No pains, no gains? mining pollution and morbidity in Mongolia*

Odmaa Narantungalag[a], Syed Hasan[a], Martin Berka[a]

[a]*School of Economics and Finance, Massey University, New Zealand*

Abstract

This paper investigates the impact of mining pollution on the likelihood of feeling sick. We link a geocoded soil pollution information with five rounds of Mongolia Household Socio-Economic Survey data and employ logistic regression models to investigate the medium- to long-run health impacts. Our results indicate that living one kilometre away from mines reduces a person's probability of feeling unwell by 11 per cent. The medical expenditure also increases as a result of feeling sick. Mining pollution impacts younger children more and generally aggravates respiratory illnesses. As expected, small-scale gold mines have a larger effect on individuals' health than medium and large scale mines. Our findings suggest that tighter environmental regulations to control mining pollution can reduce the short and long-term health risks of the people living near the mining sites.

JEL-Classification: I15, O13, Q53, Q56

Keywords: Mining pollution, Health, Development, Mongolia

*We thank Steven Peolhekke, Derrin Davis, Andrea Menclova, Stephen Lippert, Maulick Jagnani and the participants at the 2020 Inaugural New Zealand Virtual PhD Workshop, 2020 Massey Business School PhD Forum, 2021 Massey University School of Economics and Finance Seminar, 18th Midwest International Economic Development Conference for their comments on an earlier draft. We also thank Bolormaa Sugar at the National Statistics Office of Mongolia and Enkhtulga Baatar for their assistance with data collection.

Email address: O.Narantungalag@massey.ac.nz (Odmaa Narantungalag)

1. Introduction

Environmental pollution is one of the leading causes of premature deaths in the world. Around nine million premature deaths caused by pollution occurred in 2015, which is 16 percent of all deaths worldwide (Landrigan et al., 2018). Poor environmental quality result in considerable health losses and diseases caused by pollution lower productivity and reduce the gross domestic products in low and middle-income countries by up to two percent annually (Greenstone and Jack, 2015; Landrigan et al., 2018; Levasseur et al., 2021). Despite its substantial contribution to creating employment and spurring economic growth, the extractive industries exert significant negative externalities to the local communities (Aragón and Rud, 2015). Respiratory diseases (Pless-Mulloli et al., 2000; Saha et al., 2011; Hota and Behera, 2015, 2016), changes in mortality and morbidity (Cordier et al., 1983; Hendryx and Ahern, 2008, 2009) and neurological and psychological deficits in children resulting from mercury neurotoxicity (Cordier et al., 2002) are some of the adverse health outcomes resulting from mining pollution.

This paper investigates the impact of pollution from mining activities on individuals' likelihood of feeling sick when they are environmentally exposed to pollution such as mercury and arsenic in soil. For example, 1.3 kilograms of mercury is dumped into the environment for every kilogram of gold produced (Harada et al., 1999). Around 40 percent of the mercury goes to tailings, soils, stream sediments, lakes, and rivers during the initial stage of gold and mercury amalgamation. The remaining 60 percent of the lost mercury is released into the atmosphere when the amalgam is burned to extract the gold (Harada et al., 1999; Van Straaten, 2000). Artisanal and small-scale mining (ASGM) is the single largest buyer of mercury in the world, consuming around 1,400 tonnes in 2011 and releasing 17 percent of annual mercury emissions to the atmosphere. Mercury is a dangerous neurotoxin that is harmful to people, especially to developing fetuses and young children, increasing the risks of damaging their brain and nervous system development and function (Telmer and Stapper, 2012; Landrigan et al., 2018). The immediate health outcomes of mining activities may seem not to be as severe as symptoms suggest because symptoms, such as coughing, come and go. However, over the long term, chronic illnesses and cancer can develop as leading causes of death in mining regions (Cordier et al., 1983; Hendryx and Ahern, 2009). Hence, one way to examine the overall impact of mining pollution on human health is to use soil pollution data as it is correlated with air and water contamination (Li et al., 2014; Landrigan et al., 2018; Levasseur et al., 2021).

More specifically, we examine whether living in a close proximity to mining sites increases the probability of feeling sick. Pollution can affect local communities through four main channels: (i) the distance from a person's residence to a pollution source, (ii) duration of exposure to pollution (e.g., months, and years), (iii) levels of pollution (above or below a threshold) and (iv) the type of pollution (e.g., lead and mercury) (Graff Zivin and Neidell, 2013). We argue that the distance from polluted mining sites has causal effect on the probability of feeling unwell. Our paper relates to Aragón and Rud (2015) who reported that nitrogen dioxide emanating from gold mining in Ghana significantly reduced farmers' productivity and increased rural poverty. We also closely follow Von der Goltz and Barnwal (2019) who found that lead contamination from mining activities increased anemia in women by ten percentage points and stunting in children by five percentage points in 44 resource-rich developing countries. However, the impact of overall mining pollution on the population's health near mining activities in developing countries remains limited.

We investigate Mongolia, which offers both soil heavy metal pollution information from mining activities and individual morbidity data from the Mongolia Household Socio-Economic Survey (HSES). Mongolia is a resource-rich lower-middle income country that received a substantial amount of foreign direct investment into the extractive industries at the onset of the commodity price boom in the early 2000s (Li et al., 2017; Doojav et al., 2017). The country is also one of the 45 countries where mercury is the dominant pollutant at artisanal and small-scale mining sites (Caravanos et al., 2013). The main hazardous heavy metal persisting in soils around mining sites are mercury, cadmium, arsenic and lead that are harmful to human health (Vandermoere, 2008; Landrigan et al., 2018).

The impact of pollution from mining activities on human health in both developed and developing countries remains underexplored. Studies such as Chay and Greenstone (2003); Neidell (2004); Currie et al. (2009a,b); Janke et al. (2009); Currie et al. (2015) explore the effects of air pollution and toxic releases from industrial plants on human health, mainly in the United States.

However, our paper is similar to Hill (2018) who found that an additional shale gas well increases the probability of low birth weight by seven percent and the likelihood of premature birth by three percent in Pennsylvania, U.S. Our study is also closely related to (Marcus, 2021) who reported that the petroleum leaks from underground storage tanks increase the probability of low birth weight and preterm birth by seven percent in Pennsylvania, Florida, and New Jersey in the U.S. However, these studies report the impact of the extractive industries on infant health.

Therefore, our study narrows the gap in the literature in a few distinct ways. First, we use the distance from a person's residential area to the nearest mining site where soil is polluted with mercury and arsenic in Mongolia to examine the impact of mining pollution on sickness. Previous studies such as Currie et al. (2009a, 2015); Rau et al. (2015); Hill (2018); Von der Goltz and Barnwal (2019); Persico and Venator (2021) used the distance between the residential area and the pollution source to identify the treatment effect of pollution in a difference-in-differences setting. Our study relies on the exact location of pollution source and pollution level to draw a causal inference between distance and sickness. We use the distance to the pollution source as the primary variable of exposure as distance captures the effect of different heavy metals originating from the same source.

Second, our analysis examines the impact of pollution on all age groups: children up to 14 years old, the working-age population and the older people above the age of 65. Pollution affects different segments of the population disproportionately, while the impacts can be lifelong (Currie et al., 2014). While many studies examine the effect of pollution on infant health (i.e., Currie et al. (2009b); Currie (2011); Currie et al. (2014, 2015)), no previous study has investigated the impact of pollution on the three groups in combination. Hanna and Oliva (2015) and Graff Zivin and Neidell (2012) examine the effect of pollution on the working-age group. Neidell (2004), Rau et al. (2015) and Persico and Venator (2021) look at the impact of pollution on children's asthma and test scores. We study the impact of pollution from mining on children's sickness.

Third, we investigate the impact of pollution from mining on various body systems. Specifically, we explore respiratory, cardiovascular, digestive, and other body systems' responses to heavy metal pollution from mining. For example, children and individuals exposed to air pollutants from mining activities have more frequent respiratory symptoms and incur higher meidcal expenditures (Charpin et al., 1988; Pless-Mulloli et al., 2000; Hota and Behera, 2015, 2016). However, the effect of pollution from from mining on other body systems is not well-known. Next, the effect of the mine scale on individuals' health outcomes is explored. While fully-licensed medium and large mines might carry out environmental rehabilitation activities required by the law, the environmental quality in developing countries is still low. On the other hand, small-scale miners care more about their consumption for today than the quality of the environment for tomorrow (Greenstone and Jack, 2015). Our final investigation looks at the impact of different minerals mined across the country. The commonly mined minerals such as gold and limestone are examined to shed light on these minerals' impact on human health. Thus, we investigate the differential impacts of mine scales and types of minerals on sickness.

We employ a unique approach to examine soil pollution's impact on individuals' health near mining sites. By assigning pollution to individuals and using a large enough and nationally representative set of household-level survey data over several survey rounds, we optimize the precision of our causal estimate of pollution on sickness (see, e.g., Graff Zivin and Neidell, 2013). Distance as the exposure variable is preferred rather than the heavy metal pollution level because multiple heavy metals originate from the same mining site. Although we assume a linear relationship between the distance and the likelihood of feeling sick, we also explore the nonlinear effects of pollution by utilizing dummy variables based on the government's pollution standards, as suggested by Graff Zivin and Neidell (2013) and implemented by Currie et al. (2009a).

We find that mining pollution is detrimental to the health of individuals living within six kilometres of a mining site. Our preferred estimate suggests that living one kilometre away from mines would have prevented 8,700 people from feeling sick and saved over $635 thousand in private and public health expenditures annually. The impact of mining pollution is more pronounced for the vulnerable groups such as younger children (aged 0-14 years). Moreover, mining pollution generally aggravates respiratory illnesses. Interestingly, while small-scale mines have a significantly larger negative effect on individuals' health than large-scale mines, gold mines increase individuals' probability of feeling sick. Our findings are important as mine-impacted communities in resource-rich developing countries bear the burden of environmental pollution from mining activities while the entire economy may benefit from those activities. The results suggest that tighter environmental regulations to reduce mining pollution can have greater health and socioeconomic benefits for the population near the mining sites.

The rest of the paper proceeds as follows. Section 2 describes the empirical strategy and the data employed in the paper. Sections 3 and 4 discuss the results and findings, respectively. Finally, Section 5 concludes and provides policy implications.

2. Methodology

2.1. Model specification

We estimate an individual's likelihood of feeling unwell in the vicinity of a mining site by a number of specifications. The health risks of long-term exposure to heavy metals increase through various factors of transmission: (i) repeated and prolonged contact with residuals in soil (e.g. farming), air (e.g. breathing dust and particles) and water (e.g. swimming and drinking) and food chain (e.g. homegrown produce) (Li et al., 2014). Therefore, we estimate the lower-bound effect of overall pollution from mining on sickness as our analysis utilizes soil pollution data. Our empirical specification follows Neidell (2004), who uses ordinary least squares (OLS) method to estimate the impact of air pollution on the number of emergency room asthma admissions. However, our model exploits the distance from an individual's residential area to the nearest mine as the primary exposure variable of pollution. Such a specification assumes that mines, which are located further away from a residential area than the nearest mines do not affect sickness. We choose the distance as it is the only variable that captures the impact of all contaminants simultaneously. Moreover, the heavy metals are highly correlated as they mostly originate from the same source, making it challenging to investigate the effect of each pollutant separately. The main model that evaluates the impact of distance to the nearest mine on sickness is specified as follows:

$$y_{ist} = \alpha + \beta ln(distance) + \lambda_s + \gamma X_{ist} + \eta_t + \varepsilon_{ist} \quad (1)$$

where y_{ist} is the outcome variable taking the value of one if an individual i was sick in the past month in sub-province s at time t, and zero otherwise. $ln(distance)$, the natural logarithm of distance from an individual's residential area to the nearest mine, is the primary variable of interest. λ_s is sub-province fixed effects, accounting for possible omitted variables and time-invariant differences in sub-provinces that could impact a person's sickness regardless of the distance to a mine. η_t is survey year-fixed effects to capture any time-varying events in the model.

Finally, we include X_{ist} representing a person's age, gender, education, household consumption per capita, and urban/rural status, to allow for within-person differences. The individual-specific variables partially control for avoidance behavior. ε_{ist} is the independently and identically distributed error term. The model evaluates an individual's probability of feeling sick by a logistic regression as it is more efficient than the linear probability and probit models. The standard errors are clustered at the household level. The main hypothesis tests that distance, $\beta = 0$, has no effect on sickness. We estimate the marginal effect at means from the main variable of interest and the other covariates in the model. The cut-off point for the distance to the nearest mine is determined at six kilometres following Romero and Saavedra (2015) and Von der Goltz and Barnwal (2019) who found that mercury and lead contamination from mines within five and 20 kilometers had significant adverse effect on health, respectively.

Furthermore, Equation 2 investigates the impact of distance to the highest pollution level on illness.

$$y_{ist} \quad = \quad \alpha + \beta_j \sum_{j=1}^{7} \ln(distance_{ji}) + \lambda_s + \gamma X_{ist} + \eta_t + \varepsilon_{ist} \tag{2}$$

where everything stays the same as in Equation 1 but $ln(distance_j)$, which is the natural logarithm of distance from an individual's residential area to the mining area where the highest level of contamination for heavy metal j is recorded. Since most heavy metals originate from the same source, including all the distances in a single model creates multicollinearity. Therefore, Equation 2 is estimated for each single pollutant.[1]

[1] The pollution level for some heavy metals of mines are above the permissible level within six kilometers. However, for some heavy metals, the highest level of pollution is below the permissible level.

We also consider the level of heavy metal contamination in the next model to assess if both distance and level of pollution matter for the nearby inhabitants' health outcomes. Therefore, Equation 3 examines the impact of heavy metal contamination level that are above the permissible level set by Mongolian Agency for Standardization and Meteorology, MASM (2019).

$$\begin{aligned} y_{ist} &= \alpha + \sum_{j=1}^{7} ln(distance_{ji}^{\beta} \times level_{ji}^{\delta}) + \lambda_s + \gamma X_{is} + \eta_t + \varepsilon_i st \\ &= \alpha + \sum_{j=1}^{7} \beta_j ln(distance_{ji}) + \sum_{j=1}^{7} \delta_j ln(level_{ji}) + \lambda_s + \gamma X_{ist} + \eta_t + \varepsilon_{ist} \\ &= \alpha + \beta ln(distance) + \sum_{j=1}^{7} \delta_j ln(level_{ji}) + \lambda_s + \gamma X_{ist} + \eta_t + \varepsilon_{ist} \end{aligned} \tag{3}$$

where everything stays the same as in Equation 1 but Equation 3 includes interaction of the distance and the contamination level for arsenic and mercury as these heavy metals exceed the precaution value/permissible level.

Finally, the causal link between pollution and sickness can be nonlinear. Equation 4 estimates the impact of pollution on illness in a nonlinear form where we include dummy variables for the heavy metals that are above the permissible levels following the methods implemented by Currie et al. (2009a).

$$y_{ist} = \alpha + \beta \ln(\text{distance}) + \sum_{j=1}^{7} \delta D_j + \lambda_s + \gamma X_{ist} + \eta_t + \varepsilon_{ist} \tag{4}$$

where everything stays the same as in Equation 1 but includes dummy variables for individuals exposed to a particular heavy metal pollution. D_j takes the value of one for individuals exposed to mercury pollution level above the action value. Although over 30 per cent of individuals are exposed to arsenic pollution, the level of pollution does not threaten living organisms and human health (see Table 1 for reference). Therefore, we include only mercury, which significantly threatens human health.

2.2. *Endogeneity issues*

Two main endogeneity issues can arise from the causal inference of soil pollution on sickness. First, pollution is endogenous to individuals' avoidance behavior (Neidell, 2004; Graff Zivin and Neidell, 2012, 2013). People respond to pollution announcements by changing their behavior and reducing their outdoor activities to avoid pollution. For example, Metropolitan Statistical Areas with a population of more than 350,000 in the United States must report the daily air quality index to the public through the local media (newspapers, radio, television, telephone messaging) when pollution levels become unhealthy for the sensitive groups (California Air Resources Board, 1990; U.S. Environmental Protection Agency, 2006).[2] This kind of information induces behavioral changes in the population exposed to air pollution and makes it challenging to assess air pollution's impact on human health (Neidell, 2004).

On the other hand, unlike air pollution, soil heavy metal pollution from mining activities is not monitored frequently to inform the public about the harms of soil pollution in developing countries such as Mongolia. The population environmentally exposed to heavy metal pollution might not be aware of the pollution if they cannot observe it (Graff Zivin and Neidell, 2013). For example, mercury vapor is odorless and colorless, making it difficult to see and smell during evaporation until it affects the body (Solis et al., 2000). Similarly, most inorganic arsenic compounds are white or colorless powders with no smell or taste (U.S. Agency for Toxic Substances and Disease Registry, 2007).[3] Therefore, deliberate avoidance behavior would not exist when public information about soil pollution is not available, and the heavy metals in soil are not readily observable (Graff Zivin and Neidell, 2013). Unless local governments and environmental agencies notify the local community, the inhabitants would not be fully aware of soil pollution's potential dangers. Because avoidance behavior is an ex-post decision, omitting avoidance behavior in the model would not invalidate the estimates. Instead, the estimates would give the lower-bound biological effect of pollution (Currie et al., 2014).

[2] U.S. Environmental Protection Agency (2006) describe the air quality index from good to hazardous based on air quality index ranging from 0 to above 301.

[3] Inorganic arsenic occurs in minerals and ores that contain copper or lead. During the smelting of these minerals, most arsenic is released into the atmosphere as fine dust. Thus, this fine dust is colorless, tasteless, and odorless.

Second, pollution is endogenous to residential sorting in which households, who can afford it, relocate to a cleaner area to permanently reduce their exposure to pollution (Graff Zivin and Neidell, 2013; Von der Goltz and Barnwal, 2019). However, residential sorting is mainly undertaken by those who are highly educated (Currie, 2011; Marcus, 2021). For example, only highly educated white mothers relocated after local newspaper coverage about leaking underground storage tanks is released in the U.S (Marcus, 2021). Besides, pollution such as toxic air emissions from industrial plants in the U.S reduced the values of houses within 0.5 miles of a plant by 11 per cent (Currie et al., 2015). Consequently, high income people will most likely not live near polluted areas compared to those who cannot afford houses with better environment and amenities. Residential sorting, therefore, makes health outcomes endogenous to socioeconomic status (Graff Zivin and Neidell, 2013). However, residential sorting also depends on the availability of information and people's awareness about the specific pollution. Moreover, households in developing countries like Mongolia would not relocate or change their behavior easily unless their economic opportunities deteriorate and livelihoods suffer due to natural disasters and loss of income (Levasseur et al., 2021). Therefore, we do not consider residential sorting as a threat to our empirical identification and specification.

There are some omitted variables that can cause endogeneity in our causal inference. For example, prevailing winds, water flows and differences in altitude, changes in seasonal temperatures and other allergens in the environment may affect the impact of pollution on sickness (Anderson, 2020). However, sub-province fixed effects account for the permanent differences in the geography and weather conditions. In addition, our results are not sensitive to including quarter and month fixed effects in our model in account for the seasonality of sickness. Therefore, considering the above-mentioned endogeneity issues our estimates provide a lower-bound effect of overall mining pollution on the communities nearby because mining activities also pollute the air with dust and particulate matter, and stream sediments by leaching into water supplies. Soil pollution, therefore, partially captures the effect of overall pollution from mining activities.

2.3. Data

2.3.1. Contamination data

We use geo-referenced soil pollution data from mining sites in Mongolia, accessed from the Geo-Database on Ecological Health (GDEH), the Ministry of Environment and Green Development. A total of 1,315 soil samples from 262 mining sites in 95 sub-provinces across 17 provinces are recorded in the GDEH for 2002-2019.[4] However, our final sample consists of 39 mining sites that are within six kilometres of residential areas in 25 sub-provinces across 15 provinces as most samples were recorded in 2011-2012. Soil samples from gold, spar, wolfram, coal and limestone mining sites were collected and analyzed by the Geo-ecological Institute, the Central Geological Laboratory, and the Laboratory of National Agency for Meteorology and Environmental Monitoring. The atom absorption spectrophotometer method is used to determine heavy metals in the soil samples (GDEH, 2012).[5]

The database records mercury, arsenic, lead, zinc, cadmium, copper and nickel contamination in soil samples. To assess the heavy metal pollution severity, we use the following three values: (i) precaution value, (ii) trigger value, and (iii) action value set by the MASM (2019) (see Table 1). A value above the precaution value indicates that soil is polluted with heavy metals. A value exceeding the trigger value means that harm is caused to living organisms and areas of water. A value reaching the action value requires immediate action to neutralize the soil, stop current land uses, and relocate the affected population. Our analysis focuses on mercury and arsenic that exceeds the action and precaution values, respectively.

We use each soil sample point's longitude and latitude, along with a household's residential area coordinates, to calculate the distance from a household residential area to the nearest mine. We calculated the great-circle distance from the interior centroid of the location (i.e., residential

[4]There were 3,222 mining and exploration licenses issued to 2,063 mining companies in Mongolia between 1995 and 2019. The total area covered by mining licenses comprises 4.75 percent of the country's territory (Extractive Industries Transparency International, Mongolia, EITIM, 2020). Although the 262 mining sites represent 13 percent of mining companies, we limit the mines examined in the study to those located within six kilometres of a residential area.

[5]Atomic absorption spectrometry (AAS) detects heavy metals in solid samples through the application of characteristic wavelengths of electromagnetic radiation from a light source. Individual metals absorb wavelengths differently, and these absorbances are measured against the standards set to analyze the level of heavy metals (Thermo Fisher Scientific, 2021).

area) to the closest interior centroid of a soil sample site using the Haversine formula employed in Gradstein and Klemp (2020). A mining site has several sampling points for heavy metals. We calculated the distance from a household residential area to each of the sampling points. We used the shortest distance from a residential area to a sampling point as the primary variable of interest, thereby capturing a person's exposure to mining pollution.

The extent of environmental exposure to soil heavy metal contamination is recorded in Table 1. All individuals are exposed to mercury pollution level which significantly affects living organisms (columns 2 and 4). Furthermore, around 40 percent of individuals are exposed to mercury pollution level that exceeds the action value requiring the cleansing of soil and relocating of the inhabitants (column 6). While over 30 percent of the individuals are exposed to arsenic soil pollution, only around two percent of people are exposed to lead and cadmium pollution (column 2 of Table 1). The other heavy metals such as zinc, copper and nickel do not pollute the soil as they are within the permissible level.

[Insert Table 1]

2.3.2. Individual morbidity data

We use data from the most recent five rounds of the HSES, which is a nationally representative cross-sectional survey conducted by the National Statistics Office of Mongolia in every two years. The survey uses a stratified two-stage sample design based on population figures obtained from local governments' administrative records. The first stage stratifies the capital city, Ulaanbaatar, and the 21 provinces. The second stage divides the 21 provinces into two substrata: urban, comprising the provincial capitals, and rural, consisting of small towns and the countryside (National Statistics Office, 2018).[6] The five rounds of the HSES - 2008, 2010, 2014, 2016 and 2018 - included 38,425 individuals living in sub-provinces where mining occurred and soil samples had been collected. The 2012 round is excluded from the analysis as the geographic coordinates are missing from the data. The final sample, restricted to those living within six kilometres of any mine, comprises 6,713 individuals distributed as: 745 in 2008; 796 in 2010; 1,821 in 2014; 1,767 in 2016 and 1,584 in 2018.

[6]The HSES questionnaires and the primary datasets are publicly available from the NSO Census and Survey data catalog http://web.nso.mn/nada.

The types of health problems individuals experienced in the month prior to the survey are recorded. The illnesses reported by individuals fall into the following categories of body systems: (i) respiratory, (ii) digestive, (iii) cardiovascular and (iv) other illnesses including damage or intoxication by external impact. The HSES also provides information about household expenditure on medication, transportation, hospitalization, and other medical services in the prior 12 months. The summary statistics for the outcome variables are reported in Table 2. On average, eight percent of the sample said that they had been sick in the month before the survey. In addition to examining sickness from all causes, we look at illnesses of specific body systems. The average medical expenses per person is thousand MNT18 after adjusting for inflation.

[Insert Table 2]

The summary statistics for the explanatory and the control variables are shown in Table 3. The primary variable of interest that captures soil pollution exposure is the distance from an individual's residential area to the nearest mine. The average distance to the nearest mining site is three kilometres and it is similar for the different types of heavy metals. The individual-specific control variables included in the analyses are monthly consumption per capita, age, gender, education, number of household members, and household urban status.

[Insert Table 3]

3. Results

3.1. Main results

We investigate whether individuals living nearby mines, who, therefore, might be environmentally exposed to soil heavy metal pollution, have an increased likelihood of feeling unwell. The empirical analysis is carried out in two stages to test the four competing hypotheses to answer the research question. First, we estimate a simple binary outcome model with the primary variable of interest and sub-province fixed-effects to examine whether the distance to the nearest mine affects sickness. Next, our preferred estimation adds individual-specific control variables and survey-year fixed effects in the model.[7] The standard errors are clustered at the household level and the marginal effects at the means of covariates are estimated. Our investigation uses the conventional five percent significance level for testing the competing hypotheses.

We expect the variable of interest, the distance from an individual's residential area to the nearest mine, to have no impact on a person's probability of feeling sick. We estimate Equation 1 and report the results in Table 4 using the linear probability (LPM), probit and logit models. The baseline results from the LPM and the preferred model with control variables are presented in columns 1 and 2, respectively. The coefficient estimate of the distance to the nearest mine is negative and significant in column 1, indicating that the distance to the nearest mine significantly affects a person's probability of being sick: the further away from the mine, the less the chance of becoming ill. The model's results with individual-specific control variables and survey-year fixed effects are consistent with the baseline results. The magnitude of the coefficient for the distance is slightly higher, indicating that the addition of control variables were negatively correlated with the exposure variable.

[Insert Table 4]

The marginal effects from the probit model, shown in columns 3 and 4 of Table 4, are in line with the LPM results. Similarly, the marginal effects from the logit models in columns 5 and 6 are consistent with the LPM and probit models. The coefficient estimate for distance in column 6 indicates that a one per cent increase in the distance from a mining site would reduce a person's likelihood of feeling sick by 1.4 percentage points. In addition, survey year-fixed effects are included in the model to account for the duration of exposure to mining pollution. The estimates of the year-fixed effects are positively correlated with sickness but not at a significant level in 2012 and 2014. The correlation is negative in 2016 in the logit model but not significant. However, the year 2018 has a significant impact on sickness, showing the effect of duration of exposure to toxic materials. The overall goodness of fit for all models with the control variables are four percent for the LPM and seven percent for the probit and logit models and our main variable of interest is

[7] We use linear probability, probit and logit models to check the robustness of our result with different estimation techniques.

significant at the one percent level. The results from both probit and logit models indicate that soil pollution from mining has a significant adverse effect on the health of the inhabitants living near mines. We estimate a logit model in our subsequent analyses because of the model's efficiency to predict probability better than the other two models (Wooldridge, 2015).

The control variables in our models are all in line with our expectations and meaningful. Income is an important determinant of sickness but it is usually erroneous and we include consumption per capita as it is a more reliable indicator of welfare in the household data (Deaton, 1997). However, sickness can also affect medical expenses and other types of consumption, making the latter endogenous in the model. Therefore, we use the share of working-age members in the household as an instrument to predict consumption. The results indicate that individuals with higher consumption are less likely to feel sick.

Next, we examine whether distance to the highest level of heavy metal contamination (within six kilometres) significantly affects sickness. We use the distance to the mining site with the highest heavy metal contamination level instead of the shortest distance to a mine. Since the seven heavy metals coexist at most locations, the distance to the highest contamination level of any heavy metal also captures the effect of other coexisting heavy metals. Therefore, we exploit the distance to each heavy metal's highest contamination level as the variable of interest in our model by estimating Equation 2. Due to high level of correlation of the variables, all of them cannot be calculated in the same model (see Table A.1 for correlation matrix). The results from the seven separate models, each of which only includes the distance to a mine with the highest level of a particular heavy metal, are presented in Table 5. The coefficients in Table 5 are similar, indicating that our distance variables capture the effects of all heavy metals together. The results in Table 5 validate our approach to use distance to mines to capture the exposure to pollution.

[Insert Table 5]

The amount of heavy metal released into the soil can negatively impact the health of people living nearby to mines. We explore whether the level of pollution affects sickness. The assessment of heavy metal pollution is based on three values. In more than 30 percent of the sample population is exposed to arsenic pollution, whereas almost all individuals in our sample live in areas polluted with mercury (see Table 1). More concerning is that close to 40 percent of individuals reside in areas with severe mercury pollution levels, requiring cleansing of the soil and relocation of those exposed individuals. We examine these pollutants' impacts in more detail because of the particular concerns about arsenic and mercury.

The results of the investigation of arsenic and mercury contamination levels are provided in Table 6. The results are derived by interacting the general distance variable with the natural logarithm of arsenic and mercury pollution levels and estimating Equation 3. The results confirm our main finding and the primary variable of interest (distance to the nearest mine) is not affected by the inclusion of arsenic and mercury pollution levels. Overall, the results show that the distance to the nearest mine captures the effects of exposure. In contrast, the inclusion of the level of pollution does not affect the main findings.

[Insert Table 6]

Up to this point of our analysis, we assumed a linear relationship between the shortest distance and the likelihood of feeling sick in our main specification. However, the relationship between pollution level and the probability of feeling unwell can be nonlinear. Using the MASM (2019) threshold values and following Currie et al. (2009a), we include dummy variables for heavy metals that exceed the action value as that is the level most harmful to human and estimate Equation 4. The dummy variable takes the value of one for individuals exposed to mercury level that is above the action value, and zero otherwise. The effect of mercury is insignificant and the estimates in column 2 in Table 7 shows that distance significantly affects the probability of getting sick when mercury pollution is above the action value. The results show that the inclusion of level of pollution in a nonlinear form does not alter our main findings, indicating that the distance has a significant impact on sickness. Moreover, the magnitude of the effect is higher in column 2 of Table 7 than our main findings.

[Insert Table 7]

The results from Tables 4 - 7 confirm that the distance from a mining site to an individual's residential area significantly affects the probability of feeling sick. The estimates provide evidence that individuals experience adverse health outcomes from being environmentally exposed to soil pollution from mining activities. The results are in line with Hill (2018) and Marcus (2021) who examined effects of shale gas and petroleum leakage on infant health, respectively. Moreover, the findings support Aragón and Rud (2015) and Von der Goltz and Barnwal (2019) who also claim that mining activities have significant negative effects on the nearby communities. Due to the nature of our individual-level survey that only records self-reported illnesses rather than clinical records, our results fall short of explaining long-term chronic illnesses and cancer. However, it appears that the community is susceptible to feeling sick and bear the burden of the negative externalities of pollution from mining.

3.2. *Medical expenses*

We now move on to the additional analyses of medical expenses and examine whether moving away from a mining site also reduces medical expenses for individuals exposed to mining pollution. We estimate an OLS regression on the annual medical expenses and report the results in columns 1 and 2 of Table 8. The results in the baseline and the preferred models are similar. As reported in column 2, the annual medical expenses decline by 0.22 per cent as a person increases her distance from a mine by one percent. The medical expenses analyses further support our argument that pollution poses a negative externality to inhabitants affected by mining activities.

[Insert Table 8]

3.3. *The effect of mining pollution on younger children*

We next investigate whether soil pollution from mining affects different age groups disproportionately. Children and older people are more vulnerable and susceptible to feeling sick because of their poorer immune systems (Landrigan et al., 2018). Children below the age of 14 go through significant development changes that can have lasting effects on their well-being throughout their adulthood. Also, children are more vulnerable because their body size is smaller than adults, and their exposure to pollution may have more severe effects (Currie and Schmieder, 2009). Older people may have been exposed to soil pollution for a longer time. The working-age population between the ages of 15 and 64 runs the risk of occupational exposure to heavy metal pollution.

They range from miners to smelters, gold refiners, and people working in the auxiliary sectors such as trade, services, and transportation. Therefore, it is relevant to examine the effect of soil pollution on each age group separately.

A sub-sample analysis is undertaken separately for each population segment. We estimate Equation 1 and report the results for each age group in Table 9. Although the findings are consistent across all age groups, the impact is most pronounced for younger children up to 14 years old. The coefficient estimate for age group 0-14 year-old (columns 1 and 2) is slightly higher than the main finding in Table 4. As previously mentioned, children are more vulnerable, and their probability of getting sick reduces by 1.9 percentage points if their distance to a mining site increases by one per cent (column 2). Compared to younger children, soil pollution affects the working-age population to a lesser extent, which is not statistically significant (columns 3 and 4). Despite the small sample size for older people above the age of 65, the effect of soil heavy metal pollution is highest for older people but, again, not statistically significant (columns 5 and 6).[8] Therefore, mining pollution exerts a significant negative externality on younger children.[9]

[Insert Table 9]

3.4. The response of different body systems to mining pollution

Our third hypothesis tests whether the exposure to soil heavy metal pollution have a significant effect on various body systems. Testing this hypothesis helps us understand if body systems react differently to mining pollution. We estimate our primary model (Equation 1) on the illnesses of body systems. The results are presented in Table 10. The effect on the respiratory illnesses are reported in column 1, and indicates that the distance to the nearest mine adversely affects the respiratory system. The results for the digestive, cardiovascular systems and other illnesses are in columns 2-4 of Table 10. These results are not surprising as digestive and cardiovascular system illnesses are not easily diagnosed and may take longer time for a patient to discover.

[8] This finding is expected as older people might have more prolonged exposure to soil pollution and a greater build-up of heavy metals in their bodies if they have lived in the area for an extended period of time. Older people may also have underlying health conditions that are exacerbated by soil pollution from mining. Long-term exposure might result in the development of chronic illnesses in this population cohort.

[9] It is important to note that the effect of pollution on adults are confounded by other unobservable factors such as their immune system which is less susceptible to the impacts of soil pollution.

[Insert Table 10]

Compared to other body systems, respiratory illnesses are easily diagnosed as symptoms such as coughing and shortness of breath can make a person immediately uncomfortable without having to see a doctor. Besides common cold, such respiratory symptoms can also develop from exposure to heavy metal pollution. For example, acute exposure to mercury vapor leads to cough, dizziness and shortness of breath (Solis et al., 2000). Results from field surveys from artisanal and small-scale mining areas in Mongolia indicate that there is higher risks of suffering from asthma and tuberculosis among adults and increased prevalence of respiratory illnesses among children (Hugjliin Ezed NGO and Swiss Agency for Development and Cooperation, (SDC), 2010; Human Rights Commission and SDC, 2012). Although there is no *a-priori* expectation that mining activities will affect all body systems, the findings imply that living closer to a mining area significantly aggravates respiratory symptoms.

We control the seasonality of sickness by including quarter and month fixed-effects in our model because the common cold is the most prevalent sickness during the cold winter months from December to February in Mongolia (see Tables A.2 and A.3). The quarter and month fixed effects do not alter the main results, suggesting that pollution, as captured by the distance to the nearest mine, explains the likelihood of having respiratory illnesses in our model. The analysis on different body systems reveals that soil pollution is a significant contributor for feeling unwell. The respiratory tracts seem to be most affected by mining pollution, even after controlling for seasonal nature of illnesses. Such a finding is expected as mining activities also produce substantial amount of dust in the air (Li et al., 2014). The prevalence of respiratory illnesses among the inhabitants living close to mines suggest that people who are environmentally exposed to pollution have higher risks of developing more severe health conditions such as lung cancer and other chronic illnesses.

3.5. The effect of mine scale on morbidity

The next part of the analysis is concerned with the scale of mining practices. For example, medium and large scale mines rely on heavy machinery and advanced technologies to extract minerals. They are also more likely to enforce better safety standards for their workers and adhere to environmental regulations as they are licensed entities to operate mines. We refer to this medium to large-scale mines as official license holders in our study. The extent they pollute the environment can be large because of the scale of their mining practice. On the other hand, small-scale miners are mostly unlicensed individuals and individuals partnered under one mining license to extract minerals from the same mining site (Human Rights Commission and SDC, 2012). Small-scale miners rely on various tools to mine and process minerals, and may pay less attention to personal safety standards or environmental impacts due to their inadequate financial and technical capabilities. As mentioned earlier, using mercury to separate gold from gold amalgam is a common practice in developing countries.

Therefore, our next hypothesis tests whether licensed entities and small-scale mines affect sickness differently. We undertake sub-sample analyses for individuals affected by each mine scale separately. The results are reported in Table 11. The baseline results without control variables for both official license holders, and small-scale mines are shown in columns 1 and 3, respectively. Both estimates are qualitatively similar. However, the estimated impacts of small-scale mines appear to be larger than that of the official license holders in our preferred model with the control variables in column 4.

[Insert Table 11]

A person will be 1.7 percentage points less likely to feel sick if the distance from small-scale mine is increased by one per cent. Compared to our main findings, the magnitude of the estimate is slightly higher for small-scale mines (see Table 4). Together these results provide important insights into the varying effects of mine scale on the health of communities nearby, especially the severity of soil pollution from small-scale mines in developing countries.

3.6. The impact of different types of minerals mined on sickness

The final investigation looks at whether the types of minerals mined matter for feeling unwell. We test the hypothesis that all minerals have no discernible effect on sickness by undertaking sub-sample analysis for each mineral category. Gold, limestone, and coal are the primary minerals that people are exposed to within six kilometres of their residential area. The results from the analysis of different minerals are reported in Table 12. The distance from the nearest gold mine adversely affects sickness (column 2), limestone and other minerals such as coal, spar and wolfram have similar effects although they are not statistically significant. Such an outcome can be due to the lower number of observations for the other types if minerals. Taken together, these results suggest that gold mines have a significant discernible effect on the likelihood of feeling sick.

[Insert Table 11]

The results in this section indicate that pollution from mining activities captured by soil pollution adversely affects the health of nearby communities significantly. The four hypotheses we tested all reveal that there is a causal link between mining pollution and sickness. Younger children living within six kilometres of a polluted mine site are more prone to sickness. Besides, the respiratory system is particularly sensitive to mining pollution. Small-scale mines have a bigger effect on health than medium and larger mines, and gold mines also increase the chances of feeling sick. However, our lower-bound effects of mining pollution on sickness is considerable and calls for better environmental quality standards in resource-rich developing countries. Our results are robust to the application of different estimation techniques that are discussed in the next section.

3.7. Robustness checks

Additional robustness checks are undertaken to confirm that the type of models used and biases related to endogeneity and omitted variables do not drive our results. First, we repeat Tables 4-7 and Tables 9-12, using the LPM. Although the LPM does not predict a probability within the boundary of zero and one perfectly, it provides a useful comparison to the logit model. The results in Tables A.4 – A.10 are similar to the results from the logit model and support our findings. The same applies to our exercise with a probit model.

Next, we use the principal component analysis (PCA) for an additional robustness test. The PCA is a statistical technique for data reduction, which creates new uncorrelated variables-principal components with the highest variance from a large dataset (Jolliffe and Cadima, 2016). Using the PCA, we reduce the levels of seven types of heavy metals into three components, each component grouping specific heavy metals together. Table A.11 presents the results for a principal component containing copper, mercury, and nickel. The estimate of the distance to the nearest mine is negative and significant in columns 1 and 2, whereas the impact of the chosen principal component is insignificant, similar to the results in Tables 6 and 7.[10] In summary, the results from the LPM and PCA confirm that soil pollution affects sickness at a significant level and our main findings are robust.

4. Discussion

Our study provides an empirical evidence for the negative externality of mining industries that arise from mining, refining, processing minerals, and disposing of chemicals into the environment. There are three main findings in this paper. First, the distance to the nearest mine significantly affects a person's probability of feeling sick and the effect is considerably higher for children aged 0-14. This finding is concerning because early life exposure to neurotoxins such as lead and mercury can affect cognitive abilities, disrupt concentration and behavior, leading to lifetime earnings loss (Landrigan et al., 2018). These health impacts are irreversible and can have long-term and inter-generational effects on well-being and earnings which will have further implications for productivity loss.

Second, respiratory illnesses appear to be aggravated by exposure to soil pollution, as indicated by a sub-sample analysis of those in the population who experienced respiratory diseases in the past month. While we could only control for sickness seasonality through quarter and month fixed effects, other pollution sources such as dust and particulate matter in the air may affect the estimates. However, the soil pollution data and distance to the nearest mine are good indicators to measure the impact of mining activities on health.

A third finding is that small-scale mining activities affect the health of those living nearby mines, suggesting that environmental regulation monitoring, in general, is weak. While small-scale miners use mercury and cyanide extensively in their mining practices, medium to large-scale mining companies may release a significant amount of toxins and waste into the environment due to their

[10] We examined the other two components containing the rest of the heavy metals. However, the main variable of interest, the distance to the nearest mine is not significant. The reason is that these two components do not contain mercury which surpasses both trigger and action values.

operations scale. However, the impact of small-scale mining on sickness is more significant compared to fully licensed entities due to their low adherence to environmental regulations. Therefore, mining activities at smaller-scale pose significant threats to the health of communities nearby the mines.

On average, 533 people, who live within six kilometres of a mine, felt sick in the month prior to the survey in our sample. However, this number represents 76,000 people. If these individuals move one kilometre further away from a mine, they will be around one percentage point less likely to feel sick.[11] The one percentage point reduction translates into a 11.5 per cent overall reduction in the number of people feeling unwell. In other words, over 8,700 fewer people would not have felt sick if they were one kilometre away from a mining site.

The costs of feeling unwell are multiple. First, sickness deteriorates human physical and emotional well-being, which have wider impact on labor supply and productivity (Graff Zivin and Neidell, 2013; Hanna and Oliva, 2015). Second, it leads to school absences and lower performance in the short-term and loss in lifetime earnings in the long-term (Neidell, 2004; Rau et al., 2015). Third, pollution-related sicknesses and diseases incur intangible costs such as disruption of family stability when a family member becomes ill or dies because of pollution, and loss in years of life to the sick person (Landrigan et al., 2018). Moreover, both private (i.e., home care) and public costs (i.e., hospital, physician and medical costs) are incurred to treat illnesses. For example, the annual personal medical expenditures of \$63 thousand could have been saved if 8,700 people moved one kilometer away from a polluted mining site. Similarly, annual fiscal medical expenditures of \$572 thousand are incurred for those who felt sick in the past month. Therefore, being environmentally exposed to mining pollution costs over \$635 thousand annually in addition to pains and sufferings of the sick people.

Our estimates are likely to be lower-bound estimates of the impact of overall pollution from mining on health as we cannot rule out the effect of other types of pollution such as air and water pollution. Furthermore, the effect of long-term exposure to heavy metal pollutants may be undetected in the absence of biometric information and clinical examinations of the population.

[11] This result arises from an examination of the change in the estimate when the distance is increased by one kilometre.

Nevertheless, these estimates are large in magnitude and suggest that living at a greater distance from a mine would substantially benefit the communities exposed to mining activities. Therefore, substantial economic and social benefits can be realized through the reduction of pollution from mining activities with appropriate policies and institutions in place in resource-rich developing countries.

5. Conclusion

We examined the impact of soil pollution on individuals' likelihood of feeling sick when environmentally exposed to pollution from mining activities. Combining a geo-coded soil pollution information with five rounds of household socioeconomic survey data from Mongolia, we exploit the distance from an individual's residential area to the nearest mine as the pollution exposure in our model. We find that the exposure to mining pollution significantly increases the probability of feeling sick within six kilometres of a mining site. Younger children appear to suffer the most from mining pollution. Furthermore, respiratory illnesses are exacerbated by pollution, which can increase the risks of developing chronic diseases such asthma among children. In addition, small-scale gold mining activities adversely affect human health.

Our study provides a new empirical evidence on the effect of pollution from mining activities in a resource-rich developing country. Unlike previous studies, we investigate different age groups, scales of mines and mineral types that have not been examined in a holistic approach. The results are lower-bound effects of overall pollution from mining activities as we are capturing overall pollution with soil pollution data that is highly correlated with air and water pollution. Compared to air pollution, soil pollution is not directly visible to the naked human eye, making it difficult for people exposed to detect pollution and mitigate the risks of adverse health effects. Therefore, it might take years before severe irreversible health problems occur in the community exposed tp pollution. We contribute to the existing literature by investigating the impact of soil pollution from recent mining activities and examining the early symptoms of potential chronic illnesses that could be induced by exposure to heavy metal pollution in soil. The findings call for stricter environmental regulations of the mining industry and tighter monitoring of soil heavy metal pollution in resource-rich developing countries.

6. Tables

Table 1: **Summary statistics of level of contamination**

Heavy metal	(precaution) (1)	(>precaution) (2)	(trigger) (3)	(>trigger) (4)	(action) (5)	(>action) (6)
Mercury (Hg)	2.0	0.99	10.0	0.93	20.0	0.39
		(0.11)		(0.25)		(0.49)
Arsenic (As)	20.0	0.31	50.0	0.04	100.0	0.00
		(0.46)		(0.20)		(0.07)
Lead (Pb)	100.0	0.02	500.0	0.00	1,200.0	0.00
		(0.13)		(0.00)		(0.00)
Zinc (Zn)	300.0	0.00	500.0	0.00	1,000.0	0.00
		(0.00)		(0.00)		(0.00)
Cadmium (Cd)	3.0	0.02	10.0	0.00	20.0	0.00
		(0.13)		(0.00)		(0.00)
Copper (Cu)	100.0	0.00	500.0	0.00	1,000.0	0.00
		(0.00)		(0.00)		(0.00)
Nickel (Ni)	150.0	0.00	600.0	0.00	1,000.0	0.00
		(0.00)		(0.00)		(0.00)
N		6,713		6,713		6,713

Note: 1. A value that is above the precaution value indicates pollution of heavy metal in soil. Precaution value and permissible value are the same.
2. A value exceeding the trigger value harms living organisms and water surfaces. A licensed entity in a manufacturing and mining zones shall monitor the value.
3. A value exceeding the action value requires immediate action to clean the soil pollution. For example, action such as neutralizing the pollution, removing the polluted soil, stopping land use and relocating residents are required.
4. The sample consists of 39 mining sites that is within six kilometres of a residential area.
5. Standard deviations are recorded in the parentheses.

Table 2: **Summary statistics of outcome variables**

Variable name	Mean	SD
Sick in the past month	0.08	0.27
Respiratory system illness	0.02	0.15
Digestive system illness	0.01	0.09
Cardiovascular system illness	0.02	0.13
External impact & other illness	0.03	0.17
Annual medical expenses	18.04	132.72
Number of observations	6,713	

Note: 1. The mean of annual medical expenses are reported in thousand Tugrik (MNT): The exchange rate for the end of survey period (December) ranged from US$1≈MNT1,229 in 2008 to US$1≈MNT2,644 in 2018. All values are on per capita monthly basis and adjusted for inflation.

Table 3: **Summary statistics of independent variables**

Variable name	Mean	SD
Distance to the nearest mine (km)	3.10	1.74
Distance to the nearest mercury (km)	3.09	1.74
Distance to the nearest arsenic (km)	3.07	1.73
Distance to the nearest lead (km)	3.40	1.66
Distance to the nearest zinc (km)	3.55	1.59
Distance to the nearest cadmium (km)	3.24	1.63
Distance to the nearest copper (km)	3.57	1.61
Distance to the nearest nickel (km)	3.46	1.66
Consumption per capita	147.8	105.1
Individual's age (years)	29.3	19.8
Individual is female	0.51	0.50
Individual's education (years)	8.14	5.63
Number of HH members	4.26	1.61
Lives in rural area	0.47	0.50
Number of observations	6,713	

Note: 1. Consumption (monthly) per capita is reported in thousand Tugrik (MNT): The exchange rate for the end of survey period (December) ranged from US$1 ≈ MNT1,229 in 2008 to US$1≈ MNT2,644 in 2018. All values are adjusted for inflation. 3. The distance to the nearest mine with heavy metal is not the same as the distance to the nearest mine. Heavy metal contamination within six kilometers is not required to be above the precaution value. The number of observations varies for each heavy metal.

Table 4: **The effect of mining pollution on sickness**

Variable name	(1)	(2)	(3)	(4)	(5)	(6)
ln(distance to the nearest mine)	-0.015**	-0.020***	-0.011**	-0.016***	-0.011**	-0.014***
	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
Individual is female		0.009		0.009		0.009
		(0.007)		(0.006)		(0.006)
Individual's age (years)		0.008***		0.006***		0.005***
		(0.002)		(0.001)		(0.001)
Individual's education (years)		-0.012***		-0.009***		-0.008***
		(0.002)		(0.002)		(0.001)
Number of HH members		-0.295***		-0.213***		-0.182***
		(0.079)		(0.058)		(0.053)
Consumption per capita		-0.311***		-0.223***		-0.190***
		(0.085)		(0.063)		(0.057)
Lives in rural area		-0.042		-0.040		-0.037
		(0.034)		(0.047)		(0.048)
2010		0.036*		0.028*		0.022
		(0.020)		(0.015)		(0.014)
2014		0.020		0.009		0.005
		(0.019)		(0.016)		(0.015)
2016		0.012		0.000		-0.005
		(0.020)		(0.016)		(0.015)
2018		0.063***		0.044**		0.036**
		(0.024)		(0.019)		(0.017)
Model	LPM	LPM	Probit	Probit	Logit	Logit
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2/Psedu R^2	0.01	0.04	0.02	0.07	0.02	0.07
N	6,713	6,713	6,713	6,713	6,713	6,713

Note: 1. Standard errors, clustered at the household level, are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the probit and logit models.
* p <0.10, ** p <0.05, *** p <0.01.

Table 5: **The effect of mining pollution on sickness: using distance from the nearest mine with particular types of heavy metal contamination**

Variable name	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(distance to Mercury)	-0.014***						
	(0.005)						
ln(distance to Arsenic)		-0.014***					
		(0.005)					
ln(distance to Lead)			-0.014**				
			(0.006)				
ln(distance to Zinc)				-0.014**			
				(0.006)			
ln(distance to Cadmium)					-0.017***		
					(0.006)		
ln(distance to Copper)						-0.017***	
						(0.006)	
ln(distance to Nickel)							-0.015***
							(0.006)
Individual is female	0.008	0.010*	0.014**	0.014**	0.016**	0.014**	0.014**
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)
Individual's age (years)	0.005***	0.005***	0.005***	0.006***	0.006***	0.006***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Individual's education (years)	-0.008***	-0.008***	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of HH members	-0.181***	-0.175***	-0.204***	-0.211***	-0.215***	-0.208***	-0.204***
	(0.053)	(0.055)	(0.060)	(0.061)	(0.061)	(0.061)	(0.061)
Consumption per capita	-0.190***	-0.182***	-0.213***	-0.221***	-0.226***	-0.218***	-0.213***
	(0.057)	(0.060)	(0.065)	(0.066)	(0.066)	(0.066)	(0.066)
Lives in rural area	0.011	0.017	0.039	0.079	0.075	0.071	-0.030
	(0.048)	(0.051)	(0.048)	(0.052)	(0.052)	(0.051)	(0.054)
2010	0.023*	0.023	0.018	0.014	0.020	0.014	0.013
	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
2014	0.006	0.005	0.001	-0.003	0.003	-0.004	-0.006
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
2016	-0.003	-0.009	-0.018	-0.025	-0.017	-0.025	-0.026
	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
2018	0.039**	0.037**	0.037*	0.034*	0.040**	0.033*	0.032
	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Psedu R^2	0.07	0.07	0.07	0.07	0.07	0.08	0.08
N	6,703	6,320	5,611	5,424	5,565	5,446	5,358

Note: 1. The impact of contamination is estimated by distance to the highest level of each heavy metal within 6 kilometres. Since most contaminants coexist at a mining site, this analysis looks at the impact of highest contamination level of each heavy metal on sickness within 6 kilometres. We assume that the effect of other contaminants are captured by specific contaminant examined by the distance. Since some heavy metals are not reported in some sites, the number of observations differ for each heavy metal.
2. Standard errors, clustered at the household level, are recorded in the parentheses.
3. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table 6: **The effect of mining pollution on sickness: including the level of pollution in the model**

Variable name	(1)	(2)
ln(distance to the nearest mine)	-0.019**	-0.020***
	(0.008)	(0.007)
ln(Arsenic pollution level)	0.017	0.018
	(0.016)	(0.015)
ln(Mercury pollution level)	-0.045	-0.036
	(0.029)	(0.028)
Individual is female		0.009
		(0.006)
Individual's age (years)		0.005***
		(0.001)
Individual's education (years)		-0.008***
		(0.001)
Number of HH members		-0.183***
		(0.053)
Consumption per capita		-0.191***
		(0.057)
Lives in rural area		0.010
		(0.062)
2010		0.022
		(0.014)
2014		0.006
		(0.015)
2016		-0.004
		(0.015)
2018		0.038**
		(0.017)
Sub-province fixed effects	Yes	Yes
Psedu R^2	0.02	0.07
N	6,713	6,713

Note: 1. Standard errors, clustered at the household level, are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table 7: **The effect of mining pollution on sickness: controlling for the nonlinear effect of pollution level**

Variable name	(1)	(2)
ln(distance to the nearest mine)	-0.016**	-0.017**
	(0.008)	(0.007)
Hg above action value	-0.016	-0.009
	(0.017)	(0.016)
Individual is female		0.009
		(0.006)
Individual's age (years)		0.005***
		(0.001)
Individual's education (years)		-0.008***
		(0.001)
Number of HH members		-0.182***
		(0.053)
Consumption per capita		-0.190***
		(0.057)
Lives in rural area		-0.024
		(0.053)
2010		0.022
		(0.014)
2014		0.005
		(0.015)
2016		-0.004
		(0.015)
2018		0.037**
		(0.017)
Sub-province fixed effects	Yes	Yes
Psedu R^2	0.02	0.07
N	6,713	6,713

Note: 1. In order to account for the nonlinear effects of pollution level, we include a dummy variable for individuals exposed to mercury pollution level that is above the action value. 39 percent of the individuals are exposed to mercury pollution that is considered very harmful to the population exposed where as over 30 percent of individuals are exposed to arsenic pollution. Therefore, we include a dummy taking the value of one for those exposed to high mercury pollution level and zero otherwise.
2. Distance to each heavy metal is 6 kilometres.
3. Standard errors, clustered at the household level, are recorded in the parentheses.
4. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table 8: **The effect of mining pollution on medical expenses**

	ln(medical expenses)	
Variable names	(1)	(2)
ln(distance to the nearest mine)	-0.175**	-0.219***
	(0.082)	(0.079)
Individual is female		0.191**
		(0.083)
Individual's age (years)		0.077***
		(0.020)
Individual's education (years)		-0.062**
		(0.026)
Number of HH members		-1.976**
		(0.973)
Consumption per capita		-2.093**
		(1.051)
Lives in rural area		0.786
		(0.563)
2010		0.014
		(0.199)
2014		0.269
		(0.247)
2016		0.332
		(0.250)
2018		0.532*
		(0.291)
Sub-province fixed effects	Yes	Yes
R^2	0.02	0.07
N	6,713	6,713

Note: 1. Standard errors, clustered at the household level, are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

TABLE 9: **The effect of mining pollution on medical expenses and sickness for different age groups**

	Age: 0-14 years		Age: 15-65 years		Age: 65+ years	
Variable names	(1)	(2)	(3)	(4)	(5)	(6)
ln(distance to the nearest mine)	-0.023***	-0.019***	-0.006	-0.008*	-0.078	-0.074
	(0.009)	(0.007)	(0.005)	(0.005)	(0.051)	(0.053)
Individual is female		0.002		0.012*		-0.111*
		(0.010)		(0.006)		(0.058)
Individual's age (years)		-0.006*		0.006***		0.006
		(0.003)		(0.001)		(0.013)
Individual's education (years)		-0.000		-0.006***		-0.021
		(0.006)		(0.002)		(0.018)
Number of HH members		0.047		-0.221***		-0.329
		(0.163)		(0.064)		(0.558)
Consumption per capita		0.062		-0.233***		-0.353
		(0.177)		(0.069)		(0.621)
Lives in rural area		0.030		-0.010		0.357
		(0.061)		(0.048)		(0.388)
2010		0.086*		0.005		0.100
		(0.045)		(0.014)		(0.115)
2014		0.072		-0.010		-0.109
		(0.049)		(0.016)		(0.157)
2016		0.063		-0.026		-0.079
		(0.051)		(0.017)		(0.153)
2018		0.069		0.024		0.017
		(0.057)		(0.019)		(0.182)
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Psedu R^2	0.04	0.08	0.02	0.08	0.04	0.09
N	1,781	1,781	4,488	4,488	256	256

Note: 1. Standard errors, clustered at the household level, are recorded in the parentheses.
2. Sub-provinces, where sickness is not reported, are dropped from the logit model as there is no variation in the outcome variable in the sub-sample analysis.
3. Individuals living at a proximity of up to 6 km are analyzed in this table.
4. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table 10: **The effect of mining pollution on body system illnesses**

Variable name	Respiratory (1)	Digestive (2)	Cardiovascular (3)	Other illnesses (4)
ln(distance to the nearest mine)	-0.0063***	-0.0006	-0.0004	-0.0039
	(0.0021)	(0.0013)	(0.0010)	(0.0026)
Individual is female	0.0030	0.0003	0.0027	0.0020
	(0.0026)	(0.0015)	(0.0014)	(0.0033)
Individual's age (years)	0.0003	0.0002*	0.0004***	0.0010***
	(0.0003)	(0.0001)	(0.0001)	(0.0002)
Individual's education (years)	-0.0019***	0.0001	-0.0000*	-0.0007*
	(0.0004)	(0.0002)	(0.0002)	(0.0004)
Number of HH members	-0.0047***	-0.0012**	-0.0009	0.0000
	(0.0013)	(0.0006)	(0.0005)	(0.0011)
Consumption per capita	-0.0121**	-0.0001	-0.0006**	-0.0125**
	(0.0060)	(0.0020)	(0.0019)	(0.0057)
Lives in rural area	0.0218	0.0126***	-0.0048	0.0031
	(0.0171)	(0.0045)	(0.0084)	(0.0178)
2010	0.0141*	-0.0011	-0.0059*	0.0146*
	(0.0082)	(0.0029)	(0.0031)	(0.0081)
2014	0.0082	-0.0088***	-0.0059	0.0082
	(0.0078)	(0.0028)	(0.0025)	(0.0067)
2016	0.0114	-0.0074***	-0.0086	-0.0025
	(0.0077)	(0.0027)	(0.0028)	(0.0075)
2018	0.0121	-0.0037	-0.0035**	0.0136**
	(0.0076)	(0.0027)	(0.0026)	(0.0069)
Sub-province fixed effects	Yes	Yes	Yes	Yes
Psedu R^2	0.09	0.11	0.18	0.06
N	6,282	5,039	6,179	6,587

Note: 1. Standard errors, clustered at the household level, are recorded in the parentheses.
2. Sub-provinces, where sickness of for a particular body system is not reported, are dropped from the logit model as there is no variation in the outcome variable in the sub-sample analysis.
3. Individuals living at a proximity of up to 6 km are analyzed in this table.
4. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table 11: **The effect of mining scale on sickness**

Variable name	Mining license holders (1)	(2)	Small-scale miners (3)	(4)
ln(distance to the nearest mine)	-0.006	-0.008	-0.014**	-0.017***
	(0.009)	(0.008)	(0.006)	(0.006)
Individual is female		0.001		0.022**
		(0.007)		(0.009)
Individual's age (years)		0.002***		0.003***
		(0.000)		(0.001)
Individual's education (years)		-0.004***		-0.005***
		(0.001)		(0.001)
Number of HH members		-0.005*		-0.011***
		(0.003)		(0.003)
Consumption per capita		-0.026**		-0.031*
		(0.010)		(0.016)
Lives in rural area		-0.005		-0.050
		(0.041)		(0.053)
2010		0.062***		-0.023
		(0.022)		(0.018)
2014		-0.008		-0.026
		(0.018)		(0.017)
2016		-0.018		-0.033*
		(0.019)		(0.017)
2018		0.004		0.004
		(0.019)		(0.017)
Sub-province fixed effects	Yes	Yes	Yes	Yes
Psedu R^2	0.02	0.09	0.02	0.07
N	3,338	3,338	3,343	3,343

Note: 1. Standard errors, clustered at the household level, are recorded in the parentheses.
2. Sub-provinces, where sickness is not reported, are dropped from the logit model as there is no variation in the outcome variable in the sub-sample analysis.
3. Individuals living at a proximity of up to 6 km are analyzed in this table.
4. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table 12: **The effect of mineral type on sickness**

	Gold		Limestone		Other	
Variable name	(1)	(2)	(3)	(4)	(5)	(6)
ln(distance to the nearest mine)	-0.013**	-0.015***	-0.011	-0.007	0.005	-0.017
	(0.006)	(0.005)	(0.061)	(0.020)	(0.042)	(0.020)
Individual is female		0.014*		0.002		0.010
		(0.007)		(0.004)		(0.011)
Individual's age (years)		0.003***		0.001***		0.002***
		(0.001)		(0.000)		(0.001)
Individual's education (years)		-0.005***		-0.001*		-0.003**
		(0.001)		(0.001)		(0.001)
Number of HH members		-0.009***		-0.002		-0.007*
		(0.003)		(0.002)		(0.004)
Consumption per capita		-0.025*		-0.022**		-0.020
		(0.013)		(0.009)		(0.019)
Lives in rural area		-0.094*		0.030		0.061
		(0.050)		(0.043)		(0.065)
2010		-0.016		0.017**		0.031
		(0.018)		(0.008)		(0.032)
2014		-0.012		-0.027***		-0.009
		(0.015)		(0.010)		(0.031)
2016		-0.028*		-0.028**		-0.004
		(0.016)		(0.014)		(0.029)
2018		0.024		-0.041***		-0.021
		(0.016)		(0.016)		(0.032)
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Psedu R^2	0.02	0.07	0.01	0.29	0.00	0.06
N	4,327	4,327	904	904	1,450	1,450

Note: 1. Standard errors, clustered at the household level, are recorded in the parentheses.
2. Sub-provinces, where sickness is not reported, are dropped from the logit model as there is no variation in the outcome variable in the sub-sample analysis.
3. Individuals living at a proximity of up to 6 km are analyzed in this table.
4. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

7. Additional tables

Table A.1: **Cross-correlation table**

Variables	Mercury	Arsenic	Lead	Zinc	Cadmium	Copper	Nickel
Mercury	1.000						
Arsenic	-0.042	1.000					
Lead	0.506	0.348	1.000				
Zinc	0.114	0.690	0.681	1.000			
Cadmium	0.497	0.271	0.977	0.656	1.000		
Copper	0.307	0.418	0.762	0.542	0.769	1.000	
Nickel	0.439	-0.042	0.329	0.114	0.365	0.519	1.000

Table A.2: **The effect of mining pollution on sickness**

Variable name	Sick (1)	Respiratory (2)	Digestive (3)	Cardiovascular (4)	Other illnesses (5)
ln(distance to the nearest mine)	-0.013***	-0.005**	-0.001	-0.000	-0.004
	(0.005)	(0.002)	(0.001)	(0.001)	(0.003)
Individual is female	0.012**	0.003	0.000	0.003*	0.002
	(0.006)	(0.003)	(0.002)	(0.002)	(0.003)
Individual's age (years)	0.003***	0.000	0.000**	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Individual's education (years)	-0.005***	-0.002***	0.000	-0.000	-0.001*
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Consumption per capita	-0.031***	-0.010	-0.001	-0.001	-0.014**
	(0.011)	(0.007)	(0.003)	(0.002)	(0.006)
Lives in rural area	-0.006	0.019	0.013**	-0.006	0.005
	(0.046)	(0.019)	(0.006)	(0.009)	(0.018)
Quarter 2	-0.019**	-0.013***	0.004	0.001	-0.004
	(0.009)	(0.004)	(0.003)	(0.002)	(0.005)
Quarter 3	-0.042***	-0.028***	0.003	-0.001	-0.004
	(0.010)	(0.005)	(0.003)	(0.003)	(0.005)
Quarter 4	-0.044***	-0.021***	0.003	-0.002	-0.014**
	(0.010)	(0.005)	(0.003)	(0.003)	(0.006)
2010	0.024*	0.015*	-0.001	-0.007*	0.015*
	(0.014)	(0.008)	(0.003)	(0.004)	(0.008)
2014	-0.017	0.006	-0.009***	-0.006*	0.007
	(0.012)	(0.007)	(0.003)	(0.003)	(0.007)
2016	-0.026**	0.009	-0.008***	-0.009***	-0.004
	(0.013)	(0.007)	(0.003)	(0.003)	(0.007)
2018	0.011	0.012*	-0.003	-0.003	0.014**
	(0.013)	(0.007)	(0.003)	(0.003)	(0.007)
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes
Psedu R^2	0.08	0.11	0.11	0.06	0.18
N	6,713	6,282	5,039	6,587	6,179

Note: 1. Standard errors clustered at the household level are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.3: **The effect of mining on sickness**

Variable name	Sick (1)	Respiratory (2)	Digestive (3)	Cardiovascular (4)	Other illnesses (5)
ln(distance to the nearest mine)	-0.013***	-0.005**	-0.001	-0.000	-0.004
	(0.005)	(0.002)	(0.001)	(0.001)	(0.003)
Individual is female	0.012**	0.003	0.000	0.003*	0.002
	(0.006)	(0.002)	(0.002)	(0.001)	(0.003)
Individual's age (years)	0.003***	0.000	0.000**	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Individual's education (years)	-0.005***	-0.002***	0.000	-0.000	-0.001*
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Consumption per capita	-0.031***	-0.009	-0.000	-0.001	-0.014**
	(0.010)	(0.006)	(0.002)	(0.002)	(0.006)
Lives in rural area	0.006	0.024	0.011**	-0.002	0.005
	(0.046)	(0.020)	(0.006)	(0.009)	(0.018)
February	-0.012	-0.003	-0.002	-0.006	-0.002
	(0.014)	(0.005)	(0.005)	(0.004)	(0.008)
March	-0.028*	-0.006	-0.005	-0.001	-0.014
	(0.016)	(0.006)	(0.006)	(0.004)	(0.010)
April	-0.036**	-0.016**	-0.003	-0.000	-0.008
	(0.016)	(0.007)	(0.005)	(0.004)	(0.010)
May	-0.042***	-0.015**	0.001	-0.002	-0.012
	(0.014)	(0.006)	(0.004)	(0.004)	(0.009)
June	-0.020	-0.014**	0.004	-0.000	-0.006
	(0.014)	(0.006)	(0.003)	(0.003)	(0.008)
July	-0.062***	-0.026***	0.000	-0.009**	-0.009
	(0.016)	(0.007)	(0.004)	(0.004)	(0.009)
August	-0.044***	-0.053***	0.003	-0.002	0.002
	(0.016)	(0.011)	(0.004)	(0.004)	(0.008)
September	-0.050***	-0.019***	0.000	0.001	-0.021**
	(0.015)	(0.007)	(0.004)	(0.004)	(0.010)
October	-0.052***	-0.035***	0.006	0.001	-0.022**
	(0.017)	(0.010)	(0.004)	(0.004)	(0.010)
November	-0.066***	-0.019***	-0.001	-0.007*	-0.024**
	(0.015)	(0.006)	(0.004)	(0.004)	(0.010)
December	-0.048***	-0.019***	0.001	-0.005	-0.011
	(0.016)	(0.006)	(0.004)	(0.004)	(0.008)
2010	0.025*	0.015**	-0.001	-0.007**	0.015*
	(0.014)	(0.007)	(0.003)	(0.004)	(0.008)
2014	-0.015	0.008	-0.008***	-0.006**	0.007
	(0.013)	(0.007)	(0.003)	(0.003)	(0.007)
2016	-0.026**	0.010	-0.008***	-0.009***	-0.005
	(0.013)	(0.007)	(0.003)	(0.003)	(0.007)
2018	0.012	0.013*	-0.004	-0.003	0.014*
	(0.013)	(0.007)	(0.003)	(0.003)	(0.007)
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes
Psedu R^2	0.08	0.12	0.12	0.07	0.20
N	6,713	6,282	5,039	6,587	6,179

Note: 1. Standard errors clustered at the household level are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.4: **The effect of mining pollution on sickness, by distance to heavy metal**

	Sick						
Variable name	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(distance to Mercury)	-0.021***						
	(0.007)						
ln(distance to Arsenic)		-0.020***					
		(0.007)					
ln(distance to Lead)			-0.019**				
			(0.008)				
ln(distance to Zinc)				-0.019**			
				(0.008)			
ln(distance to Cadmium)					-0.024***		
					(0.008)		
ln(distance to Copper)						-0.023***	
						(0.008)	
ln(distance to Nickel)							-0.020**
							(0.008)
Individual is female	0.014**	0.016**	0.021***	0.020***	0.022***	0.020***	0.020***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Individual's age	0.004***	0.004***	0.004***	0.004***	0.005***	0.004***	0.004***
(years)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Individual's education	-0.006***	-0.006***	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***
(years)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Consumption per	-0.042***	-0.039***	-0.046***	-0.048***	-0.049***	-0.047***	-0.046***
capita	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Lives in rural area	0.000	-0.021	0.002	0.040	0.036	-0.034	-0.005
	(0.031)	(0.022)	(0.039)	(0.048)	(0.048)	(0.021)	(0.024)
2010	0.029	0.028	0.019	0.013	0.020	0.013	0.013
	(0.019)	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
2014	-0.014	-0.013	-0.021	-0.028*	-0.021	-0.029*	-0.031*
	(0.015)	(0.015)	(0.016)	(0.017)	(0.016)	(0.017)	(0.017)
2016	-0.021	-0.025*	-0.038**	-0.047***	-0.038**	-0.047***	-0.046***
	(0.015)	(0.015)	(0.016)	(0.017)	(0.016)	(0.017)	(0.017)
2018	0.014	0.014	0.008	0.002	0.009	0.001	0.003
	(0.016)	(0.016)	(0.017)	(0.018)	(0.017)	(0.018)	(0.018)
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.04	0.04	0.04	0.04	0.05	0.05	0.05
N	6,772	6,528	5,680	5,493	5,634	5,547	5,459

Note: 1. The impact of contamination is estimated by distance to the highest level of each heavy metal within 6 kilometres. Since most contaminants coexist at a mining site, this analysis looks at the impact of highest contamination level of each heavy metal on sickness within 6 kilometres. We assume that the effect of other contaminants are captured by specific contaminant examined by the distance.
2. Standard errors clustered at the household level are recorded in the parentheses.
3. Marginal effects are calculated at the means of all other covariates from the linear probability model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.5: **The effect of mining pollution level on sickness**

	Sick	
Variable name	(1)	(2)
ln(distance to the nearest mine)	-0.024**	-0.029***
	(0.010)	(0.010)
ln(Arsenic pollution level)	0.021	0.027**
	(0.017)	(0.010)
ln(Mercury pollution level)	-0.050	-0.049
	(0.038)	(0.037)
Individual is female		0.014
		(0.009)
Individual's age (years)		0.004***
		(0.001)
Individual's education (years)		-0.007***
		(0.001)
Consumption per capita		-0.043***
		(0.012)
Lives in rural area		0.063
		(0.058)
2010		0.029
		(0.036)
2014		-0.012
		(0.018)
2016		-0.021
		(0.024)
2018		0.016
		(0.024)
Sub-province fixed effects	Yes	Yes
R^2	0.01	0.04
N	6,713	6,713

Note: 1. Standard errors clustered at the household level are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the linear probability model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.6: **The effect of mining pollution level on sickness in a nonlinear form**

	Sick		
Variable name	(1)	(2)	(3)
ln(distance to the nearest mine)	-0.033***	-0.025***	-0.024**
	(0.009)	(0.008)	(0.009)
As above trigger value	0.062**	0.053**	
	(0.026)	(0.026)	
Hg above action value	-0.024		-0.009
	(0.021)		(0.024)
Individual is female	0.014**	0.014**	0.014
	(0.006)	(0.006)	(0.009)
Individual's age (years)	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)
Individual's education (years)	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
Consumption per capita	-0.042***	-0.042***	-0.042***
	(0.014)	(0.014)	(0.012)
Lives in rural area	0.041	0.007	0.013
	(0.042)	(0.031)	(0.034)
2010	0.029	0.029	0.029
	(0.019)	(0.019)	(0.036)
2014	-0.013	-0.014	-0.013
	(0.015)	(0.015)	(0.018)
2016	-0.021	-0.022	-0.020
	(0.015)	(0.015)	(0.023)
2018	0.015	0.014	0.015
	(0.016)	(0.016)	(0.024)
Sub-province fixed effects	Yes	Yes	Yes
R^2	0.04	0.04	0.04
N	6,713	6,713	6,713

Note: 1. In order to allow nonlinear effects of pollution, we estimate the intensity of pollution by creating dummy variables for different levels of pollution. The mean of mercury is above the action value whereas the mean of arsenic is above the precaution (permissible) value. All other heavy metals are within the precaution level, indicating that soil is not polluted by cadmium, copper, nickel, lead and zinc.
2. Distance to each heavy metal is 6 kilometres.
3. Standard errors clustered at the household level are recorded in the parentheses.
4. Marginal effects are calculated at the means of all other covariates from the linear probability model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.7: **The effect of mining pollution on sickness for different age groups**

	Age: 0-14 years		Age: 15-65 years		Age: 65+ years	
Variable names	(1)	(2)	(3)	(4)	(5)	(6)
ln(distance to the nearest mine)	-0.035**	-0.035**	-0.008	-0.012	-0.090	-0.083
	(0.015)	(0.014)	(0.007)	(0.007)	(0.063)	(0.064)
Individual is female		-0.000		0.021***		-0.096*
		(0.011)		(0.007)		(0.057)
Individual's age (years)		-0.007***		0.004***		0.002
		(0.002)		(0.001)		(0.006)
Individual's education (years)		0.001		-0.001		-0.015*
		(0.003)		(0.001)		(0.009)
Consumption per capita		0.019		-0.027**		-0.062
		(0.031)		(0.013)		(0.078)
Lives in rural area		0.011		0.033		0.269
		(0.031)		(0.049)		(0.384)
2010		0.062*		0.005		0.140
		(0.035)		(0.022)		(0.134)
2014		0.047*		-0.046***		-0.123
		(0.027)		(0.017)		(0.110)
2016		0.034		-0.057***		-0.087
		(0.030)		(0.018)		(0.115)
2018		0.046		-0.015		-0.025
		(0.031)		(0.019)		(0.118)
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.03	0.04	0.01	0.04	0.05	0.11
N	1,915	1,915	4,536	4,536	262	262

Note: 1. Standard errors clustered at the household level are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the linear probability model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.8: **The effect of mining pollution on body system illnesses**

Variable name	Respiratory (1)	Digestive (2)	Cardiovascular (3)	Other illnesses (4)
ln(distance to the nearest mine)	-0.0127***	-0.0012	-0.0006***	-0.0056***
	(0.0012)	(0.0046)	(0.0014)	(0.0006)
Individual is female	0.0049	0.0006	0.0067	0.0025
	(0.0051)	(0.0012)	(0.0042)	(0.0032)
Individual's age (years)	0.0002	0.0004***	0.0017***	0.0017***
	(0.0003)	(0.0001)	(0.0004)	(0.0003)
Individual's education (years)	-0.0028***	-0.0002	-0.0021***	-0.0016***
	(0.0006)	(0.0003)	(0.0004)	(0.0005)
Consumption per capita	-0.0105*	-0.0011	-0.0100***	-0.0201***
	(0.0056)	(0.0023)	(0.0047)	(0.0052)
Lives in rural area	-0.0076	0.0393***	-0.0132**	-0.0161**
	(0.0053)	(0.0073)	(0.0049)	(0.0066)
2010	0.0202***	-0.0030	-0.0112	0.0206
	(0.0072)	(0.0024)	(0.0050)	(0.0137)
2014	0.0094	-0.0112**	-0.0092**	0.0120**
	(0.0064)	(0.0043)	(0.0068)	(0.0047)
2016	0.0146**	-0.0104**	-0.0126	0.0017
	(0.0068)	(0.0041)	(0.0059)	(0.0077)
2018	0.0152*	-0.0048	-0.0023*	0.0199*
	(0.0073)	(0.0050)	(0.0070)	(0.0102)
Sub-province fixed effects	Yes	Yes	Yes	Yes
R^2	0.02	0.01	0.04	0.02
N	6,713	6,713	6,713	6,713

Note: 1. Standard errors clustered at the household level are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the linear probability model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.9: **The effect of mining scale on sickness**

	Mining license holders		Small-scale mining	
Variable name	(1)	(2)	(3)	(4)
ln(distance to the nearest mine)	-0.010	-0.013	-0.018**	-0.024***
	(0.014)	(0.014)	(0.009)	(0.009)
Individual is female		0.002		0.031***
		(0.008)		(0.011)
Individual's age (years)		0.004***		0.004***
		(0.001)		(0.001)
Individual's education (years)		-0.006***		-0.008***
		(0.001)		(0.002)
Consumption per capita		-0.043**		-0.047**
		(0.017)		(0.024)
Lives in rural area		-0.028		-0.077***
		(0.025)		(0.030)
2010		0.091***		-0.021
		(0.028)		(0.025)
2014		0.004		-0.031
		(0.017)		(0.023)
2016		-0.004		-0.045*
		(0.017)		(0.024)
2018		0.019		0.032
		(0.018)		(0.027)
Sub-province fixed effects	Yes	Yes	Yes	Yes
R^2	0.01	0.05	0.01	0.05
N	3,370	3,370	2,835	2,835

Note: 1. Standard errors clustered at the household level are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the linear probability model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.10: **The effect of mineral type on sickness**

	Gold		Limestone		Other	
Variable name	(1)	(2)	(3)	(4)	(5)	(6)
ln(distance to the nearest mine)	-0.016**	-0.020***	-0.010	-0.028	0.005	-0.011
	(0.007)	(0.007)	(0.060)	(0.052)	(0.046)	(0.049)
Individual is female		0.017**		0.002		0.013
		(0.008)		(0.011)		(0.012)
Individual's age (years)		0.004***		0.006***		0.003**
		(0.001)		(0.001)		(0.001)
Individual's education (years)		-0.007***		-0.007***		-0.005***
		(0.001)		(0.002)		(0.002)
Consumption per capita		-0.033*		-0.083***		-0.026
		(0.018)		(0.029)		(0.028)
Lives in rural area		-0.099***		0.077		0.046
		(0.013)		(0.160)		(0.117)
2010		-0.009		0.101**		0.056
		(0.022)		(0.046)		(0.043)
2014		-0.007		-0.052*		0.005
		(0.019)		(0.028)		(0.036)
2016		-0.023		-0.050*		0.014
		(0.019)		(0.029)		(0.037)
2018		0.040*		-0.061**		-0.004
		(0.021)		(0.027)		(0.038)
Sub-province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.01	0.04	0.00	0.13	0.00	0.03
N	4,359	4,359	904	904	1,450	1,450

Note: 1. Standard errors clustered at the household level are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

Table A.11: **The effect of mining pollution on sickness – Principal component analysis**

	Sick	
Variable name	(1)	(2)
ln(distance to the nearest mine)	-0.017**	-0.019***
	(0.007)	(0.007)
Component: Cu, Hg, Ni	-0.015	-0.009
	(0.013)	(0.012)
Individual is female		0.013**
		(0.005)
Individual's age (years)		0.003***
		(0.000)
Individual's education (years)		-0.004***
		(0.001)
Consumption per capita		-0.026***
		(0.009)
Lives in rural area		0.038
		(0.081)
2010		0.026*
		(0.014)
2014		-0.010
		(0.012)
2016		-0.017
		(0.013)
2018		0.014
		(0.013)
Sub-province fixed Effects	Yes	Yes
Psedu R^2	0.02	0.07
N	6,713	6,713

Note: 1. Standard errors clustered at the household level are recorded in the parentheses.
2. Individuals living at a proximity of up to 6 km are analyzed in this table.
3. Marginal effects are calculated at the means of all other covariates from the logit model.
* p <0.10, ** p <0.05, *** p <0.01.

References

Anderson, M. L. (2020). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association*, 18(4):1886–1927.

Aragón, F. M. and Rud, J. P. (2015). Polluting industries and agricultural productivity: Evidence from mining in Ghana. *The Economic Journal*, 126(597):1980–2011.

California Air Resources Board (1990). Air Resources Board Sets New Warning Level for Urban Smog. News release, California Air Resources Board, California, USA. Available from https://bit.ly/37ddfam [Accessed on: 14 December 2020].

Caravanos, J., Ericson, B., Ponce-Canchihuamán, J., Hanrahan, D., Block, M., Susilorini, B., and Fuller, R. (2013). Rapid assessment of environmental health risks posed by mining operations in low-and middle-income countries: selected case studies. *Environmental Science and Pollution Research*, 20(11):7711–7718.

Charpin, D., Kleisbauer, J., Fondarai, J., Graland, B., Viala, A., and Gouezo, F. (1988). Respiratory symptoms and air pollution changes in children: the Gardanne coal-basin study. *Archives of Environmental Health: An International Journal*, 43(1):22–27.

Chay, K. Y. and Greenstone, M. (2003). The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *The Quarterly Journal of Economics*, 118(3):1121–1167.

Cordier, S., Garel, M., Mandereau, L., Morcel, H., Doineau, P., Gosme-Seguret, S., Josse, D., White, R., and Amiel-Tison, C. (2002). Neurodevelopmental investigations among methylmercury-exposed children in French Guiana. *Environmental Research*, 89(1):1–11.

Cordier, S., Theriault, G., and Iturra, H. (1983). Mortality patterns in a population living near a copper smelter. *Environmental Research*, 31(2):311–322.

Currie, J. (2011). Inequality at birth: Some causes and consequences. *American Economic Review*, 101(3):1–22.

Currie, J., Davis, L., Greenstone, M., and Walker, R. (2015). Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings. *American Economic Review*, 105(2):678–709.

Currie, J., Hanushek, E. A., Kahn, E. M., Neidell, M., and Rivkin, S. G. (2009a). Does pollution increase school absences? *The Review of Economics and Statistics*, 91(4):682–694.

Currie, J., Neidell, M., and Schmieder, J. F. (2009b). Air pollution and infant health: Lessons from New Jersey. *Journal of Health Economics*, 28(3):688–703.

Currie, J. and Schmieder, J. F. (2009). Fetal exposures to toxic releases and infant health. *American Economic Review*, 99(2):177–83.

Currie, J., Zivin, J. G., Mullins, J., and Neidell, M. (2014). What do we know about short-and long-term effects of early-life exposure to pollution? *Annual Review of Resource Economics*, 6(1):217–247.

Deaton, A. (1997). *The analysis of household surveys: A microeconometric approach to development policy*. World Bank, Maryland, USA.

Doojav, G.-O., Tsenguunjav, B. Tsenddorj, D., and Munkhbayar, B. (2017). Mongolia's economic cycle and the factors affecting it. Research report, Bank of Mongolia, Ulaanbaatar, Mongolia.

EITIM (2020). Report Data Analysis - Infographics. Database, EITIM, Mongolia Extractive Industries Transparency Initiative, Ulaanbaatar, Mongolia. Available from http://bit.ly/3oTukMA [Accessed on: 24 November 2020].

GDEH (2012). Geo-Database on Ecological Health. Database, GDEH, Geo-Database on Ecological Health, Ministry of Environment and Green Development, Ulaanbaatar Mongolia. Available from http://bit.ly/3r43aV2 [Accessed on: 16 March 2020].

Gradstein, M. and Klemp, M. (2020). Natural resource access and local economic growth. *European Economic Review*, 127:103441.

Graff Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–73.

Graff Zivin, J. and Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, 51(3):689–730.

Greenstone, M. and Jack, B. K. (2015). Envirodevonomics: A research agenda for an emerging field. *Journal of Economic Literature*, 53(1):5–42.

Hanna, R. and Oliva, P. (2015). The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City. *Journal of Public Economics*, 122:68–79.

Harada, M., Nakachi, S., Cheu, T., Hamada, H., Ono, Y., Tsuda, T., Yanagida, K., Kizaki, T., and Ohno, H. (1999). Monitoring of mercury pollution in Tanzania: relation between head hair mercury and health. *Science of the Total Environment*, 227(2-3):249–256.

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.

Hendryx, M. and Ahern, M. M. (2009). Mortality in Appalachian coal mining regions: the value of statistical life lost. *Public Health Reports*, 124(4):541–550.

Hill, E. L. (2018). Shale gas development and infant health: evidence from Pennsylvania. *Journal of Health Economics*, 61:134–150.

Hota, P. and Behera, B. (2015). Coal mining in Odisha: An analysis of impacts on agricultural production and human health. *The Extractive Industries and Society*, 2(4):683–693.

Hota, P. and Behera, B. (2016). Opencast coal mining and sustainable local livelihoods in Odisha, India. *Mineral Economics*, 29(1):1–13.

Hugjliin Ezed NGO and Swiss Agency for Development and Cooperation, (SDC) (2010). Socio-economic baseline survey of Bornuur Soum, Tuv province, Mongolia. Research report, Hugjliin Ezed NGO and Swiss Agency for Development and Cooperation, Ulaanbaatar Mongolia. Available from https://bit.ly/3f0vGTS [Accessed on: 17 May 2021].

Human Rights Commission and SDC (2012). The human rights situation in small-scale mining in Mongolia. Research report, Human Rights Commission and Swiss Agency for Development and Cooperation, Ulaanbaatar Mongolia. Available from https://bit.ly/2RsfbGU [Accessed on: 17 May 2021].

Janke, K., Propper, C., and Henderson, J. (2009). Do current levels of air pollution kill? The impact of air pollution on population mortality in England. *Health Economics*, 18(9):1031–1055.

Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society*, A374(2065):20150202.

Landrigan, P. J., Fuller, R., Acosta, N. J., Adeyi, O., Arnold, R., Baldé, A. B., Bertollini, R., Bose-O'Reilly, S., Boufford, J. I., Breysse, P. N., et al. (2018). The lancet commission on pollution and health. *The Lancet*, 391(10119):462–512.

Levasseur, P., Erdlenbruch, K., and Gramaglia, C. (2021). The health and socioeconomic costs of exposure to soil pollution: evidence from three polluted mining and industrial sites in europe. *Journal of Public Health*, pages 1–14.

Li, B. G., Gupta, P., and Yu, J. (2017). From natural resource boom to sustainable economic growth: Lessons from Mongolia. *International Economics*, 151:7–25.

Li, Z., Ma, Z., van der Kuijp, T. J., Yuan, Z., and Huang, L. (2014). A review of soil heavy metal pollution from mines in China: Pollution and health risk assessment. *Science of the Total Environment*, 468:843–853.

Marcus, M. (2021). Going beneath the surface: Petroleum pollution, regulation, and health. *American Economic Journal: Applied Economics*, 13(1):1–37.

MASM (2019). Soil quality. Soil Pollutants Permissible Value MNS 5850:2019. Mongolian national standard, MASM, Mongolian Agency for Standardization and Meteorology, Ulaanbaatar, Mongolia. Available from https://bit.ly/37Q9XsZ [Accessed on: 12 August 2020].

National Statistics Office (2018). Mongolia Household Socio-Economic Survey 2016. Report, NSO, National Statistics Office, Ulaanbaatar, Mongolia.

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, 23(6):1209–1236.

Persico, C. L. and Venator, J. (2021). The effects of local industrial pollution on students and schools. *Journal of Human Resources*, 56(2):406–445.

Pless-Mulloli, T., Howel, D., King, A., Stone, I., Merefield, J., Bessell, J., and Darnell, R. (2000). Living near opencast coal mining sites and children's respiratory health. *Occupational and Environmental Medicine*, 57(3):145–151.

Rau, T., Urzúa, S., and Reyes, L. (2015). Early exposure to hazardous waste and academic achievement: Evidence from a case of environmental negligence. *Journal of the Association of Environmental and Resource Economists*, 2(4):527–563.

Romero, M. and Saavedra, S. (2015). The effect of gold mining on the health of newborns. *Unpublished manuscript. Stanford University, Department of Economics, Stanford, CA.*

Saha, S., Pattanayak, S. K., Sills, E. O., and Singha, A. K. (2011). Under-mining health: Environmental justice and mining in India. *Health & Place*, 17(1):140–148.

Solis, M. T., Yuen, E., Cortez, P. S., and Goebel, P. J. (2000). Family poisoned by mercury vapor inhalation. *The American Journal of Emergency Medicine*, 18(5):599–602.

Telmer, K. and Stapper, D. (2012). Reducing mercury use in artisanal and small-scale gold mining. A practical guide, Artisanal Gold Council and United Nations Environment Programme, Geneva, Switzerland. Available from https://bit.ly/3woXeZr [Accessed on: 4 April 2021].

Thermo Fisher Scientific (2021). Atomic Absorption Spectrometry (AAS) Information. Technical report, Thermo Fisher Scientific. Available from https://bit.ly/3fY0THM [Accessed on: 6 June 2021].

U.S. Agency for Toxic Substances and Disease Registry (2007). Arsenic CAS:7440-38-2. Public health statement, United States Agency for Toxic Substances and Disease Registry, Atlanta, USA. Available from https://bit.ly/3qR62Ve [Accessed on: 12 August 2020].

U.S. Environmental Protection Agency (2006). Guideline for Reporting of Daily Air Quality - Air Quality Index (AQI). Guideline, United States Environmental Protection Agency, North Carolina, USA. Available from https://bit.ly/37hesxi [Accessed on: 14 December 2020].

Van Straaten, P. (2000). Mercury contamination associated with small-scale gold mining in Tanzania and Zimbabwe. *Science of the Total Environment*, 259(1-3):105–113.

Vandermoere, F. (2008). Psychosocial health of residents exposed to soil pollution in a flemish neighbourhood. *Social Science & Medicine*, 66(7):1646–1657.

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

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach.* Nelson Education, OH, USA.