difference-in-differences tests that are tailored to prior knowledge on the toxic properties of mining pollution. Our purpose in designing these tests is two-fold. Firstly, they help bolster our claims to observing a causal impact of mining on health. In particular, we believe that they are largely immune to residential sorting. Secondly, our tests provide evidence to suggest that the observed health effects are due to pollution, not other mechanisms. To devise the tests, we (i) leverage knowledge on the association of specific mine types with lead contamination – and by extension, health impacts specific to lead – to conduct falsification tests. We show both that we only observe those health impacts that

are expected from exposure to lead pollution, and that we only observe them near mine types strongly associated with the release of lead. Furthermore, we (ii) exploit detailed information on the birth history of women to describe a pattern of impaired ability to recover from blood loss after pregnancy among women living in mining communities, as compared to those living in the general vicinity. We argue that this effect is consistent with a known pathophysiological pattern of lead toxicity in adults, but not easily consistent with other mechanisms.

Our results show that, at the global mean, long-run asset wealth in mining communities rises by about 0.1 standard deviations 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 0.3 standard deviations. We illustrate that these are considerable effects, given the high variation in asset ownership within survey rounds. Wealth effects are strongly concentrated in the direct vicinity of the mine; there are benefits across the wealth distribution, although in the long run, the wealthiest households benefit the most; across countries, wealth gains are greatest near mines where the overall economic environment is poor.

We find clear evidence of two health impacts that are known consequences of exposure to lead and other heavy metals that may be present near mines. Thus, women in mining communities show depressed blood hemoglobin, and increases in the incidence of anemia of three to ten 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 important adverse growth outcomes from *in utero* exposure, with a five percentage point increase in the incidence of stunting – although there is very little evidence of lower birth weight. Growth impacts are weaker 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. By way of contrast to these specific health impacts, we find no effects on health

outcomes that are not linked to heavy metal pollution, nor are mining communities differentially affected by other known causes of anemia, or under-served by health care.

Because our paper shows reduced form impacts (that is, it allows for the effect of exposure to a mining environment on health to play out through any channel, including greater wealth), our health results should be interpreted as the *compensated* impact of mining. By implication, since living in mining communities goes hand in hand with economic benefits across the distribution, there is at least no indication that ill health is caused by deprivation. In an illustrative cost-benefit analysis, we show that in the expectation, economic gains outweigh the cost of health impacts, if gains are sufficiently permanent. However, among those who do suffer health impacts, their cost may exceed any economic gains.

This paper seeks to make three contributions to the literature. It is the first to demonstrate that residents of mining communities in developing countries face a trade-off between real economic benefits and specific health costs. Secondly, we add to the limited evidence on the consequences of industrial pollution in developing countries. Finally, we complement the toxicological and epidemiological literature by showing that the health effects of mining pollution are salient in a study of the general population near a large number of mines (rather than in local treatment effects found in case studies), and are robust to tests that require weak identifying assumptions.

The remainder of the paper is organized as follows. Section 2 describes what is known about welfare in mining communities. Section 3 introduces results from environmental science and toxicology that guide the way we develop hypotheses, measure impacts, and interpret results. Section 4 discusses data, and Section 5 summarizes econometric methods. Section 6 presents results; Section 7 concludes.

2 Mining, wealth, and health

This section discusses the state of knowledge on mining and wealth (Section 2.1), and on health in mining communities (Section 2.2). Section 3 provides additional background on individual links

in the causal chain from mining to ultimate health impacts, namely (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.

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 economic 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 0.3 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. Productivity losses in the general vicinity 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 a 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,

[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 services, and among men, 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 (2010) 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. 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 in the general population 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 assesses the issue across many mine sites in a manner that allows for a causal interpretation of results.

Prior work in economics on the issue is 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 their recent working paper on Ghana, 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 twelve gold mines (perhaps due to air pollution around the mines studied), but no impact on stunting and diarrhea (Aragón and Rud, 2013a). 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. 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, p. 14) 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. Among school-age children living near a lead smelter in Belgium, Roels et al. (1976) find changes in sensitive biomarkers that indicate an incipient disruption of the process of blood formation, but not overt anemia. Both papers show comparisons to a matched control group in addition to dose-response relationships. Dose-response relationships alone have also been reported between blood lead (PbB) and lower blood hemoglobin (Hgb)[2], as well as reduced nerve conductivity, 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 of IQ to PbB 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, without clear causal claims.

Pollution due to mining is a special case of industrial pollution, and the latter has been analyzed in large and well-identified studies. Yet, most of these investigate developed countries (see Currie et al. (2013) for a major recent contribution); studies of developing countries – especially using large samples – remain 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 gastro-intestinal cancer rates in China; and Rau et al. (2013) show cognitive losses from lead exposure near an abandoned toxic waste site in Chile. Hanna and Oliva (2011) describe reductions in air pollution from the closure of a large refinery in Mexico city, and an associated increase in labor demand. Studies of overall urban pollution (Arceo-Gomez et al., 2012; Greenstone and Hanna, 2011) are related, but not specific to industry, while studies of air pollution from urban traffic (e.g. Gallego et al., 2013) are less closely related. We seek to contribute to this nascent literature

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

by presenting a multi-site micro-data study of the comparative health and wealth impacts of an important industry, across many developing countries.

3 Scientific background

This section first discusses environmental pollution near mines, and its relationship to the body burden of toxicants (Section 3.1). We then establish that metals, and in particular lead, are of most interest as pollutants in our sample, and discuss the toxic effects of lead (Section 3.2).

3.1 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 and smelters of different types. We base the following discussion on Alloway (2013), Ripley et al. (1996), and Wright and Welbourn (2002).

Local communities can be exposed to pollution through a multitude of channels. These include dust from mining, handling and processing; mine waste water; direct exposure to abandoned mine spoil and tailings; metals leached from tailings 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, pollutants are used in processing, such as in the case of cyanide leaching of gold, or gold and silver extraction by mercury amalgamation. Finally, sometimes the concern is with toxicants co-located with the mineral mined and released either in processing or weathering of mine spoils, such as in the case of heavy metals in non-ferrous metal mining.

Two stylized facts on pollution near mines are essential to the way we analyze the health impacts of mining.

(i) The kinds of pollutants near a given mine can be predicted well from the ore mined.

Table 1 summarizes pollutants 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 "non-ferrous metalliferous mining and smelting industries ... with very high levels of heavy metal(loid) contamination of the environment." (Alloway, 2013, p. 43) Thus, 'polymetallic' mines, where any combination of copper, gold, lead, silver, and zinc are extracted, are linked 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 4.2.1 for coding notes.) Pollution near polymetallic mines is of particular concern both because heavy metals are important toxicants, but also 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 (Salomons, 1995).

(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 been associated in the literature with distances from the point source of emissions of 0.5 to 4km. Mean blood lead levels (PbB) among children in the highly exposed communities ranged from 13 μg/dL to more than 40 μg/dL. (Table 2) All mean PbB values far exceed the reference value of 5 μg/dL (the 97.5$^{\text{th}}$ 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 this paper, we do not directly observe 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. In line with the empirical evidence reviewed above, 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 also corresponds to the extent of high-exposure buffer zones around mines in van Geen et al. (2012). It is considerably tighter than in other current studies of mining in economics, as is appropriate for our focus on health impacts.[3] A key benefit of working with our large multi-country dataset is that it allows us to restrict our treatment group in this manner, while retaining sufficient statistical power.[4]

3.2 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 important and observable in our data, and second, because a

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

[4] With perfect data, we might define closeness even more restrictively. In the context of available data, a tighter cutoff would risk introducing noise, both because of the practice of jittering cluster geolocations in our socio-economic data, and because of the fact that we work with (imperfectly recorded) mine point locations, while mining operations can measure several kilometers across.

large share of mines in our sample is associated with this type of pollution (40% of mines in the cross-section, and 70-90% in the panel, depending on definitions). Among heavy metal pollutants, lead takes a central role, because it is known that the lead body burdens previously measured near smelters (reported above) are high enough to cause health problems that we observe in our data.

3.2.1 Sequalae of lead exposure observed in our data

The toxic properties of lead have long been studied, and are well understood.[5] The wide-ranging effects on adults include reduced blood hemoglobin (Hgb) and overt anemia, cognitive defects, hypertension, and impaired renal function. In our data, we observe only one of these conditions, namely low blood Hgb and 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 use for falsification tests some health outcomes that have not been linked to lead (cough, fever), or linked only weakly (all-cause mortality), or at very high exposure (gastro-intestinal problems). Regrettably, we do not have a good measure of impaired cognitive performance and behavioral problems due to neurological damage in children. However, while the health impacts we do observe – anemia and growth deficits are known to require high blood lead, "there is no evidence of a threshold for the adverse consequences of lead exposure" for intellectual development (Lanphear et al., 2005, p. 899). Hence, demonstrating overt anemia or growth deficits implies a strong likelihood that the affected individuals – and presumably others with lower PbB – also suffer some cognitive and behavioral impairment. (Appendix Table E summarizes how the health consequences we observe affect the well-being of those exposed, and what their economic cost might be.)

[5] See ATSDR (2007) for a full discussion.

3.2.2 Hematologic toxicity of lead

Lead depresses blood Hgb levels both by shortening red blood cell life spans, and by interfering with enzymes essential to the synthesis of the heme component in hemoglobin. Enzyme activity begins to be disrupted at very low PbB, but is not measured in our data. Effects on Hgb – which we can observe – have previously been reported at high PbB levels: in excess of 40μg/dL in children, and 50μg/dL in adults (ATSDR, 2007, pp. 69, 71f). That is, we expect the hurdle to finding impacts on Hgb to be quite high.

Therefore, 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. demonstrate this effect by comparing Hgb recovery after blood donation in lead factory workers and a control group. In our study, we show that analogously, Hgb recovery is similarly impaired among women in mining communities after another kind of blood loss, namely pregnancy and delivery

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 (in the case of lead ingested with food) absorb far larger portions of lead. In the case of anemia, however, we expect effects to be *harder* to demonstrate in children than in adults. This is because, by contrast to adults, children are able to compensate for erythrocyte loss by increasing production of the hormone erythropoietin (EPO), and thus boosting the generation of red blood cells. (Factor-Litvak et al., 1998)

In summary, based on the state of scientific knowledge, we 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 levels in children.

3.2.3 Effects of lead on child growth

While there is an epidemiological link between lead and anemia, and several hematotoxic mechanisms are known, studies are in less agreement on the effect of lead on growth in children, and "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 gestational age, as well as a wide range of measures of height and weight from birth to adolescence. (ATSDR, 2007; Bellinger et al., 1991; Hernandez-Avila et al., 2002; Sanín et al., 2001; Zhu et al., 2010) However, other studies have failed to show such correlations; indeed, it is common for a study to find impacts on some dimension of growth, but not on others, with no conclusive pattern of which indices are sensitive.

In this paper, we seek to exclude both endogeneity and small-sample variation as potential sources of ambiguous results. However, while we are able to show that *in utero* exposure affects one dimension of growth (height for age), our results mirror the existing evidence in that we find no clear effects on another key measure of growth (birth weight). In addition, in our study sites, growth effects are concentrated among infants, but abate in older children. As context for this finding, we note that, while, as a stylized fact, "blood lead levels [peak] in the age range of 1 to 3 years" (Bellinger, 2004, p. 1017), there is an important earlier path of exposure, through transfer of lead from the mother's body through cord blood and breast milk. Indeed, "infants are born with a lead body burden that reflects the burden of the mother," (ATSDR, 2007, p. 223) with correlations as high as 0.8 between maternal and infant PbB (Lauwerys et al., 1978, p. 280). Finding health impacts among infants is therefore particularly plausible if there is evidence of significant maternal lead burdens.

4 Data

4.1 Socio-economic and health data

We obtain socio-economic and health data by pooling all 104 available geo-coded Demographic and Health Surveys (DHS) 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 3) Their location is shown in Figure I.

The DHS data has some notable strengths: it covers a very broad 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, hemoglobin (Hgb). These features currently make DHS an obvious choice to study health and development at the micro level across multiple countries.[6]

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 pairwise common sets would lead to tedious repetition. We seek to strike a balance, and present side-by-side comparisons for core results. (iii) Finally, we stress again that the data are cross-sectional.

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

Therefore, while our difference-in-difference tests are designed to yield evidence of causal effects, they always compare across different individuals.

Our core measure of wealth is a standard asset index computed over household durables and housing characteristics. (Filmer and Pritchett (2001); see Appendix B for details.) We base it on the largest set of wealth proxies available within each survey round, but do not include slow-moving or immutable traits of the household head, such as gender, marital status, or education.

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. 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 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 in 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.

In addition to our core outcomes, we collect data on a range of general adult and child health outcomes, on health care, sexual risk taking, nutrition, and employment and occupation. Finally, we construct infant and under-five mortality data for all children whose births were recorded in any survey module.[7]

[7]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.

4.2 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 datasets of mine output that permit us to estimate mine-level panels; and an additional dataset of mine locations that serves to ensure robustness of our findings to measurement error in geo-location. In total, we observe communities near 838 mines in the cross-section, and 515 mines in the panel, though the set of mines that enters our estimating samples is generally smaller.[8] (Table 3)

4.2.1 Cross-sectional data on mine location and characteristics

In the cross-section, we work with the United States Geological Survey's Mineral Resource Database (United States Geological Survey, 2005). It contains the point locations of a very large set of mines, legacies, deposits, and smelters (about 25,000 locations in total) across developing countries. The data records 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 sample, we include all active mines, legacies (that is, former mines that are now dormant), and smelters. We include smelters because they are often located close to mines, and it is intuitive to think of a single mineral extraction and processing chain from mining to smelting.[9] 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

[8] 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.

[9] In Appendix Table G, we show that our core results are nearly fully robust to excluding smelters. In one case (panel results on women's hemoglobin), the effect is not significant, although it is consistent in sign and approximate size.

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 minerals present in a given location to sort mines into larger groups that share the same expected pollutants and hence, the same health effects. We remove from our baseline sample all quarries (see Appendix A for a definition). 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 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 a 'heavy metal' mine if (i) lead is being mined or smelted, or (ii) lead, though not targeted for extraction, is known to be present in significant amounts, or (iii) any two of the metals copper, gold, silver, and zinc are being mined or processed. 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. For instance, among gold-producing mines, it would aim to exclude alluvial gold deposits, where gold is typically the only metal of interested, and we expect mercury contamination from processing to be the primary concern, rather than lead pollution.

4.2.2 Mine-level production data

Since the USGS data provides virtually no time variation, we draw additional information from two business intelligence firms: Infomine (2013), and IntierraRMG (2013) – for whose product we henceforth write 'RMD', for 'Raw Materials Data'. Both sources record dates of operation and annual production, 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. 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 quite comprehensive, and the mines recorded in the dataset account for a very large share of global metal production. 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 use mine geolocations from an additional dataset, Mining Atlas (2014), for three purposes. First, we add geolocations for RMD mines wherever location is missing in the original data. Secondly, we use company records and Google Earth images to investigate the small number of cases where there are very large discrepancies in location between the two sources; we discard a few records where location is plainly not recorded with any precision in either dataset. Thirdly, we use the two independent but noisy measures of location to check robustness of our results to measurement error in geolocation (see Appendix D).

4.3 Other data

For the purpose of constructing a time-varying instrumental variable, we retrieve data on mineral prices from various sources, summarized in Appendix A. In order to describe how the wealth effects of mining vary with the economic environment, we obtain country-level data on GDP and governance from the World Development Indicators; data on the efforts a given country made toward compliance with the Extractive Industries Transparency Initiative (EITI) from the Initiative's website (www.eiti.org); and state-level data on governance, geography, infrastructure, and education from Gennaioli et al. (2013).

5 Econometric Specification

5.1 Baseline treatment definition

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 pollution – which we do not observe.

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. As 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 greatly eases the stringency of identifying assumptions required for a causal interpretation of our results. The cost of working with these definitions is that we can only achieve reasonable sample size by allowing our panels to be unbalanced. We argue that this is a reasonable price to pay for the sake of working with a treatment definition that is in line with prior scientific knowledge, and a control group definition that promises to provide a credible counterfactual.[10]

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 conservatively impute inactivity – see Appendix A.) That is, we consider only extensive margin impacts of production. We do so because year-on-year variation in output is likely to be more weakly associated with health outcomes. In this, mines differ from sources of pollution studied

[10]For the study of wealth benefits alone, a natural alternative would be to study effects of mine density in (hopefully quite balanced) panels of administrative units. This would, however, vitiate the purpose of studying health effects.

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 pollute. 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 hope to do it justice by studying extensive margin variation alongside the cross-sectional 'once on, always on' measure.

5.2 Cross-sectional model

Identification in the cross-section rests on a conservative choice of control group, and restrictive group effects. Because they cannot decisively address the possibility of residential sorting, the correct way to read our cross-sectional results is 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, we believe they can be interpreted as causal; and to the degree that regional disparities matter, they are of policy interest. Our difference-in-differences models then provide evidence that impacts are unlikely to be driven by sorting, and allow us to make stronger claims about the well-being of 'people exposed to mining'.

In our baseline specification, we consider outcomes y for individuals or households i in sampling cluster j within no more than 20km of a mine, conditional on whether the cluster is *close* (within 5km) 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 clus-

tering error terms at the mine level (not the mine-year level). Wherever the outcome of interest is binary, we model it using a linear probability model.

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

Identifying assumptions would be violated if mining towns differed from neighboring communities in geography, institutions or other characteristics in ways that correlate with potential outcomes. However, differences would have to arise even compared to locations very close by, because we restrict control locations to those no more than 20km away from the nearest mine. 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, through infrastructure construction, or the emergence of institutions).

5.3 Pseudo-panel model

We have argued that our cross-sectional setup offers valid estimates of the long-run impact of mining on communities. Still, it says less than is desirable about mechanisms of treatment transmission, and due to the possibility of sorting, it does not allow us to make claims about the impact of mining on individuals. 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 ('mine-level panel'). Secondly, we compare children born to the same mother at different times ('mother-level panel'). Plainly, comparisons in each case are across different individuals.

Equations 2 and 3 describe the mine-level and mother-level models. We analyze outcomes for individuals i in cluster j at time t.

$$y_{i(t)} = \beta_1 close_j + \beta_2 operating_{j(t-\tau)} + \beta_3 close_j * operating_{j(t-\tau)} + \beta_4 X_{i(t-\tau)} + \gamma_{mine} + f(t) + \epsilon_{i(t)} \quad (2)$$

$$y_{i(t)} = \beta_1 operating_{j(t-\tau)} + \beta_3 close_j * operating_{j(t-\tau)} + \beta_4 X_{i(t-\tau)} + \gamma_{mother} + f(t) + \epsilon_{i(t)} \quad (3)$$

In Equation 2, we allow for time-invariant effects γ_{mine} for each mine, and model outcomes at time t as being conditional on whether the respondent lived in a community *close* to a mine during the time period relevant for treatment, $t - \tau$, and whether the mine was *operating* at time $t - \tau$.[11] The time periods of interest t and $t - \tau$ depend on the outcome being investigated. For instance, where we analyze height-for-age in children, the outcome is measured in the survey year t, and may be modeled conditional on exposure to mining operations during the survey year (τ = 0), the birth year (τ = age), or while the child was *in utero* (τ = age + 1). The model also includes time-specific effects $f(t)$. We believe country-year dummies are sufficiently flexible and appropriate for sample size. We use these in our baseline models, and show robustness to using different time effects. Modifications in the mother-level model are immediate (Equation 3); because we do not observe location of prior residence for migrants, no coefficient on *close* can be estimated, and because of the much smaller sample sizes in the mother-level model, we include country linear trends $f(t)$ in our baseline model.

[11] For each respondent in our sample, we only observe current residence, and how long the household has been resident there. 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 for at least τ years. (Although they may have moved to their present location at a time before $t - \tau$.)

5.4 Difference-in-differences tests tailored to the health conditions studied

For some indicators, our sample is small near mines where there is production information, so that the pseudo-panel tends to be highly unbalanced. We therefore leverage the scientific understanding of the health conditions of interest to our study to construct additional difference in differences tests. Like the pseudo-panel, they compare the impact of mining across groups that are and are not expected to show effects. However, unlike the pseudo-panel, they do not rely on the use of time-varying production data, and hence, tend to preserve sample size better. Because they each build upon a different insight into the likely nature of exposure and the organism's reaction to it, they generate distinct control groups, and hence, further "reduce the importance of biases or random variation in a single comparison group" (Meyer, 1995, p.157).

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 none is expected, as in Equation 4. (The effect of *heavy metal mine* alone is collinear with mine-year effects.)

$$y_i = \beta_1 close_j + \beta_2 heavy\ metal\ mine_j + \beta_3 close_j * heavy\ metal\ mine_j + \beta_4 X_i + \gamma_{mine-year} + \epsilon_i \tag{4}$$

Identification rests on the assumption that potential outcomes vary among those close and not close to the mine in similar ways near mines of different types. Most obviously, if wealth effects varied systematically among mine types, health results might be confounded. With respect to preference-based sorting, the assumption would be violated if respondents were aware of how mine types differ in health outcomes, and sorted accordingly. We address the issue in two ways. Firstly, we compare DiD results on health to those on wealth, and show that differences arise for health outcomes, but not wealth. Secondly, we show that there are DiD effects only on specific expected health outcomes, not general health.

Maternal Hgb recovery: Secondly, 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 stark 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 6.2.)

$$y_i = \beta_1 close_j + \beta_2 pregnant\ or\ postpartum_i + \beta_3 close_j * pregnant\ or\ postpartum_i + \beta_4 X_i + \gamma_{state-year} + \epsilon_i \tag{5}$$

5.5 IV models

Finally, we use both cross-sectional and panel IV strategies to study wealth effects. Our purpose for the IV estimates is somewhat narrow: they provide reassurance against endogenous choice of location (even within 20km) in the cross-section, and endogenous decisions to produce in the panel. However, because they do not help address residential sorting, we discuss results relatively briefly, and for wealth only – for health impacts, we instead rely on the additional DiD tests described above.

5.5.1 Cross-sectional IV

In the cross-section, to instrument for whether a cluster is within 5km of a mine, we use the dummy (Wald) instrument *deposit* that simply indicates whether there is a mineral deposit within

5km of a given cluster (Equation 6).[12] The sample is restricted to clusters within no more than 20km of a deposit.

$$\begin{cases} y_i = \beta_1 close_j + \beta_2 X_i + \gamma_{state-year} + \epsilon_i \\ close_j = \phi deposit_j + \delta_{state-year} + \eta_j \end{cases} \tag{6}$$

Because coverage of deposit locations in the cross-sectional data is very broad, we can think of our IV estimates as general population effects. Because there can be no mine without a mineral deposit, there are neither 'defiers' nor 'always-takers', and we can interpret IV estimates as the effect of treatment on the treated. (Imbens and Wooldridge, 2009) Unsurprisingly, the dummy instrument is exceedingly strong. Since the true global distribution of mineral deposits is exogenous to human activity, the instrument is also exogenous, as long as there is no preferential *prospecting* for minerals. We believe this is likely the case, since all anecdotal evidence suggests that mining companies will seek out promising deposits in virtually any location, regardless of geographic or political obstacles. We also believe that the instrument satisfies the exclusion restriction. The most likely violations would be due to topographical features such as land quality, gradient, or water availability. Because we work at small spatial scales and across many countries, potential violations are hard to test directly. Yet, since we strongly restrict our analysis in space, characteristics would have to vary systematically over small scales to cause any problems.

5.5.2 Panel IV

Our cross-sectional IV strategy extends very naturally to the panel setting, by interacting the presence of mineral deposits with world minerals prices. Our panel data does not have very high coverage of mineral deposits, but it does include some deposits that are being explored or prepared for exploitation. We adjust the panel IV sample to include such deposits. Hence, in Equation

[12] This is similar in spirit to the geographic instrument in Duflo and Pande (2007).

7, we treat the variable *deposit* that records whether cluster j was within 5km of any deposit as exogenous.

$$y_{i(t)} = \beta_1 deposit_j + \beta_2 operating_{j(t-\tau)} + \beta_3 deposit_j * operating_{j(t-\tau)} + \beta_4 X_{i(t-\tau)} + \gamma_{mine} + f(t) + \epsilon_{i(t)} \tag{7}$$

We then instrument for whether the mine was *operating*, and for the interaction of closeness and operating status, using world mineral prices *price*, and their interaction with *deposit*. (See Appendix A for a full description of the instrument.)

6 Results

6.1 Effects on wealth

Mining towns are wealthier than neighboring communities, both in the long run and the medium term

Households in mining communities are at the mean considerably wealthier in terms of asset ownership than control households. The magnitude of the cross-sectional effect at the global average is on the order of 0.11 standard deviations of the asset index. (Table 4, Column 1) In the mine-level panel, 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. (Column 3) Since survey rounds are typically about five years apart, we interpret this as a medium-term effect.

The effect size is appreciable, 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 cross-sectional 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. The panel effect is comparable to the impact on the index 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. (See Appendix B for a description of the index and for examples of factor loadings.)[13]

We argue below that, because of the spatial pattern of long-run wealth effects, the cross-sectional baseline estimate should be interpreted as a lower bound. In Appendix C, we show that our unweighted baseline estimates are smaller than estimates obtained by (i) weighting each mine equally, or (ii) weighting by estimates of the mine-year population. In Appendix D, we use two independent measures of the geolocation of mines to instrument with one distance measure for the other, and show that our baseline results likely carry substantial attenuation bias – in our preferred specification, some 18% of the estimate. Cross-sectional IV estimates yield results that are close to and not statistically different from both our baseline results, and the OLS benchmark estimated on the IV sample. Panel IV estimates are somewhat larger than the benchmark, but not significantly different. (Table 5)

We have argued that, if the object of interest is the effect of mining on household welfare, rather than on the spatial distribution of wealth, the most salient identification concern in the cross-section is residential sorting. Panel results can be presumed to be more robust, but with about five years between survey rounds, there is still the possibility that sufficiently rapid sorting could influence results. We therefore separately study results for households that report never having moved from their current location. Effects are somewhat smaller and weaker (if not significantly different) among never-movers in both the cross-section and the panel. (Table 4, Columns 2 and 4)

[13]Regrettably, the DHS surveys have no wage data, and limited coverage of employment. The sample of men living near mines in our sample for whom employment data was collected is small. In consequence, an in-depth analysis of effects on these core dimensions of welfare is not possible. In the cross-section, unemployment among men is virtually unaffected, consistent with long-run general equilibrium. As is intuitive, the sectoral share of agriculture decreases alongside ownership of agricultural land. In the panel, employment effects tend to be adverse in sign – consistent with queuing – but we caution that the estimates are noisy and not stable. (Results available upon request.) We refer the reader to Kotsadam and Tolonen (2013) for a detailed discussion of effects on women and sectoral shifts in sub-Saharan Africa.

We interpret this as limited 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.[14]

Spatial extent of the wealth effect

Wealth effects decay steeply with distance to the nearest mine. In the panel, effects are limited to those living with 5km; in the cross-section, there is a gradient in wealth up to a distance of 15-20km. (Figures II and III) That is, in the long-run, communities in the general vicinity are economically affected to some degree, although less so than those in the direct vicinity. Hence, the cross-sectional treatment effect in our baseline model is smaller than the wealth effect on the direct vicinity of mines, as compared to those living *outside* of the general vicinity, within 20-40km of a mine (0.4σ – results not shown). Conversely, it is larger than the *average* effect of living either in the direct or general vicinity of the mine, as opposed to living at 20-40km (0.05σ).

The difference in spatial patterns between the cross-section and the panel allows for a number of explanations. If both patterns are well-identified, one would argue that the discrepancy reflects the contrast between medium-term 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. We note that, 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, p. 26), 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 Aragón and Rud study a policy change that can be presumed to be very favorable for local welfare; or the fact that they consider the case of a very large mine in a region with reportedly high transport cost. In

[14]For background, we note that there is only weakly more migration in mining communities than in neighboring communities. However, in both mining and control 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.

addition, Aragón and Rud have income data available; presumably, a more sensitive measure of well-being than our asset index.

Effects on the distribution of asset wealth

Mining is associated with wealth benefits across the distribution, though in the long run, there are much higher gains for the top quantiles, and a mild increase in wealth inequality. Benefits are more evenly distributed among never-movers. The distributional pattern might, for instance, reflect slow sorting of high-income households into mining communities, or the gradual emergence of economic opportunities that are open only to a select few.

We obtain quantile regression estimates using the two-step procedure described in Canay (2011). The results suggest that closeness to mines raises long-run asset wealth quite evenly across the distribution, with effect sizes for most quantiles close to the mean effect. (Figure IV That said, the top 5-10% benefit the most, with gains about three times as large as those at the median. Gains at the top are more limited among never-movers. In the panel, if anything, benefits are progressive, and the top quantiles gain less than others (Figure V); this pattern is comparable to the distribution of income effects found in Aragón and Rud (2013b).

Secondly, we directly consider effects on a simple measure of within-cluster inequality, namely the absolute deviation of a household's asset index value from the cluster mean.[15] In the cross-section, the mean absolute deviation increases moderately among all households, by 0.03 standard deviations of the asset index, or one-fourth of the cross-sectional wealth effect. (Table 4, Column 5) There is no effect among never-movers, nor in the panel. (Columns 6-8)

Correlates of long-run effects across countries

Long-run wealth effects vary greatly across mining communities. Table 6 shows correlations of

[15]This simple index seems more appropriate than more familiar inequality indices both due to the small number of households in many clusters, and to the nature of the mean-zero standardized asset index.

mine-level wealth effects with measures of the larger economic, geographic and policy environment. Gains are greatest where the economic environment is weak, across a range of indicators – GDP, education, access to infrastructure, some dimensions of remoteness, and (directionally only) measures of institutional quality.[16] While these correlations cannot be interpreted as causal relationships, they raise the question whether the local economic effect of mining might be 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.[17] With the same caveat regarding causal interpretation, we also note that we do observe stronger wealth effects in surveys conducted in countries at a time when the country had completed a report for the Extractive Industries Transparency Initiative,[18] or (weakly) when it had participated in the EITI in any way.

6.2 Evidence of hematologic toxic 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 hematologic toxicity – blood Hgb concentrations. As argued in Section 3.2.2, we would expect most strongly to see a reduced ability to *recover* from blood loss in adults, perhaps alongside 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 weaker evidence of lower Hgb levels in children.

[16] Appendix L shows the distribution of treatment effects across world regions and countries; correlations with measures of overall development empirically supersede regional patterns.

[17] We emphasize that, because we study effects purely at the local level, the correlation between local benefits and a weak economic environment cannot be read to contradict findings from the resource curse literature. Our findings have no implications for whether, beyond the local level, resource revenue creates corrupt structures or drives Dutch disease.

[18] See www.eiti.org. The 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.'

Hemoglobin levels in adult women are strongly depressed in mining communities

In the cross-section, blood hemoglobin (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. (Table 7) Appendix L shows the distribution of mine-level effects across countries.

Panel results confirm these patterns. Point estimates are larger, with DiD coefficients of a 0.33 g/dL decrease in blood Hgb, and a ten percentage point increase in the incidence of anemia in our preferred specification. (Table 7, Columns 3 and 6) 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 high 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 on Hgb levels 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 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 any mine). The increase in the incidence of anemia is a large effect in absolute terms, though it must be seen in the context of a baseline proportion of anemic women of 36% in control locations. That is, the cross-sectional effect amounts to an 7% relative increase in incidence, and the panel effect, to a 27% 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 surprising, given that our wealth results showed a zero or weak negative effect in mining communities when the mine is not operational. (Table 4) However, it further reassures us against any concerns that geographic features, for instance altitude, might be driving cross-sectional results.

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. (ii) Secondly, we provide direct evidence of reduced ability to recover Hgb after blood loss — 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

Table 8 shows that the effect on Hgb levels of living in mining communities are statistically zero (and mildly negative) in women living near mines where there is less reason to expect heavy metal contamination. 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. (Column 3) 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). (Column 4) 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.[19] As noted (in Section 4.2.1), our definition of heavy metal mines is best thought of as a meaningful but far from perfect proxy of the presence of lead and other toxic metals. In consequence, DiD estimates are likely attenuated.

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

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. (Columns 5-6) 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.[20] We further estimate the DiD 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. (Column 7) Finally, we do not observe similar differential effects of living near a mine associated with heavy metal contamination on two general indicators of ill health among women, namely miscarriage, and grave sickness (Columns 8-9).

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

The left panel in Figure VI 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 decrease in the ability to recover Hgb after blood loss, as described in Grandjean et al. (1989) and discussed above (in Section 3.2.2), 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

[20] As a further robustness check, Appendix L demonstrates that the median difference between heavy metal mines and non-heavy metal mines is always at least weakly negative in each individual country for which sufficient mine-level estimates can be computed.

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 9 show that the DiD coefficient is negative, large (0.21 g/dL), and significant. (Column 1) That is, the difference in Hgb levels between 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, but not stable on the small sub-sample of women in the model. 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 fixed effects, as shown in Column (2). Mine-level results do not always reach significance, but are as stable as the state-level results when we include controls, vary the treatment definition, or conduct placebo tests. 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 VI suggests that differences even out only in the second year of the child's life. It seems more appropriate than longer periods because the more we extend the time window, the stronger are the identifying assumptions required. Results 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. (Results not shown.)

Alternative explanations for the pattern of Hgb recovery are harder to come by than those for cross-sectional differences in Hgb levels. Because the test uses as a counterfactual women whose most recent birth lies at most three years 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) note that Column (7) shows that there are no significant DiD effects on wealth. Secondly, (ii) the DiD effect is robust to controlling directly for the woman's height as a slow-moving wealth proxy, or for whether she gave birth in an 'improved' setting. (Columns 3 and 4). Finally, we (iii) show a placebo regression to test whether a similar recovery pattern emerges when we compare mothers in households in the bottom wealth quintile (placebo treatment) to those in the top quintile (placebo control). We generate two samples: a small sample designed to match the baseline sample particularly tightly, and a larger sample designed to allow for more power. Both placebo samples include women who are pregnant or have given birth within the past three years, and reside at least 20km away from the nearest mine. The small sample is restricted to observations in the same state-year pairs as those observed in the main model, and the large sample, to observations within the same survey rounds. As expected, Columns (5) and (6) show that women in poor households always have lower Hgb levels than those in wealthy households – but there is no indication of an adverse time pattern around pregnancy and postpartum, with placebo DiD coefficients either near zero, or with an opposite sign.

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 regarding mechanisms of treatment transmission and regarding 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, then nutritional behavior and infection rates would have to vary across distance groups in systematically different ways near metal and non-metal mines, and among pregnant and non-pregnant women – despite the fact that socio-economic outcomes do not vary in such ways. The results also provide reassurance on identification, most importantly because they are very hard to explain with sorting.

Because mine types differ in health impacts, but not in wealth and non-specific health 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.

Residents of mining communities are not differentially affected by causes of anemia other than lead exposure, do not bear a higher burden of disease unrelated to pollution, and are not under-served by health care

The high dimensionality of the DHS data allows for diverse falsification tests that could yield evidence against our contention that the observed hematologic effects are due to pollution, not other mechanisms. Across a range of tests, we find no such evidence.

Firstly, we show in Appendix F that here is no conclusive pattern in mining communities in the leading causes of anemia other than lead toxicity (nutritional iron deficiency, malaria, and intestinal worm infections). Secondly, we test whether residents of mining communities suffer ill health that is unlikely to be attributable to pollution. Significance patterns are very sparse in the cross-section, and there are no significant adverse health impacts at all in the panel (Table 10). Appendix Tables I and J show additional specifications with similarly sparse patterns.[21]

Finally, we note that residents of mining communities are at least as well off in terms of health care as those living farther afield. As appendix K demonstrates, in the long run, women are more likely to have health insurance coverage, and to give birth with some level of skilled assistance. Panel results suggest that such benefits, along with access to health care, may extend beyond the immediate vicinity of the mine. 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

[21] We find no indication of greater alcohol abuse among men or women, and at most a mild indication of increased sexual risk taking, consistent with Wilson (2012). (Results available upon request.)

panel evidence contradicts this finding. However, we mention it here because it is at odds with our otherwise consistent evidence on wealth. We note that our discussion of maternal Hgb recovery explicitly sought to exclude the potential effect of differences in maternal health care.

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 3.2.2) that children can effectively compensate for the hematologic toxicity of lead by increasing production of EPO and red blood cell production. 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). In the cross-section, we observe insignificant decreases in Hgb on the order of 0.07 g/dL (Table 11, Column 1); the effect is strongly concentrated near heavy metal mines, but the DiD coefficient is again not significant. (Column 2) The panel shows statistically insignificant losses from current exposure to mining, but is highly sensitive to changes in the treatment definition. (Results omitted.)

Next, we ask whether infants might be more strongly affected by pollution than older children. There are two reasons to expect this pattern. Firstly, we have attributed anemia among women – and particularly, pregnant women – in mining communities to lead exposure, and it is known that children are born with a lead burden mirroring that of their mothers. Secondly, it has been previously shown that compensatory over-production of EPO and Hgb in lead-exposed children does not quite start at birth, but at some point during infancy. (Wasserman et al., 1992) When we consider impacts on infants only, we find a larger but insignificant effect on Hgb (0.13 g/dL), relative to infants living farther away from the mine. (Column 4) However, the differential impact on infants near heavy metal mines (Column 5) is both significant and large. The triple-difference coefficient shows a 0.60 g/dL difference in Hgb levels, with a nearly identical and significant difference in differences between the effect on infants near heavy metal mines and infants near other mines. Falsification results show that infants near these mines are not indiscriminately less

healthy. (Columns 7-9) However, we caution that infants born in the direct vicinity of heavy metal mines tend to live in poorer households. (Column 6) As shown above, we did not find such a correlation between mine type and wealth in our analysis of hematologic toxic effects among women living near heavy metal mines. The fact that we do find it here makes it somewhat less compelling to interpret the difference among mine types as evidence that the health impacts are due to pollution.

6.3 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 and severe stunting (height more than two or three standard deviations below the age-appropriate median, respectively). We find strong evidence of lower height among children exposed to a mining environment *in utero*, but also evidence of a compensatory positive growth effect of living in mining communities after birth. Appendix H reports that we observe an effect on birth weight in the mother-level panel, but lack corroborating evidence from our other models.[22]

Without regard to time patterns of exposure, children in mining communities grow taller than their peers

In the simple cross-section, we observe *better* outcomes for height among children of less than five years of age in mining communities than in the controls. (Table 12, Column 1) This may not be surprising: growth is strongly linked with nutrition (both the mother's and the child's), and with greater wealth in mining communities, there may also be better diets. There is also no differential impact near 'heavy metal' mines. (Column 3)

[22]Prior studies have observed that adverse conditions *in utero* can impair long-run well-being without being reflected in birth weight. (Schulz, 2010)

However, the evidence is somewhat more subtle. Firstly, as Column 2 makes obvious, there is no indication of a positive effect among never-movers. Secondly, although infants are not more affected than older children when we consider *all* types of mines (Column 4), there is at least some indication of an adverse effect on infants of living near a 'heavy metal' mine. The triple-difference effects are adverse, and significant for 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, although none of these effects reach significance. (Columns 5-7) There are no significant differences between mine types in the economic status of families with infants. (Column 8)

The cross-sectional evidence alone is thus not easy to read. There clearly are growth benefits to be had for children in mining communities, and it seems obvious to connect these to the wealth increases enjoyed by residents. However, not all children appear to benefit. The question is whether this is because some children are simply left out from economic gains, or whether they suffer countervailing direct health damage. The absence of a DiD effect between mine types and of a differential effect on infants near all mines may suggest the former. Yet, the appreciable effect on infants near mines associated with heavy metals points toward the latter. Similarly, the difference between never-movers and the general population is consistent with lower economic benefits among never-movers. (See Table 4.) However, since the differential in wealth effects is not very large, it is reasonable to note that children born to never-movers are also more likely to have been exposed to pollution, particularly *in utero*, through the maternal body burden of lead. We look to the panel for more conclusive evidence of the impacts of different exposure patterns.

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

Results from the mine-level panel suggest that there is an effect of mining activity on height, that the effect is chiefly due to exposure *in utero*, and that it 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' *in utero* effects 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. (Table 13, Columns 1-3) The effect on the discrete outcomes is significant; the one on the continuous measure not significant (t = 1.39), but stable. With a baseline incidence of 23% and 8%, respectively, the impact on stunting is appreciable, and the impact on severe stunting dramatic.

We next note that, in the case of the continuous index and of stunting, the effect of *in utero* exposure is larger and stronger when we estimate it for infants only. (Columns 4-6) This points either to a balancing effect – perhaps due to household wealth – in older children, or a spontaneous attenuation of *in utero* impacts with time. We shed some further light on this question by studying the effect of different exposure patterns. Thus, the estimated effect of exposure during the first year of life alone is centered near zero. (Column 7) Results when estimating *in utero* and birth-year effects jointly are more instructive. We find robust and large adverse effects of *in utero* exposure on the continuous index (0.5σ), alongside beneficial effects of birth-year exposure. (Column 8) 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. It points less toward a mere attenuation of impacts. 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 (conversely, operational status is highly serially correlated).[23] To 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 9), and then compare this estimate to the one obtained when we include also the effect of the mine operating during gestation. (Column 10)

[23]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.

Finally, when we estimate the effects of *in utero* and birth-year exposure with mother-level effects, the results match the pattern in the mine-level panel, but do not reach significance. (Columns 11-13) This is 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, there are few mothers with recorded births both while the mine was operational and while it was not operational.

7 Conclusion

We present the first systematic empirical assessment of the health-wealth trade-off facing mining communities, using micro-data from 44 developing countries. In communities in the vicinity of mines, we find 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 are consistent with exposure to lead contamination, and have previously been observed at body burdens of lead that are known also to cause cognitive deficits in children.

We obtain estimates of long-run effects from a cross-sectional fixed effects model; medium-term estimates come from mine-level and mother-level panels. We confirm our wealth results with an IV approach that uses deposit location and world mineral prices to instrument for mine locations and operating times. We then develop additional difference-in-difference tests that exploit (i) the association of certain mine types with lead pollution, and (ii) known pathological patterns of Hgb recovery in adults exposed to lead. These additional tests are intended both to allow for weaker identifying assumptions, and to demonstrate that the observed health impacts are due to pollution, rather than other mechanisms.

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. Medium-term benefits to households in communities near operating mines are larger, on the order of 0.3σ. Benefits are strongly concentrated

within the immediate vicinity (5km) of mines, and we find no asset wealth effects at all beyond some 15-20km. Wealth rises quite evenly across the distribution, with modest increases in inequality in the long run. Benefits in terms of health care may extend beyond the most direct vicinity of mines, although mining communities do at least as well as communities farther afield.

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 ten percentage points in adult women. The ability to recover hemoglobin levels after blood loss due to pregnancy and delivery is particularly impaired. There is weaker but consistent evidence of hematologic toxic effects in children. Children in mining communities are not disadvantaged in all aspects of physical growth. Yet, young children exposed to a mining environment *in utero* are 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 not persist. By way of contrast to these specific health impacts, there is no general pattern of ill health in mining communities.

We conclude by highlighting some conceptual and policy implications of our results.

Firstly, the presence of adverse compensated health impacts in a generally wealthier population poses an important question. The most straightforward explanation might be to suggest that the cost of avoiding exposure to pollution is high, perhaps due to the structure of settlements and the quality of public transport. We can speculate whether the decision on living in mining towns in developing countries might resemble less the choice of an optimal distance along a continuum, and more a discrete choice between two stark options – namely living either in relatively unpolluted communities outside of a reasonable commuting distance to the mine, or in a highly polluted but bustling community adjacent to the mine. The high spatial concentration of medium-term economic benefits is certainly consistent with such a situation, as is the fact that we observe the greatest wealth effects near mines in environments that are economically less active. An alternative explanation might point to limited information. Pollutant levels near mines vary greatly, even over

small distances (van Geen et al., 2012). Hence, contamination may not always be easily observed. In addition, the health impacts of pollution 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 are not making very decisive health investments to compensate for exposure to pollution. We also note that we find no differences in wealth across mine types, and hence, no *prima facie* indication of the kind of compensating wage differential one might expect if residents were widely aware of health risks.

Secondly, while our estimates of health cost and wealth benefits are not directly comparable without strong assumptions, we can offer some observations. Thus, (i) we have argued above that the effects of mining on asset wealth reflect meaningful differences in household welfare. Similarly, however, (ii) the cost to affected individuals of the health consequences we observe is very significant. The contemporaneous productivity loss due to anemia in adults has been estimated to be on the order of 5-17% (Horton and Ross, 2003), while the persistent economic impact of stunting can be dramatic (if childhood stunting persists through adulthood) – perhaps as large as an annual 53% loss in adult wages (Hoddinott et al., 2011). The permanent annual productivity loss due to lead-induced cognitive deficits expected at levels of PbB associated with overt anemia or stunting may be on the order of 1.6-13%.[24] (See Appendix E.) At the same time, (iii) it is also clear that the health burden imposed by mining pollution is very unequally distributed: at least in our compensated reduced-form estimates, relatively small population groups are affected. In consequence, the *expected* cost of health impacts is far more modest than the steep individual cost on those afflicted. However, (iv) the cost-benefit balance tilts dramatically toward costs if economic gains are less than permanent, or if legacy effects of pollution after operations cease outlast economic benefits. This is because the health cost due to cognitive losses and stunting is permanent (and the cost due to anemia may be persistent if there are adverse legacy effects of pollution after operations cease).

[24] While we do not find strong evidence of an effect of mining on the prevalence of other health conditions recorded in our data, mining communities may obviously suffer health impacts – or enjoy health benefits – that we do not observe.

In consequence, we can conclude that the decision to live in mining communities is a risky choice. Whether it is rational depends on whether economic benefits are sufficiently persistent. Furthermore, while we have shown that economic gains are quite equally distributed, the *net* benefits of mining look to be very unequally distributed. Thus, mining makes winners and losers not only between communities that benefit and communities that suffer consequences, but also *within* mining communities.

From a policy perspective, our evidence suggests that – on the global average – residents of mining communities can expect to benefit from the industry. (This is of course not to say that there are not instances of egregious local environmental damage and gross welfare loss.) Still, the presence of a health externality due to normal operations at mines in our sample that is observable in compensated health outcomes suggests that the management of mining pollution deserves renewed scrutiny. Our results yield two leads as to what effective interventions might look like. One, health concerns are most acute in the immediate vicinity of mines. Proven but expensive engineering solutions to contain and remediate pollution therefore might deserve a second look. Similarly, policy approaches need not be too broad in spatial scope to allow residents to live away from the worst pollution, while still working in or near the mine. At least for some countries in our sample, there may be a case for experimentation with programs to improve public transport, road infrastructure, or flexibility in local housing markets. Secondly, the highly uneven distribution of damages may imply that there is a premium on interventions that reduce risk. We note that the uneven distribution of costs mirrors the great spatial variation in pollution around mines described in van Geen et al. (2012), and it is tempting to posit that it might be causally related. If so, then *testing* of pollution levels in residential areas might enable residents to avoid the most dangerous sites, at a comparatively low cost.

COLUMBIA UNIVERSITY

COLUMBIA UNIVERSITY

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.

Arceo-Gomez, E. O., Hanna, R., and Oliva, P. (2012). Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. Working Paper 18349, National Bureau of Economic Research.

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.

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.

Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal*, 14(3):368–386.

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.

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.

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

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.

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.

Greenstone, M. and Hanna, R. (2011). Environmental regulations, air and water pollution, and infant mortality in India. Working Paper 17210, National Bureau of Economic Research.

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

Hanna, R. and Oliva, P. (2011). The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City. Working Paper 17302, National Bureau of Economic Research.

Hartwick, J. M., Olewiler, N. D., et al. (1986). *The Economics of Natural Resource Use*. Harper & Row.

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.

Hoddinott, J., Maluccio, J., Behrman, J. R., Martorell, R., Melgar, P., Quisumbing, A. R., Ramirez-Zea, M., Stein, A. D., and Yount, K. M. (2011). The consequences of early childhood growth failure over the life course. Technical report, International Food Policy Research Institute (IFPRI).

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

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

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.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13(2):151–161.

Mining Atlas (2014). Mining industry data and maps.

Monteiro, J. and Ferraz, C. (2009). Does oil make leaders unaccountable? 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.

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.

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).

Salomons, W. (1995). Environmental impact of metals derived from mining activities: processes, predictions, prevention. *Journal of Geochemical Exploration*, 52(1):5–23.

Sanín, L. H., González-Cossío, T., Romieu, I., Peterson, K. E., Ruíz, 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.

Schulz, L. C. (2010). The dutch hunger winter and the developmental origins of health and disease. *Proceedings of the National Academy of Sciences*, 107(39):16757–16758.

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., 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.

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

van Geen, A., Bravo, C., Gil, V., Sherpa, S., and Jack, D. (2012). Lead exposure from soil in peruvian mining towns: a national assessment supported by two contrasting examples. *Bulletin of the World Health Organization*, 90(12):878–886.

Wasserman, G., Graziano, J., Factor-Litvak, P., Popovac, D., Morina, N., Musabegovic, A., Vrenezi, N., Capuni-Paracka, S., Lekic, V., Preteni-Redjepi, E., others, et al. (1992). Independent effects of lead exposure and iron deficiency anemia on developmental outcome at age 2 years. *The Journal of Pediatrics*, 121(5):695–703.

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.

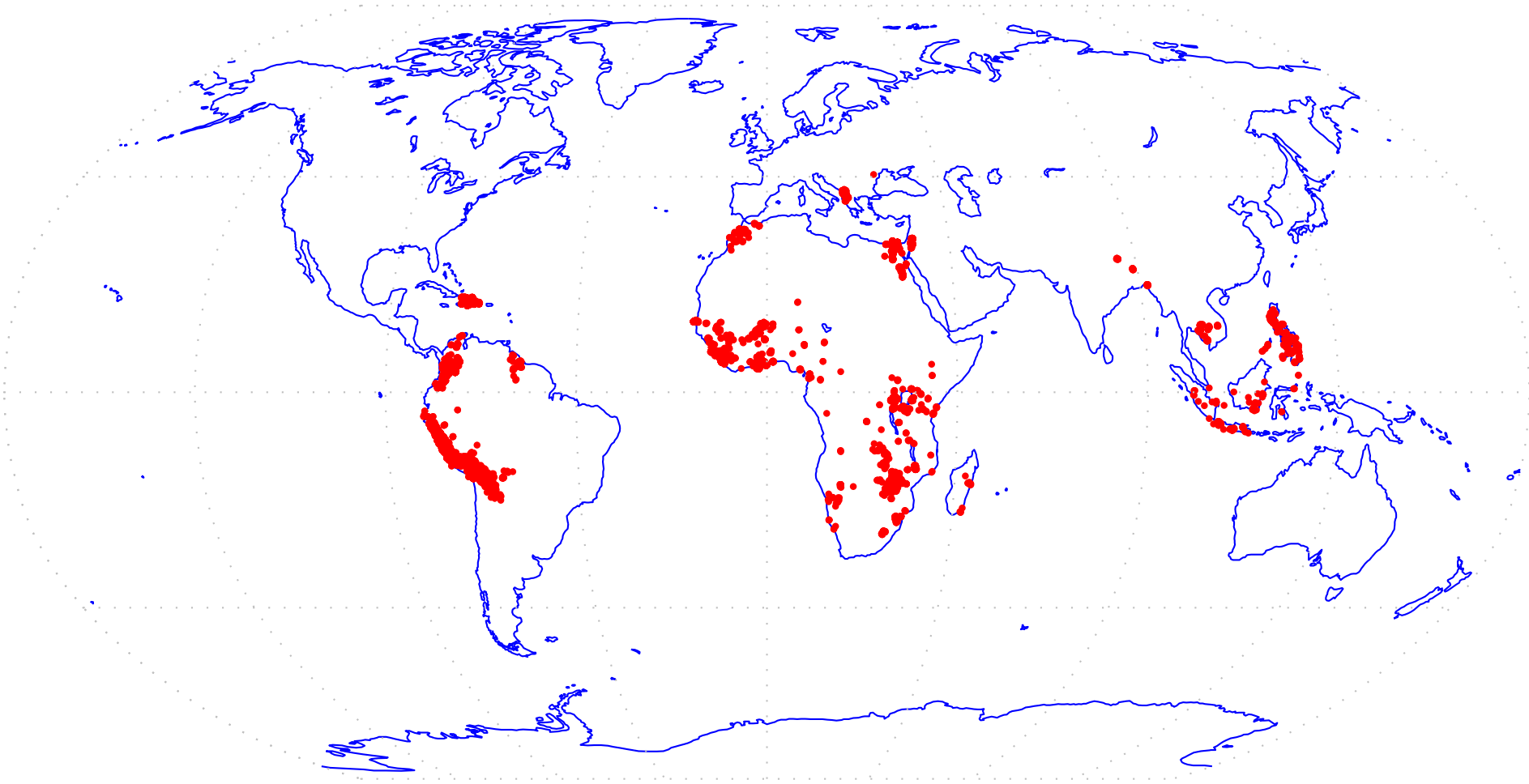

Figure I: DHS clusters within no more than 20km of a mine in the sample

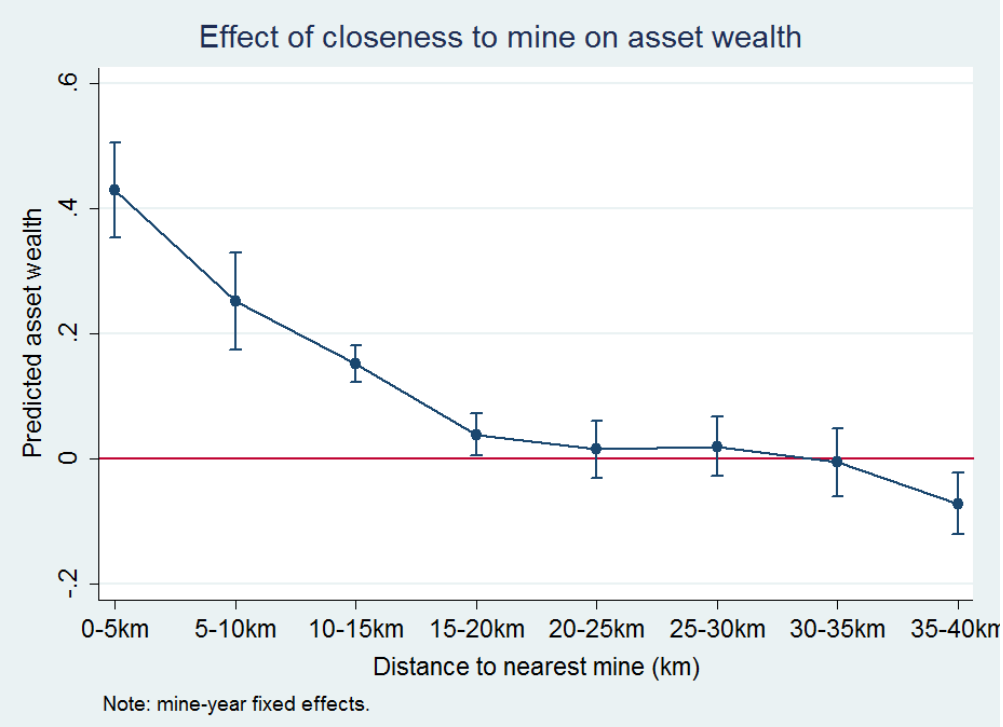

Figure II: Effect of closeness to mine on asset wealth in the cross-section

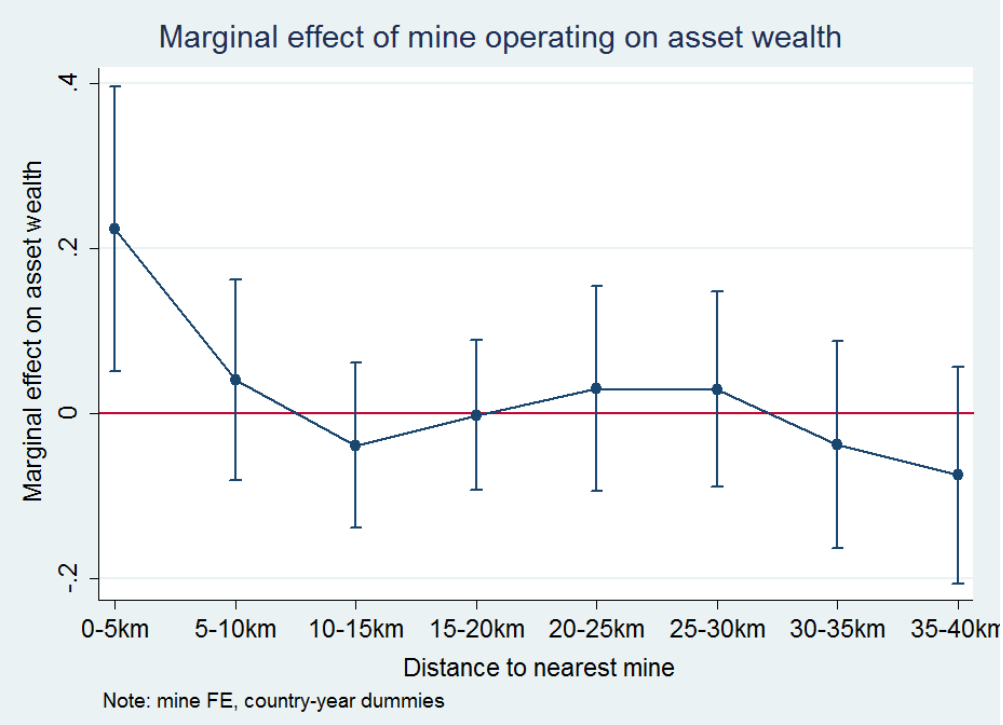

Figure III: Marginal effect of mine operating on asset wealth in the panel

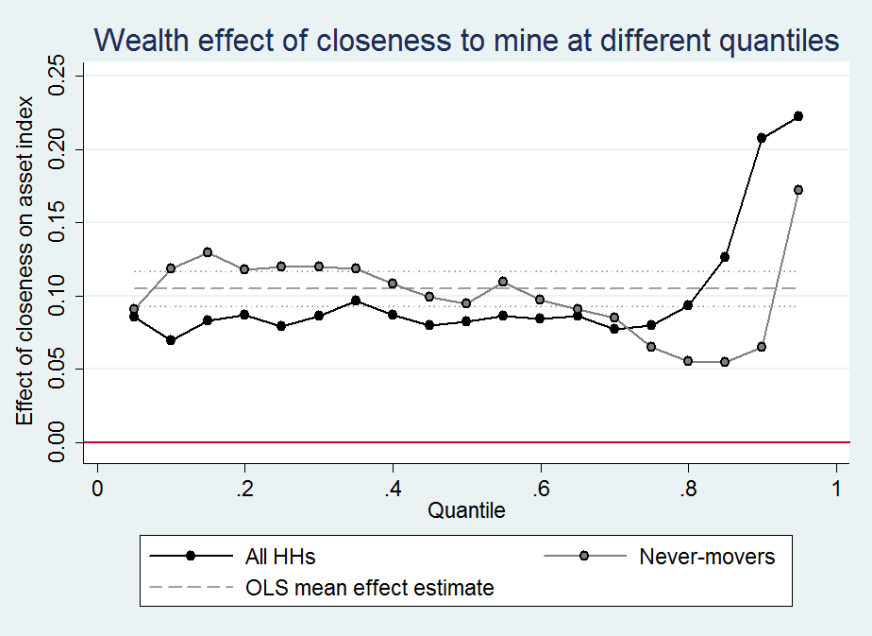

Figure IV: Cross-sectional effect of closeness to mine on asset wealth at different quantiles of the wealth distribution

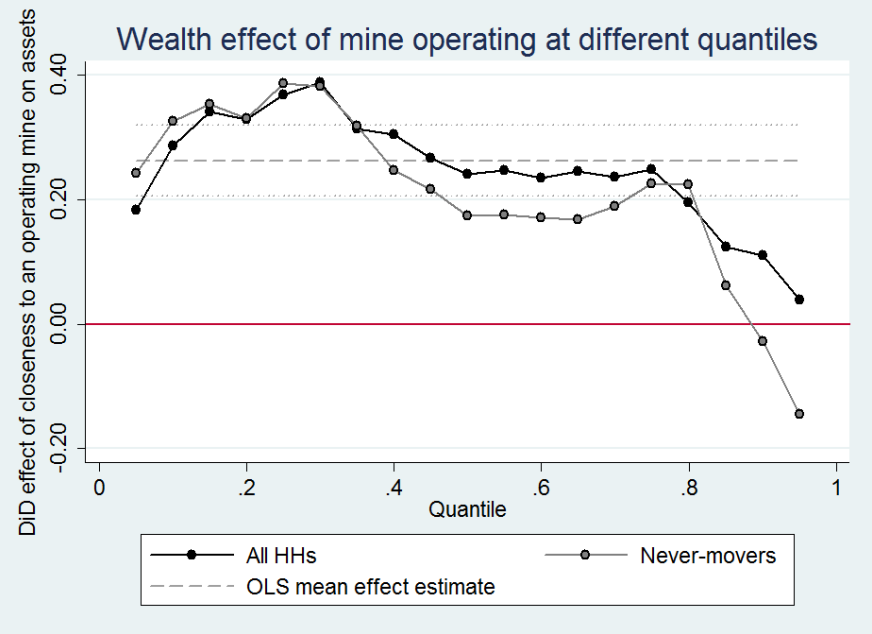

Figure V: Panel effect of mine operating on asset wealth at different quantiles of the wealth distribution

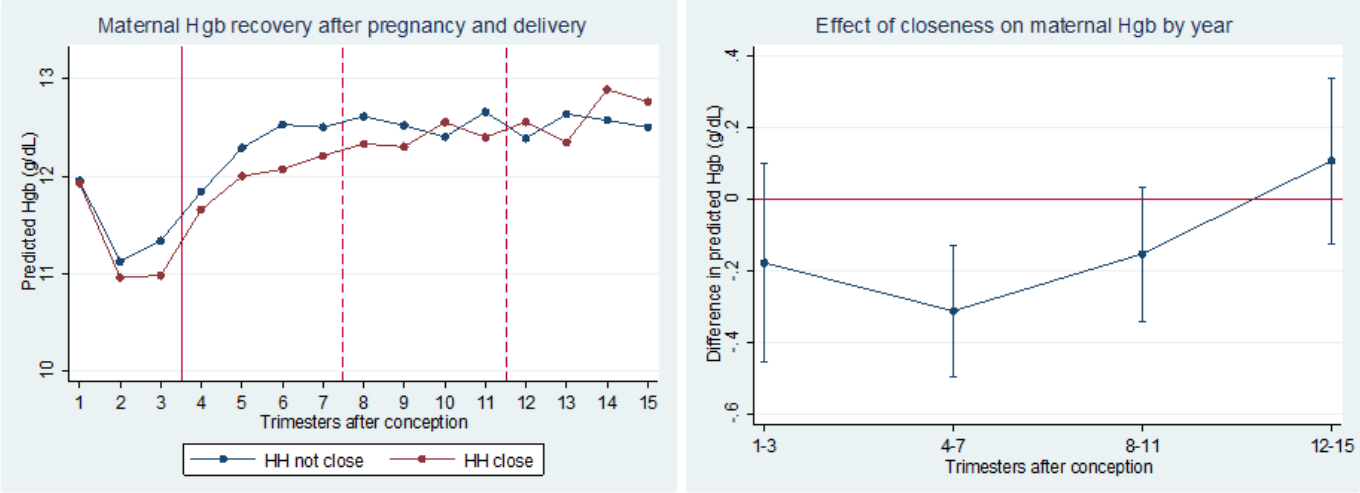

Figure VI: Effect of closeness to heavy metal mine on maternal Hgb recovery

Table 1

Mine types, associated pollutants, and health effects

Mine type	Pollutants of concern	Health effects
Polymetallic mines	Heavy metals, especially lead	Neurodevelopmental damage, anemia, growth deficits, renal disease
Small-scale gold and silver mining	Mercury	Renal disease, neurological conditions
Large-scale gold mining	Cyanide	Heart irregularities, thyroid problems
Bulk metal mines, gemstone mines	Particulates	Respiratory disease, GI problems associated with turbid water
Phosphate rock	Radionuclides	Lung cancer and non-malignant respiratory disease
Coal	Particulates, radionuclides	Respiratory disease, GI problems, lung cancer
Metal smelters	*Heavy metals, SO_2*	*As shown for polymetallic mines, and respiratory disease*

Notes. Not all mine types are mutually exclusive. Mapping based on ATSDR Toxicological Profiles for the respective pollutants, Alloway (2013), Ripley (1996), and Wright and Welbourn (2002). Health effects as reported from chronic low-level environmental exposure.

Table 2

Prior literature on blood lead levels in communities near smelters

	Distance to smelter	Mean PbB
Fontúrbel et al., 2011	0.5-1.8km	n/a
Roels et al., 1980	1-2.5km	13-30 μg/dL
Recio-Vega et al., 2012	2km	14-19 μg/dL
Factor-Litvak et al., 1999	2-4km	28-39 μg/dL
Benin et al., 1999	3km	20-40 μg/dL
Landrigan and Baker, 1981	4km	*≥ 40 μg/dL in 87% of subjects*

Notes. The table summarizes prior studies of lead levels in communities near smelters. It shows the maximum distance between the smelter and the communities considered highly exposed, alongside mean blood lead in highly exposed communities. Ranges of mean PbB refer to means for population 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. (1999), PbB was predicted from observed environmental pollution; in all other studies, PbB was measured directly.

Table 3

Sample size

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

	Number of households		
	Full sample	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

	Mines and smelters near DHS sampling clusters		
	USGS data	RMD data	Infomine data
DHS cluster within 20km	838	508	7
DHS cluster within 0-5km	339	225	4
DHS cluster within 5-20km	687	455	6
DHS cluster in both distance categories	226	172	3

Notes. Sample size based on all types of mines, smelters and legacies, excluding quarries. Not all variables used in this study are available for the entire sample. The count of locations from Infomine includes only those mines not covered in the RMD data.

Table 4
Effects on mean asset wealth and wealth disparities

	Mean asset wealth				Mean absolute wealth deviation			
	All HHs	Never-movers	All HHs	Never-movers	All HHs	Never-movers	All HHs	Never-movers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HH close to mine	**0.105*** **	**0.0784***	-0.113	-0.0352	**0.0274***	**0.00746**	-0.0451	-0.0492
	(0.035)	**(0.0423)**	(0.089)	(0.0892)	**(0.0156)**	**(0.0151)**	(0.0362)	(0.0493)
Mine operating			-0.0296	-0.0585*			-0.00611	-0.0303
			(0.0348)	(0.0354)			(0.0300)	(0.0255)
Mine operating * HH close (DiD)			**0.262*** **	**0.173***			**0.0384**	**0.0223**
			(0.0958)	**(0.0897)**			**(0.0380)**	**(0.0505)**
Number of households	90,319	31,079	22,579	9,459	90,319	31,079	22,579	9,459
Number of groups	1,562	1,371	218	205	1,562	1,371	218	205
R-squared	0.094	0.081	0.13	0.141	0.010	0.004	0.029	0.034

Notes. The table reports estimates of equations (1) and (2). Cross-sectional estimates in columns (1-2) and (5-6) use indicator variables for each mine-year pair as fixed effects. Panel estimates in columns (3-4) and (7-8) use mine-level indicators for area fixed effects, and country-year indicators for time effects. The dependent variable in Columns (1-4) is the asset factor index, with units expressed in standard deviations. In Columns (5-8), it is the absolute deviation of a household's asset factor index from the sampling cluster mean, in units of standard deviations. Controls include a quadratic in the household head's age and an indicator for urban/rural status. Columns marked 'Never-movers' restrict the sample to households with at least one respondent who had always been resident in the current location at survey time. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 5
Instrumental variables estimates of the effect on asset wealth in the cross-section and in the panel

	Cross-section		Panel				
					IV robustness checks		
	OLS benchmark	IV	FE benchmark	IV	Full sample	Baseline with all deposits	Baseline with smelters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HH close	0.0758**	0.0925**	-0.191**	-0.288*	-0.242	-0.284	-0.530**
	(0.0322)	(0.0448)	(0.0962)	(0.153)	(0.167)	(0.182)	(0.257)
Mine operating in survey year			-0.0514	0.115	0.267	0.101	-0.00301
			(0.0442)	(0.150)	(0.251)	(0.146)	(0.172)
HH close * mine operating in survey year			**0.370*** (0.110)**	**0.524** (0.208)**	**0.465** (0.232)**	**0.653* (0.396)**	**0.728*** (0.281)**
Number of households	126,434	126,434	14,671	14,671	14,735	20,593	19,780
Number of groups	618	618	187	187	188	253	200
R-squared	0.181		0.157				
First-stage F statistic		77.45		10.5	2.54	7.33	3.28
Mine-level first stage relationship		0.686***		0.215***	0.174***	0.217***	0.189***
		(0.0265)		(0.0447)	(0.0361)	(0.0445)	(0.0422)
Number of observations		1616		487	491	490	523
R-squared		0.64		0.041	0.042	0.042	0.034

Notes. The table shows cross-sectional and panel IV estimates of wealth effects, alongside fixed effects benchmark estimates computed for the IV sample, and first-stage relationships. Column (2) shows cross-sectional IV estimates following Equation (6); the instruments is an indicator recording whether there is a mineral deposit within 5km of a sampling cluster. Column (1) shows results from the cross-sectional baseline model in Equation (1), estimated on the same sample. For the panel, Columns (4-7) show IV results as in Equation (7). The instrument is an index of the world market price of minerals produced at a given mine, weighted by the preceding year's production. Column (3) shows the benchmark. The dependent variable is the asset index, expressed in units of standard deviations. The baseline panel IV sample in column (4) includes all mines and deposits that are in a stage of exploration or development. It excludes one outlier with an extremely high value of the instrument. Columns (5-7) show IV estimates using modified samples. Column (5) includes the outlier; Column (6) includes a broader range of un-mined deposits; Column (7) includes smelters, despite the poor first-stage fit. All regressions include a quadratic in the household head's age, and an indicator for urban/rural status. The cross-sectional regressions include state-year indicators, and the panel, mine fixed effects and country-year indicators. In addition to first-stage F statistics, we show simple mine-level (as opposed to household-level) OLS results on the first stage relationship between the presence of a deposit and the presence of a mine in the cross-section, and between the price index and operational status in the panel. Standard errors are clustered at the state level in the cross-section, and at the mine level in the panel. Significant at * 10%, ** 5%, *** 1%.

Table 6

Correlation of mine-level wealth effects with measures of development

Country log GDP	-0.0928*	State average years of education	-0.0794**
	(0.0493)		(0.0328)
Number of mines	228	Number of mines	135
R-squared	0.015	R-squared	0.043
State inverse distance to coast	-1.004**	State power line density (log)	-0.127**
	(0.483)		(0.0622)
Number of mines	137	Number of mines	137
R-squared	0.031	R-squared	0.030
Travel time to nearest city	0.0885	Access to land is an obstacle	0.504
	(0.0599)		(0.488)
Number of mines	137	Number of mines	66
R-squared	0.016	R-squared	0.033
State institutional quality	-0.771	Country completed an EITI report	1.500**
	(0.781)		(0.644)
Number of mines	70	Number of mines	228
R-squared	0.023	R-squared	0.039
Country ever participated in EITI	0.122		
	(0.0898)		
Number of mines	228		
R-squared	0.023		

Notes. The dependent variable in the table is composed of mine-level estimates of the cross-sectional effect of closeness to a mine on asset wealth, obtained by estimating equation (1) for each mine separately. (See Appendix E for details.) The regressors either record characteristics of the country in which the mine is located, during the year in which the survey was taken, or of the state in which the mine is located, during the year closest to the survey time for which data was available. All regressions also include log GDP. OLS estimates with conventional standard errors are shown throughout. We find no correlations with the World Bank's CPIA, and omit results for conciseness. Conventional standard errors. Significant at * 10%, ** 5%, *** 1%.

Table 7
Hematologic toxic effects on women in the cross-section

	Altitude-adjusted hemoglobin (g/dL)			Anemia		
	All HHs	Never-movers	All HHs	All HHs	Never-movers	All HHs
	(1)	(2)	(3)	(4)	(5)	(6)
HH close to mine	**-0.0863****	**-0.131**	0.396***	**0.0262****	**0.0495***	-0.107***
	(0.0438)	**(0.0838)**	(0.146)	**(0.0126)**	**(0.0268)**	(0.0292)
			0.0852			-0.0277
			(0.136)			(0.0309)
			-0.330*			**0.0966****
			(0.173)			**(0.0390)**
Number of women	38,217	13,506	9,845	36,225	13,204	9,845
Number of groups	934	785	122	934	784	122
R-squared	0.0001	0.001	0.007	0.0003	0.001	0.006

Notes. The table reports estimates of equations (1) and (2). Cross-sectional estimates in columns (1-2) and (4-5) use mine-year fixed effects, while panel estimates in columns (3) and (6) use mine-level indicators as area fixed effects, and country-year indicators as time effects. Columns (1-3) show effects on hemoglobin levels at survey time among adult women; columns (4-6) show results for the incidence of anemia (defined as Hgb below 12 g/dL in non-pregnant women, and Hgb below 11 g/dL in pregnant women). Controls include a quadratic in the respondent's age, and an indicator for urban/rural status. Columns (2) and (5) restrict the sample to respondents who had always been resident in the current location at survey time. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 8

Hematologic toxic effects on women near different mine types

	Benchmark		Effect near 'heavy metal' mines		Additional interactions		Falsification tests		
	Hgb (g/dL)	Asset index	Hgb (g/dL)	Anemia	Hgb (g/dL)		Asset index	Miscarriage	Grave illness
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HH close to mine	-0.0863**	0.145**	-0.0317	0.0123	-0.161*	-0.0285	0.140*	0.00347	0.00366
	(0.0438)	(0.0576)	(0.0533)	(0.0155)	(0.0886)	(0.0507)	(0.0756)	(0.00519)	(0.00595)
HHs close to a 'heavy metal' mine (DiD)			**-0.192****	**0.0466***	**-0.253*****	**-0.192****	**0.0176**	**-0.00377**	**-0.00286**
			(0.0944)	**(0.0247)**	**(0.0876)**	**(0.0902)**	**(0.101)**	**(0.0109)**	**(0.00898)**
Additional interactions					Region	Pregnancy			
Number of women	38,217	25,695	38,217	36,225	38,217	36,225	25,695	117,118	11,022
Number of groups	934	932	934	934	934	934	932	1,469	151
R-squared	0.0001	0.111	0.001	0.0004	0.001	0.027	0.111	0.061	0.011

Notes. Columns (1) and (2) show estimates of equation (1), and columns (3-9) estimates of equation (4), using indicators for each mine-year pair as fixed effects. In columns (3-9), the treatment variable is interacted with an indicator recording whether there is a high expectation of environmental contamination with heavy metals at the nearest mine. Columns (1), (3), (5), and (6) show effects on hemoglobin levels at survey time among adult women. Columns (2) and (7) show effects on asset wealth in households in which respondents in the health sample live. Column (4) shows results for the incidence of anemia (defined as Hgb below 12 g/dL in non-pregnant women, and Hgb below 11 g/dL in pregnant women). Columns (8) and (9) show effects on the incidence of two health conditions not specific to lead exposure. The dependent variable in column (8) is an indicator for whether a woman of reprductive age has ever suffered a miscarriage; in column (9), it is an indicator for whether the respondent was gravely ill for three months or more in the year preceding the survey. Where the dependent variable is a health condition, controls include a quadratic in the respondent's age at survey time and an indicator for urban/rural status. Where the dependent variable is the asset index, a quadratic in the household head's age replaces the respondent's age. Columns (5) and (6) show results from including additional interactions of the treatment variable, as indicated. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 9

Recovery of maternal Hgb after birth near heavy metal mines

	Altitude-adjusted hemoglobin (g/dL)						Asset index
	Baseline	Mine-year group effects	Control for height-for-age	Control for delivery setting	Placebo treatment (small sample)	Placebo treatment (large sample)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pregnancy and infancy	-0.558***	-0.581***	-0.547***	-0.429***	-0.470***	-0.652***	-0.0644**
	(0.0421)	(0.0375)	(0.0436)	(0.0454)	(0.0771)	(0.0621)	(0.0255)
HH close to mine	-0.0277	-0.00274	-0.0143	-0.0639			0.0304
	(0.0791)	(0.101)	(0.0839)	(0.0860)			(0.0523)
Pregnancy and infancy * HH close (DiD)	**-0.253****	**-0.185***	**-0.262****	**-0.330*****			**-0.0122**
	(0.0982)	**(0.107)**	**(0.104)**	**(0.116)**			**(0.0562)**
Placebo - HH in lowest wealth quintile					-0.169*	-0.232***	
					(0.0862)	(0.0551)	
Pregnancy and infancy * placebo					-0.0231	0.0783	
					(0.0825)	(0.0538)	
Number of women	5,004	5,004	4,700	3,928	6,851	14,857	4,892
Number of groups	167	521	161	158	139	269	167
R-squared	0.045	0.044	0.043	0.037	0.028	0.031	0.125

Notes. The table shows estimates of equation (5), using indicators for state-year pairs as fixed effects in all columns except column (2). Column (2) shows results using mine-year indicators. The dependent variable is Hgb at survey time among adult women in columns (1-6), and the asset index in column (7). In all columns, the sample consists of observations near mines where heavy metal contamination is to be expected. Where Hgb is the dependent variable, the sample is restricted to women who are currently pregnant, or have given birth within the three years preceding the survey, and who are known to have been resident in the current location since conception. In column (7), the sample is restricted to households in which the women included in the regression in column (1) reside. Columns (5) and (6) give results from a placebo regression, in which the treatment variable is replaced with a placebo indicator that takes value one if the respondent's household is in the bottom wealth quintile, and value zero if it is in the top wealth quintile. In column (5), the placebo sample is restricted to women who are pregnant or have given birth in the past three years, but live in households at least 20km from the nearest mine, and are observed in state-year pairs also represented in the sample in (1). In column (6), it is restricted to women who are pregnant or have given birth in the past three years, live at least 20km from the nearest mine, and are observed in country-year pairs also represented in the sample in (1). Controls include a quadratic in the respondent's age at survey time and an indicator for urban/rural status in columns (1-6). In addition, the model in column (3) includes the respondent's height for age z-score, and that in column (4) includes an indicator for whether she most recently gave birth in an improved setting. In column (7), a quadratic in the household head's age replaces that in the respondent's age. Standard errors are clustered at the mine level in column (2), and at the state level, otherwise. Significant at * 10%, ** 5%, *** 1%.

Table 10

Health outcomes not specifically linked to heavy metal pollution

	Panel A: Child health outcomes - cross-section				
	Infant mortality	Under-five mortality	Diarrhea	Cough	Fever
	(1)	(2)	(3)	(4)	(5)
HH close to mine	**-0.00246**	**-0.00305**	**0.0112***	**0.00480**	**0.00191**
	(0.00223)	**(0.00270)**	**(0.00579)**	**(0.00963)**	**(0.00788)**
Number of children	298,373	298,373	61,567	60,305	59,494
Number of groups	1,566	1,566	1,510	1,503	1,384
R-squared	0.002	0.003	0.029	0.007	0.01

	Panel B: Child health outcomes - panel				
	Infant mortality	Under-five mortality	Cough	Diarrhea	Fever
	(1)	(2)	(3)	(4)	(5)
Mine operating in exposure period * HH close (DiD)	**-0.00499**	**-0.00819**	**0.00392**	**-0.00260**	**-0.0234**
	(0.00745)	**(0.00864)**	**(0.0299)**	**(0.0282)**	**(0.0258)**
Exposure period	In utero	In utero	Survey year	Survey year	Survey year
Number of observations	43,057	43,057	15,325	15,449	15,576
Number of mines	259	259	236	237	230
R-squared	0.003	0.006	0.025	0.034	0.021

	Panel C: Adult health outcomes - cross-section			
	Ever miscarried	Night blindness during pregnancy	Female respondent very sick	Male respondent very sick
	(6)	(7)	(8)	(9)
HH close to mine	**0.00263**	**0.00254**	**0.00328**	**0.0120**
	(0.00460)	**(0.0104)**	**(0.00527)**	**(0.00977)**
Number of respondents	117,118	29,317	11,022	9,808
Number of groups	1,469	1,185	151	151
R-squared	0.061	0.001	0.011	0.011

	Panel D: Adult health outcomes - panel			
	Ever miscarried	Night blindness during pregnancy	Female respondent very sick	Male respondent very sick
	(6)	(7)	(8)	(9)
Mine operating in exposure period * HH close (DiD)	**-0.00236**		**-0.00845**	
	(0.0152)		**(0.0119)**	
Exposure period	Survey year	n/a	Survey year	n/a
Number of observations	29,666		4,111	
Number of mines	202		63	
R-squared	0.065		0.005	

Notes. The table reports estimates of equation (1) in the rows marked 'cross-section', and estimates of equation (2) in the rows marked 'panel'. In the latter, treatment variables are defined using the time period of exposure to pollution most appropriate to each health condition, as indicated. Only the difference in differences coefficient is reported. Cross-sectional models use indicator variables for each mine-year pair as group fixed effects; panel models, mine fixed effects and survey round dummies. The dependent variable in columns (1) and (2) is an indicator for whether a child died within the first year and the first five years after birth, respectively. In the other columns, it is an indicator for whether the respondent suffered the condition indicated - over the two weeks preceding the survey (3-5); at any point during her reproductive life (6); during the most recent pregnancy (7); or for three months or more during the year preceding the survey (8-9). Controls in columns (1-5) include an indicator for urban/rural status in all columns, a quadratic in the mother's age at birth, an indicator for gender, birth-order indicators, as well as indicator variables for the child's age (columns 3-5 only). In columns (6-9), they include an urban/rural indicator, and a quadratic in the respondent's age at survey time. In cells marked 'n/a', the model could not be estimated. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 11
Hematologic toxic effects on children

	Effects on all children under five years of age		Effects on infants					
	Cross-section	Mine-type DiD	Cross-section	Mine-type DiD	Mine-type DiD falsification tests			
	Hgb (g/dL)	Hgb (g/dL)	Hgb (g/dL)	Hgb (g/dL)	Asset index	Diarrhea	Cough	Fever
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HH close to mine	-0.0653	-0.0418	-0.0487	-0.0464	0.179***	0.00739	-0.00328	0.00196
	(0.0460)	(0.0516)	(0.0505)	(0.0569)	(0.0647)	(0.00732)	(0.0116)	(0.00923)
HHs close to a 'heavy metal' mine (DiD)		-0.109		-0.0184	0.0887	-0.000835	0.0492**	0.00500
		(0.108)		(0.119)	(0.115)	(0.0139)	(0.0216)	(0.0192)
Child in infancy			-0.358***	-0.359***	0.0318	0.0339***	0.0107*	0.0132*
			(0.0598)	(0.0651)	(0.0240)	(0.00579)	(0.00613)	(0.00699)
HH close to mine, and child in infancy			-0.0823	0.0643	-0.0175	0.0120	-0.00890	-0.00822
			(0.102)	(0.110)	(0.0617)	(0.0134)	(0.0120)	(0.0134)
Nearest mine (≤ 20km) is a 'heavy metal' mine,				0.00959	0.0850	-0.0198*	0.00249	0.00593
and child in infancy				(0.155)	(0.0750)	(0.0112)	(0.0119)	(0.0126)
HH close to a 'heavy metal' mine, and child in infancy (triple difference)				**-0.597*** (0.199)**	**-0.309** (0.147)**	**-0.0163 (0.0240)**	**-0.0212 (0.0262)**	**-0.0122 (0.0234)**
Number of children	18,029	18,029	18,029	18,029	12,697	61,567	60,305	59,494
Number of mines	907	907	907	907	901	1,510	1,503	1,384
R-squared	0.068	0.068	0.021	0.022	0.120	0.006	0.002	0.002
DiD effect on infants				**-0.616*** (0.201)**	**-0.221 (0.158)**	**-0.0171 (0.0232)**	**0.028 (0.0283)**	**-0.0072 (0.0261)**

Notes. Columns (1) and (3) show estimates of equation (1); columns (2) and (4-8) show estimates of equation (4). In column (3), the treatment variable in equation (1) is interacted with an indicator for whether the child was in her first year of life at survey time. In columns (4-8), the treatment variable and its interaction in equation (4) are interacted with the indicator variable for infancy. All columns indicators for each mine-year pair as fixed effects. Columns (1-4) show effects on hemoglobin levels at survey time. Columns (5-8) show effects on asset wealth and health conditions not specific to lead exposure; in columns (6-8), the dependent variables record whether a child suffered from the respective condition in the two weeks preceding the survey. Where the dependent variable is a health condition, controls include an indicator for urban/rural status in all columns, a quadratic in the mother's age at birth, an indicator for gender, birth-order indicators, as well as indicator variables for the child's age. Where the dependent variable is the asset index, controls include a quadratic in the household head's age, and an urban/rural indicator. The row labeled "DiD effect on infants" shows the sum of the coefficients "HHs close to a 'heavy metal' mine" and "HH close to a 'heavy metal' mine, and child in infancy", and tests the hypothesis that the sum is zero. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 12
Growth effects on children in the cross-section

	All children under five years of age			Infants				
	Benchmark	Never-movers only	Effect near 'heavy metal' mines	Benchmark	Effect on infants near 'heavy metal' mines			
	Height for age	Height for age	Height for age	Height for age	Height for age	Stunting	Severe stunting	Asset index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HH close to mine	0.0828**	-0.00153	0.0754*	0.0760*	0.0598	-0.0116	-0.00884	0.111**
	(0.0389)	(0.0497)	(0.0430)	(0.0418)	(0.0464)	(0.0103)	(0.00742)	(0.0469)
HHs close to a 'heavy metal' mine (DiD)			0.0378		0.0726	-0.0263	-0.00668	-0.0384
			(0.0988)		(0.102)	(0.0241)	(0.0163)	(0.0929)
Child in infancy				0.805***	0.781***	-0.164***	-0.0717***	0.0156
				(0.0368)	(0.0432)	(0.0112)	(0.00675)	(0.0156)
HH close to mine, and child in infancy				0.0304	0.0719	-0.00459	0.00597	0.00945
				(0.0528)	(0.0631)	(0.0136)	(0.00824)	(0.0403)
Nearest mine (≤ 20km) is a 'heavy metal'					0.102	-0.0571***	-0.0175	-0.0263
mine, and child in infancy					(0.0729)	(0.0217)	(0.0108)	(0.0317)
HH close to a 'heavy metal' mine, and					**-0.170**	**0.0696****	**0.0280***	**-0.0710**
child in infancy (triple difference)					**(0.110)**	**(0.0327)**	**(0.0166)**	**(0.0693)**
Number of children	40,552	16,927	40,552	40,552	40,552	40,552	40,552	28,540
Number of groups	1,244	1,041	1,244	1,244	1,244	1,244	1,244	1,243
R-squared	0.064	0.054	0.064	0.062	0.062	0.036	0.015	0.084
DiD effect on infants					**-0.097**	**0.043**	**0.021**	**-0.109**
					(0.131)	**(0.029)**	**(0.018)**	**(0.107)**

Notes. Columns (1), (2) and (4) show estimates of equation (1), and columns (3) and (5-8), estimates of equation (4). In column (4), the treatment variable in equation (1) is interacted with an indicator for whether the child was in her first year of life at survey time. In columns (5-8), the treatment variable and its interaction in equation (4) are interacted with the indicator variable for infancy. All models use indicators for each mine-year pair as fixed effects. Columns (1-5) show effects on height for age z-scores. The dependent variable in Column (6) is the prevalence of stunting, defined as a height of two σ or more below the median; in Column(7), it is the prevalence of severe stuting, defined as a height of more than three σ below the median. Column (8) shows effects on asset wealth in households in which children in the height-for-age sample live. Where the dependent variable is a health condition, controls include an indicator for urban/rural status in all columns, a quadratic in the mother's age at birth, an indicator for gender, birth-order indicators, as well as indicator variables for the child's age. Where the dependent variable is the asset index, controls include a quadratic in the household head's age, and an urban/rural indicator. The row labeled "DiD effect on infants" shows the sum of the coefficients "HHs close to a 'heavy metal' mine" and "HH close to a 'heavy metal' mine, and child in infancy", and tests the hypothesis that the sum is zero. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 13
Comparative growth effect of in utero and birth-year exposure in the panel

	In utero exposure			In utero exposure, infants only			In utero vs. birth year exposure		In utero vs. survey-year exposure		In utero and birth year exposure, mother fixed effects		
	Height	Stunting	Severe stunting	Height	Stunting	Severe stunting	Height		Height		Height		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
HH close to mine	0.191**	-0.0501**	-0.0424***	0.535**	-0.126***	-0.0448	0.0972	0.141	0.0743	0.137			
	(0.0808)	(0.0251)	(0.0110)	(0.209)	(0.0472)	(0.0346)	(0.0902)	(0.0896)	(0.107)	(0.0963)			
Mine operating during pregnancy	0.0212	0.000151	-0.00102	0.0930	-0.0369	-0.0135		0.0950		0.0953	0.175		0.0855
	(0.0614)	(0.0178)	(0.0124)	(0.124)	(0.0307)	(0.0141)		(0.0712)		(0.0826)	(0.161)		(0.173)
Mine operating during pregnancy * HH close	**-0.136**	**0.0534***	**0.0496*****	**-0.371***	**0.149*****	**0.0512**		**-0.494*****		**-0.417****	**-0.460**		**-0.422**
	(0.0981)	**(0.0274)**	**(0.0140)**	**(0.217)**	**(0.0510)**	**(0.0346)**		**(0.139)**		**(0.185)**	**(0.361)**		**(0.359)**
Mine operating in second exposure period							-0.00443	-0.0843	-0.00185	-0.0908		0.197	0.169
							(0.0595)	(0.0692)	(0.0546)	(0.0922)		(0.154)	(0.172)
Mine operating in second exposure period * HH close							-0.00860	0.423***	0.0474	0.356*		-0.163	-0.0540
							(0.110)	(0.149)	(0.130)	(0.189)		(0.424)	(0.418)
Second exposure period							Birth year		Survey year			Birth year	
Number of children	11,629	11,629	11,629	2,426	2,426	2,426	11,629	11,629	11,155	11,321	11,629	11,629	11,629
Number of fixed effects	200	200	200	186	186	186	200	200	188	191	9,408	9,408	9,408
R-squared	0.113	0.072	0.055	0.091	0.100	0.058	0.113	0.114	0.073	0.117	0.204	0.204	0.205

Notes. Columns (1-10) report estimates of equation (2), and Columns (11-13), estimates of equation (3). Columns (1-10) use indicators for each country-year pair as time effects, and columns (11-13) use country linear time trends. In columns (1-6) and (11), treatment is defined as exposure to mining *in utero*. Columns (7-10) and (12-13) compare this to the effect of exposure during a second exposure period, as indicated. The dependent variable in columns (1), (4), and (7-13) is the height-for-age z-score. In columns (2) and (5), it is the prevalence of stunting, defined as a height of two σ or more below the median; in Columns (3) and (6), it is the prevalence of severe stuting, defined as a height of more than three σ below the median. Controls include an indicator for urban/rural status in all columns, a quadratic in the mother's age at birth, an indicator for gender, birth-order indicators, as well as indicator variables for the child's age. For consistency across models, the sample is restricted to those observations where the operating status of the mine is known both in the birth year and during gestation (this removes about 3% of observations). Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.