Policy Research Working Paper 7250

# The Local Socioeconomic Effects of Gold Mining

## Evidence from Ghana

*Punam Chuhan-Pole*
*Andrew Dabalen*
*Andreas Kotsadam*
*Aly Sanoh*
*Anja Tolonen*

WORLD BANK GROUP
Africa Region
Office of the Chief Economist
April 2015

Policy Research Working Paper 7250


## Abstract

Ghana is experiencing its third gold rush, and this paper sheds light on the socioeconomic impacts of this rapid expansion in industrial production. Using a rich dataset consisting of geocoded household data combined with detailed information on gold mining activities, the authors conduct two types of difference-in-differences estimations that provide complementary evidence. The first is a local-level analysis that identifies an economic footprint area very close to a mine, and the second is a district-level analysis that captures the fiscal channel. The results indicate that men are more likely to benefit from direct employment as miners compared to men further away, and that women in mining communities may more likely gain from indirect employment opportunities and earn cash for work. Authors also find that infant mortality rates decrease significantly in mining communities, compared to the evolution in communities further away.


This paper is a product of the Office of the Chief Economist, Africa Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at andreas.kotsadam@econ.uio.no, anja.tolonen@economics.gu.se, and pchuhan@worldbank.org.

# The Local Socioeconomic Effects of Gold Mining: Evidence from Ghana

JEL Classification: J16, J21, O13, O18

## 1 Introduction

The mining sector in Africa is growing rapidly and is the main recipient of foreign direct investment (World Bank 2011). The welfare effects of this sector are not well understood, although a literature has recently developed around this question. The main contribution of this paper is to shed light on the welfare effects of gold mining in a detailed, in-depth country study of Ghana, a country with a long tradition of gold mining and a recent, large expansion in capital-intensive and industrial-scale production.

A second contribution of this paper is to show the importance of decomposing the effects with respect to distance from the mines. Given the spatial heterogeneity of the results, we explore the effects in an individual-level, difference-in-differences analysis by using spatial lag models to allow for nonlinear effects with distance from mine. We also allow for spillovers across districts, in a district-level analysis. We use two complementary geocoded household data sets to analyze outcomes in Ghana: the Demographic and Health Survey (DHS) and the Ghana Living Standard Survey (GLSS), which provide information on a wide range of welfare outcomes.

The paper contributes to the growing literature on the local effects of mining. Much of the academic interest in natural resources is focused on country-wide effects, and this research discusses whether the discovery of natural resources is a blessing or a curse to the national economy. Natural resource dependence at the national level has been linked to worsening economic and political outcomes, such as weaker institutions, and more corruption and conflict (see Frankel 2012 and van der Ploeg 2011 for an overview). While all these effects can have household-level implications, fewer analyses have, thus far, analyzed the geographic dispersion of such impacts. A recent literature on the local and subnational effects of natural resources contributes to the understanding of such effects (for example Aragón and Rud 2013, 2015; Axbard et al., 2016; Benshaul-Tolonen 2018, 2019; Caselli and Michaels 2013; Corno and de Walque 2012; Fafchamps et al. 2016; Kotsadam and Tolonen 2016; Loyaza et al 2013; Michaels 2011; von der Goltz and Barnwal 2019; Wilson 2012). A growing number of papers explore the mining industry, in particular, see Aragón, Chuhan-Pole, and Land (2015) for an overview. We contribute to this literature by showing the importance of analyzing local level effects in addition to district level effects in a one-country case study.

Aragón and Rud (2013) provided the seminal work exploring the economic effects of one very large mine in Peru. They find that the expansion of the mine had poverty-reducing effects, but only in conjunction with policies for local procurement. Moreover, some of the mining-related papers have focused on mining in an African context, exploring a range of outcomes, including HIV-transmission and sexual risk taking (Corno and de Walque 2012; Wilson 2012), women's empowerment (Benshaul-Tolonen 2018), infant mortality (Benshaul-Tolonen, 2019) and labor market outcomes (Kotsadam and Tolonen 2016). Mining is also associated with more economic activity measured by nightlights (Benshaul-Tolonen, 2019; Mamo et al, 2019).

Kotsadam and Tolonen (2016) use DHS data from Africa, and find that mine openings cause women to shift from agriculture to service production and that women become more likely to work for cash and year-round as opposed to seasonally. Continuing this analysis, Benshaul-Tolonen (2018) explores the links between mining and female empowerment in eight gold-producing countries in East and West Africa, including Ghana. Women in gold mining communities have more diversified labor markets opportunities, better access to health care, and are less likely to accept domestic violence. In addition, infant mortality rates decrease with up to 50% in mining communities, from very high initial levels (Benshaul-Tolonen, 2019). In a study that focuses exclusively on Ghana, Aragón and Rud (2013) explore the link between pollution from mining and agricultural productivity. The results point toward decreasing agricultural productivity because of environmental pollution and soil degradation, which could have negative welfare effects on households that do not engage in mining activities or in indirectly stimulated sectors. Lower productivity in agriculture could potentially push households to engage in mining-related sectors, in addition to pull factors such as higher wage earnings in the stimulated sectors.

We explore the effects of mining activity on employment, earnings, expenditure, and children's health outcomes in local communities and in districts with gold mining. We combine the DHS and GLSS with production data for 17 large-scale gold mines in Ghana. We find that a new large-scale gold mine changes economic outcomes, such as access to employment and cash earnings. In addition, it raises local wages and expenditure on housing and energy.

An important welfare indicator in developing countries is infant mortality, and we note a large and significant decrease in mortality rates among young children, at both the local and district levels.[1] We hypothesize that increased access to prenatal care is one of the mechanisms behind the increased survival rate.

[1] In the 2010 Ghana population census average district size is 112,000

We suggest interpreting the local effects as being additional to the district-level effects; that is, the mine affects the mining district predominantly through the fiscal channel, and local mining communities mainly through employment generation, and other localized factors. Overall, the results are more robustly estimated at the district level than at the individual level, and we find no indications of positive spillover effects across districts. This is in line with a public spending hypothesis, where mining districts benefit more than adjacent non-mining districts through the fiscal revenue channel, since 10 percent of mining royalties are redistributed to mining districts.

**2 Gold mining in Ghana**

Ghana has a long tradition of gold mining and has produced a substantial portion of the world's gold for over 1,000 years (see Hilson [2002] for an extensive overview of gold production in Ghana). During colonial British rule, the country was named the Gold Coast Colony, and gold production was booming. The first gold rush occurred between 1892 and 1901, and the second after World War I. Gold production decreased at the dawn of independence in 1957, and remained low until the 1980s. Over the last 20 years, Ghana has been experiencing its third gold rush. During this period, annual gold production has increased by 700 percent, as shown in Figure 1. It is the expansion that has happened during this recent gold rush that is used in this analysis to understand the socioeconomic effects of mining. The high international gold price was a driving factor in the expansion of small-scale mining, such as the 2,700% increase in gold mining territory around the Offin River between 2008 and 2012 (Hausermann et al., 2018). Between 2006 and 2012, two large-scale mines opened in Ghana, but no mine closed down (Table 1) possibly due to the high gold price increasing profitability and extending life length.

The expansion across artisanal small-, medium-, and large-scale mining contributed to an increase in total production that rose from 541,147 oz in 1990 to 3,119,823 oz in 2009 according official Ghana statistics (Bloch and Owusu, 2012). This production increase led to an increased sector contribution to GDP from 4,83% (1990) to 5,78% (2009), alongside export value of US$304m in 1990, US$702m in 2000, and US$2246 m in 2008, reaching 43% of national exports in 2008. Mining related foreign direct investment (FDI) also rose from US$165m to US$762m between 1995 and 2009. Mining was the dominating sector with between 48% and 94% of mining FDI to total FDI from 1995 to 2007, until the country saw an incredible increase in non-mining foreign direct investment (Bloch and Owusu, 2012), following the discovery of oil in 2007.

**Figure 1 Ghana's annual gold production and world price of gold**


Ghana is the second-largest gold producer in Africa after South Africa, with gold production averaging 77 tons per year (Gajigo, Mutambatsere, and Mdiaya 2012). In 2011, Ghana's mineral sector accounted for about 14 percent of total tax revenues and 5.5 percent of the gross domestic product (GDP) (Bermúdez-Lugo 2011), as well as 44 percent of Ghanian exports (Gajigo, Mutambatsere, and Mdiaya 2012). This makes the gold mining industry one of the country's most important industries, and an essential industry to study.

Similar to gold mining in other African countries (see Gajigo, Mutambatsere, and Mdiaya [2012] for an overview), the sector is highly capital intensive, and direct employment generation is, relative to its economic importance, limited. In 2010, it was estimated that about 20,000 Ghanaian nationals—0.08 percent of the population—were employed in large-scale mining (Bermudez-Lugo 2010), despite accounting for 5.5 percent of GDP. Nonetheless, the spillovers to other sectors of the economy may be substantial, because nonnationals also work in the mines and wages are relatively high. Aryee (2001) estimates that, between 1986 and 1998, large-scale mining injected over US$300 million into the national economy from salaries alone.

Beyond direct and indirect employment effects, the mining industry is connected to the wider economy via taxes and royalties. Ghana has been highlighted as a good example of how mineral-rich countries can distribute mining wealth, since a proportion of the rents are distributed to the local communities (Standing and Hilson 2013). The mining royalty paid by mining companies in Ghana was 3 percent until 2010, which was the average rate for gold production in Africa (Gajigo, Mutambatsere, and Mdiaya 2012), but increased to 5 percent in 2010 (Standing and Hilson 2013). Of this money, 80 percent goes to the general government budget, 10 percent goes to the administration of mining oversight, and 10 percent supports district administration (Garvin et al. 2009). Between 1993 and 1998, about US$17 million was distributed to local mining communities (Aryee 2001). While considered a model of best practice, there is still a worry that the beneficial effects of allocations to the districts are undermined by elite capture and corruption at the district level (Standing and Hilson 2013). For our analysis, the scheme implies that it may be necessary to conduct a district-level analysis in addition to the more local-level analyses.

12 currently active mines dominate the sector, and there are an additional five suspended mines that have been in production in recent decades. Table 1 presents a full list of the mines, the year they opened, and their status as of December 2012. Company name and country are for the main shareowner in the mine. Most of these 17 mines have foreign ownership, such as Australian, Canadian, or South African, sometimes in partnership with Ghanaian firms or the Ghanaian state. Most are open-pit mines, although a few consist of a combination of open-pit and underground operations.

**Table 1 Gold Mines in Ghana**

| Name | Opening year | Closing year | Company | Country |
|---|---|---|---|---|
| Ahafo | 2006 | active | Newmont Mining Corp. | USA |
| Bibiani | 1998 | active | Noble Mineral Resources | Australia |
| Bogoso Prestea | 1990 | active | Golden Star Resources | USA |
| Chirano | 2005 | active | Kinross Gold | Canada |
| Damang | 1997 | active | Gold Fields Ghana Ltd. | South Africa |
| Edikan (Ayanfuri) | 1994 | active | Perseus Mining | Australia |
| Iduapriem | 1992 | active | AngloGold Ashanti | South Africa |
| Jeni (Bonte) | 1998 | 2003 | Akrokeri-Ashanti | Canada |
| Konongo | 1990 | active | LionGold Corp. | Singapore |
| Kwabeng | 1990 | 1993 | Akrokeri-Ashanti | Canada |
| Nzema | 2011 | active | Endeavour | Canada |
| Obotan | 1997 | 2001 | PMI Gold | Canada |
| Obuasi | 1990 | active | AngloGold Ashanti | South Africa |
| Prestea Sankofa | 1990 | 2001 | Anglogold Ashanti | South Africa |
| Tarkwa | 1990 | active | Gold Fields Ghana Ltd. | South Africa |
| Teberebie | 1990 | 2005 | Anglogold Ashanti | South Africa |
| Wassa | 1999 | active | Golden Star Resources | USA |

*Source:* InterraRMG 2013.
*Note:* Active is production status as of December 2012, the last available data point.

Alongside the large-scale, capital-intensive mining industry in Ghana, there is an artisanal and small-scale mining sector (ASM). ASM activities were legalized in 1984, when the state loosened its monopoly on gold mining. In Ghana, as in many other African countries, the sector is an important employer (ILO 1999). It is estimated that around 1 million people in Ghana support themselves with revenues from ASM activities.

The sector is associated with several hazardous labor conditions, however. This includes child labor, mercury exposure, and risk of mine collapse (Hilson 2009). The ASM and the large-scale mining sector sometimes thrive side by side, but sometimes competing interests lead to conflict between the two sectors, such as around Prestea, where domestic *galamsey* miners (informal small-scale miners) have been in conflict with the multinational concession owner (Hilson and Yakoleva 2007).

In this analysis, we focus solely on large-scale mining. We understand, however, that small- and large-scale operations may be geographically correlated. Assuming that the start of a large-scale mine does not affect the likelihood or viability of artisanal and small-scale mining, it is not a threat to our identifying assumptions. However, should ASM respond to large-scale activities, either by increasing or decreasing activity in the close geographic area, we will end up estimating the impact of these sectors jointly. In a later stage, should the opportunity arise, we encourage researchers to try to disentangle the effects of small-scale and large-scale mining.

**3 Data**

To conduct this analysis, we combine different data sources using spatial analysis. The main mining data is a dataset from InterraRMG covering all large-scale mines in Ghana, explained in more detail in section 3.1. This dataset is linked to survey data from the DHS and GLSS, using spatial information. Geographical coordinates of enumeration areas in GLSS are from Ghana Statistical Services (GSS).[2] Point coordinates (global positioning system [GPS]) for the surveyed DHS clusters[3] allow us to match all individuals to one or several mineral mines. We do this in two ways.

First, we calculate distance spans from an exact mine location given by its GPS coordinates, and match surveyed individuals to mines. These are concentric circles with radiuses of 10, 20, and 30 kilometers (km), and so on, up to 100 km and beyond. In the baseline analysis where

[2] The data was shared by Aragón and Rud (2013)

[3] Both the DHS and GLSS enumeration area coordinates have a 1-5 km offset. The DHS clusters have up to 10km displacement in 1% of the cases.

we use a cutoff distance of 20 km, we assume there is little economic footprint beyond that distance. Of course, any such distance is arbitrarily chosen, which is why we try different specifications to explore the spatial heterogeneity by varying this distance (using 10 km, 20 km, through 50 km) as well as a spatial lag structure (using 0 to 10 km, 10 to 20 km, through 40 to 50 km distance bins).[4]

Second, we collapse the DHS mining data at the district level.[5] The number of districts has changed over time in Ghana, because districts with high population growth have been split into smaller districts. To avoid endogeneity concerns, we use the baseline number of districts that existed at the start of our analysis period, which are 137. Eleven of these districts have industrial mining. Because some mines are close to district boundaries, we additionally test whether there is an effect in neighboring districts.

### 3.1 Resource data

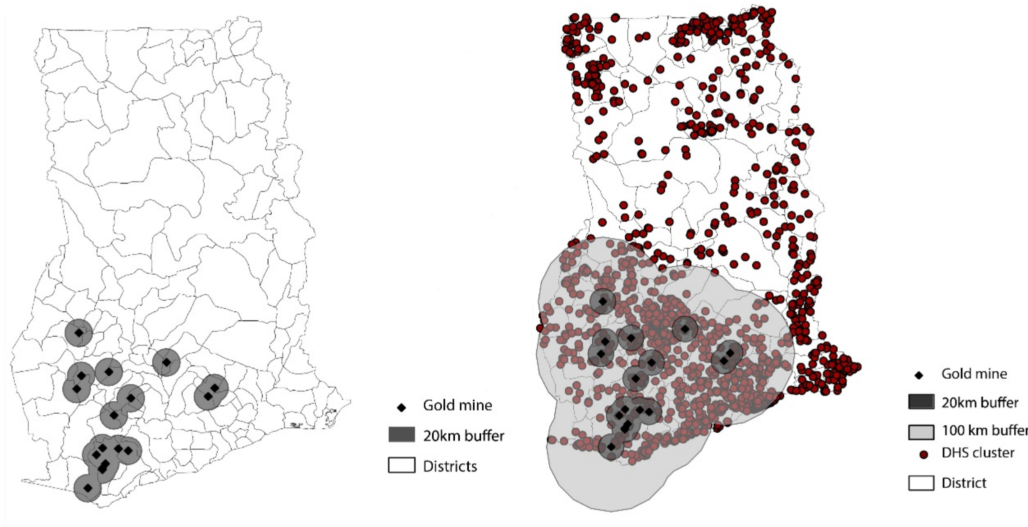
The Raw Materials Data are from InterraRMG (2013). The data set contains information on past or current industrial mines. All mines have information on annual production volumes, ownership structure, and GPS coordinates on location. We complete this data with exact geographic location data from MineAtlas (2013), where satellite imagery shows the actual mine boundaries, which allows us to identify and update the center point of each mine. The production data and ownership information are double-checked against the companies' annual reports.

For Ghana, this exercise results in 17 industrial mines tracked over time. We have annual production levels from 1990 until 2012. As mentioned, Table 1 shows the mining companies active in Ghana during recent decades, with opening and closing years (although some were closed in between, and are not presented in the table). Figure 2 shows the geographic distribution of these mines.

**Figure 2 Gold mines and DHS clusters in Ghana**

**Panel A Gold mines and 20 km buffer zones   Panel B Gold mines, DHS clusters, and 100 km buffer zones**

[4] The distances are radii from mine center point, and form concentric circles around the mine.
[5] The DHS and the GLSS data are representative at the regional level, and not at the district level. Since the regional level is too aggregated, we do the analysis at the district level, but note that the sample may not be representative.




*Note:* Panel A shows the location of the gold mines that were active during the study period. Around each circle, a 20-km radius is marked. These 40-km-wide areas are the baseline treatment areas in the analysis. Panel B shows the 100-km treatment areas and the distribution of the DHS clusters. Road data is an alternative way of defining distance from mines, but time series data on roads is not available.

### 3.2 Household data

We use microdata from the DHS, obtained from standardized surveys across years and countries. We combine the respondents from all four DHS standard surveys in Ghana for which there are geographic identifiers. The total data set includes 19,705 women (of which 12,392 live within 100 km of a mine) aged 15–49 from 137 districts. They were surveyed in 1993, 1998, 2003, and 2008,[6] and live in 1,623 survey clusters. Since the DHS surveys focus on women, the surveys of women will be the main source of data. However, we also use the surveys of men, which give us data from the same four survey years, but with a total number of 12,294 individuals, of which 7,491 men live within 100 km of a mine. In addition, the DHS data collect records of all children born within the five years prior to the surveying. Of the 12,174 children born to the surveyed women within the last five years, 6,888 were born to women currently residing within 100 km of a mine. See Appendix table 1 for definition of outcome variables.

We complement the analysis with household data from the GLSS collected in the years—1998–99, 2004–05, and 2012–13. These data are a good complement to the DHS data, because they

[6] The first mines were opened in 1990, prior to the first household survey. Ten mines were opened after the first DHS in 1993. There is less variation in the data set using GLSS where the first households were surveyed in 1998, i.e. 8 years after the first mine opened. However, the DHS data include births recorded from 1987, which is prior to all mine openings.

have a stronger focus on all households' members, rather than focusing only on women and young children. In addition, they provide more detailed information on labor market participation, such as exact profession (where, for example, being a miner is a possible outcome), hours worked, and a wage indicator. The data estimate household expenditure and household income. Wages, income, and expenditure can, however, be difficult to measure in economies where nonmonetary compensation for labor and subsistence farming are common practices.

## 4 Empirical Strategies

### 4.1 Individual-level difference-in-differences

Time-varying data on production and repeated survey data allow us to use a difference-in-differences approach.[7] However, due to the spatial nature of our data and the fact that some mines are spatially clustered, we use a strategy developed by Benshaul-Tolonen (2018). The difference-in-difference model compares the treatment group (close to mines) before and after the mine opening, while removing the change that happens in the control group (far away from mines) over time under the assumption that such changes reflect underlying temporal variation common to both treatment and control areas.

We limit the data to include households within 100 km of a mine location and estimate the following:

$$Y_{ivt} = \beta 0 + \beta_1 \cdot active_t + \beta_2 \cdot mine + \beta_3 \cdot \; active_t * mine + \alpha_d + g_t + \lambda X_i + \varepsilon_{ivt}, \qquad (1)$$

where the outcome of an individual $i$ in cluster $v$, and for year $t$ is regressed on district and year fixed effects, a dummy for whether the respondent lives within 20 km of a mine (which is a current or future mine[8]), a dummy for whether the mine is active at the time of the survey (active), an interaction term between active mines and living close to a mine ($active_t * mine$), and a vector of individual-level control variables. *Mine* is the terminology chosen to capture a known gold resource in the ground, regardless of whether it is being extracted or not. In all regressions, we also control for living in an urban area, years of education, and age.

[7] We have not done a synthetic control approach because of limited ability to explore pretreatment trends.

[8] A current or future mine could also be called a deposit. We have chosen against this terminology as there may be known deposits in Ghana that never started actively producing. Such deposits are not included in our dataset and thus not in the analysis. In addition, deposits may be considered all existing geological deposits whether or not known to man, or those that have been discovered. The latter being truly exogenous, while known deposits and "mines" according to our definitions are not truly exogenous.

The choice of district – rather than cluster – fixed effect is informed by the understanding that meaningful time-invariant factors - such as mining laws, level of development, local political institutions, norms regarding environment, women's participation in the labor market, etc. - that influence exploitation of the mine happens at the district level. Including district fixed effects, we control for various institutional and cultural factors at the district level that are stable over time. Including district fixed effects also ensures that we are not only capturing effects from transfers or the fiscal system as we compare individuals within the same districts. With this method we capture the geographic spillover effects in the vicinity of the mine. Moreover, cluster fixed effects are not possible because of clusters are not repeatedly sampled over time. However, since the estimation is at individual level, all standard errors are clustered at the DHS cluster level.

The sample is restricted to individuals living within 100 km of a deposit location (*mine*), so many parts of Northern Ghana where there are few gold mines are not included in the analysis. The sample restriction is created by using the time-stable continuous distance measure that we calculate from each mine location to each DHS cluster. This is also the distance measure that we use to create the "mine" dummy, which captures whether the cluster lies within 20 km of a known gold deposit. Note that we only consider deposits that have been in production at some point until December 2012.

All households are thus within 100 km of one, or several, gold deposits. To ascertain whether there is any gold production in these potential mining sites, we construct an indicator variable *active*, which takes a value of 1 if there is at least one mine within 100 km that was extracting gold in the year the household was surveyed, and 0 otherwise. While the *mine* dummy captures some of the special characteristics of mining areas (for example, whether mines tend to open in less urban areas), the *active* dummy captures long-range spillovers of mining.

The treatment effect that we are mostly interested in is captured with the *active*mine* coefficient. The coefficient for $\beta_3$ tells us what the effect of being close to an actively producing mine is. Since the inclusion of the three dummies (*active*, *mine*, and *active*mine*) captures the difference between close and far, and before and after mine opening, we have created a difference-in-differences estimator.

Panel B of figure 2 shows this strategy in a map, where the small blue circles show the treatment areas, and the 100-km-radius green circles show the geographic areas that constitute the control group. As is common in difference-in-differences analysis, the estimation relies on treatment and control groups being on similar trajectories before mine opening. This assumption is discussed below when we investigate the balance of treatment and control areas. In particular, we test for differences in outcomes in areas where mining has not started and compare this to areas farther away.

While we cannot show the exogeneity of the opening year to local socioeconomic variables, this assumption has been made in previous literature (e.g. Aragon and Rud, 2015; Benshaul-Tolonen, 2018, 2019; Kotsadam and Tolonen, 2016, von der Goltz and Barnwal, 2019). In addition, Benshaul-Tolonen (2018, 2019) who explore gold mining, in particular, point to (i) the rapid increase in large-scale gold mining that occurred during the recent mineral price supercycle, (ii) the dominance of large multinational firms who are not relying on local labor market conditions, (iii) and their lower reliance on local infrastructure compared with bulkier metals and minerals, as gold mining firms may fly out their resources. Despite this, the assumption of exogenous opening year or exact location remain untested.

In a second method, we use a spatial lag model. Such a model allows for nonlinear effects with distance. We divide the plane into 10-km distance bins and estimate the model with a full set of distance bin dummies.

$$Y_{ivt} = \beta_0 + \sum_d \beta_d mine + \sum_d \beta_d active_t \cdot mine + \propto_d + g_t + \lambda X_i + \varepsilon_{ivt} \quad (2)$$

for d ∈ {0–10, 10–20, …, 80–90}.

This method, in addition to varying the cutoff point in the baseline estimation strategy, allows us to identify in more detail the spatial structure of the data. Using this method, we can support our choice of baseline cutoff distance.

Two limitations to both individual level analysis are that (i) clusters are not repeatedly sampled, so cluster fixed effects cannot be included, (ii) the data is not representative below the regional level and no weighting can be undertaken to ensure representability.

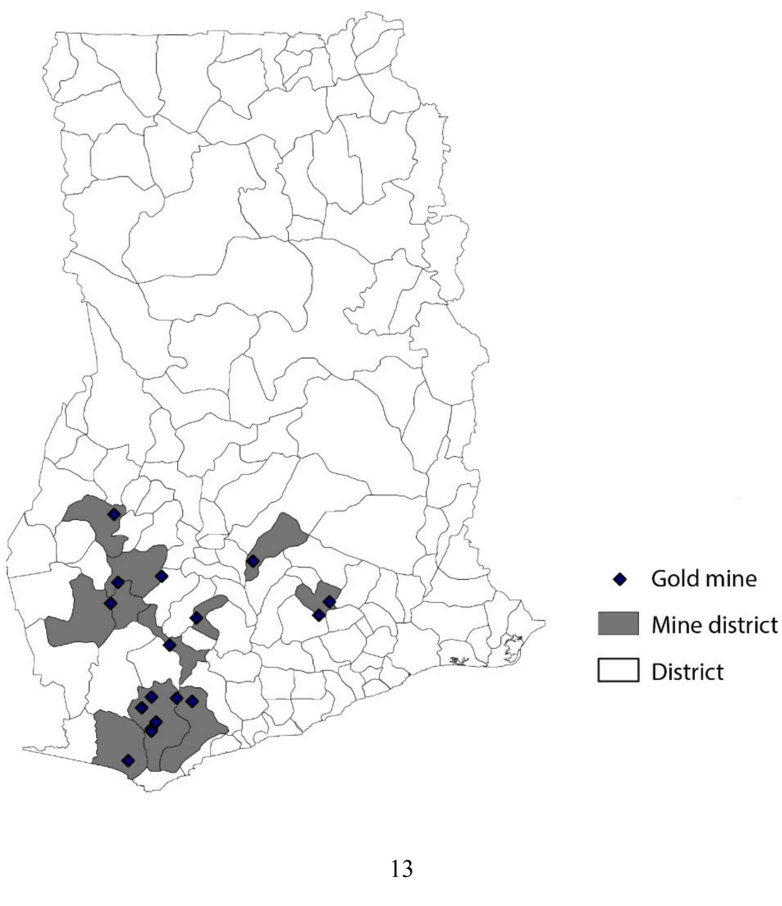
### 4.2 District-level analysis

While the estimation strategy in (1) captures some spillovers beyond the 20 km, and strategy (2) maps the economic footprint of the mine up to 100 km, it does not capture district-level treatment. District level is an additional interesting level of analysis, since it captures effects through the fiscal channel, and has previously been employed in, for example, Loyaza, Mier y Teran, and Rigolini (2013) to measure income inequality across mining districts in Peru. With Ghana's tax-sharing rules, district tax income and royalties increase with expansion in gold extraction (for more information, see section 2). In the district-level analysis, we first use mine openings as the independent variable, and then we use a richer specification with production levels. Given the spatial location of mines near district borders, we additionally analyze spatial spillovers from mining districts to neighboring districts.

### 4.2.1 Difference-in-differences at the district level

The first approach is similar to the local-level approach, only an individual is defined as being treated by a mine opening if she or he lives in a district with at least one active mine. In total, our mines are located in 11 mining districts (see figure 3). For districts with several mines, we define the whole district as active whenever at least one mine is active. Later, we will also consider district total annual production (tons of gold extracted), and thereby the effect of the intensity of production is recognized.

**Figure 3 District-level analysis in Ghana**



*Note:* This figure shows the mine locations and the district in which the mines are located.

The baseline specification is shown in the following equation:

$$Y_{idt} = \beta_1 active\_district_{dt} + \propto_d + g_t + \lambda X_{it} + \varepsilon_{idt} \qquad (3)$$

The outcome for individual *i* in district *d* in time period *t* is regressed on district and year fixed effects, an indicator for whether the individual lived in an active mine district at the time of the interview, and time varying individual-level factors. Even though the treatment is defined at the district level, we use individual-level data to be able to control for individual- level factors and explore heterogeneity at the individual level. The standard errors are, however, clustered at the district level to take into account the interdependence induced by the higher-level treatment. Since the treatment variable is at the same level as our district fixed effects, the $\beta 1$ coefficients are directly interpretable as difference-in-differences estimates. That is, they capture the difference between mining districts and nonmining districts before and after mining starts.

In estimating the district-level effects of mine openings on birth outcomes, we control for birth-year fixed effects instead of survey-year fixed effects, as we are interested in the effect of mining at birth. In investigating the effects on birth outcomes and infant mortality, we further classify a child as treated if he or she is born in a district with active mining during the birth year (in contrast to whether the mine is active when the mother was interviewed). We also include controls for the age of the child in the survey year in the child and birth outcome regressions (but, naturally, not in the infant mortality regressions).

## 5. Results using individual-level difference-in-differences strategy

In this section, we present results using the two difference-in-differences strategies. Since the individual analysis contains district fixed effects, the two strategies are complementary. While the district-level analysis informs us about differences across and within districts over time, the local-level analysis gives us the additional impact at the very local level. This means that any differences in effects across district and local analysis should not be interpreted as inconsistencies, but rather as differential and additional impacts.

In a difference-in-differences setting, it is important that the sample is balanced, assuming that the treatment and control groups are on similar trajectories. Table 2 shows the summary statistics for the women's surveys across four different groups, close and far away, and before and during the mine's production phase. Columns 1 and 3 show mean values of the population that live far away from mines, before and during mining respectively. Columns 2 and 4, in contrast, show the univariate regression coefficients using OLS, highlighting the difference between the population living close (e.g. Column 2) and far away (e.g. Column 1) before mining.

In the pre-period, women in communities that are close to mines are less urban, poorer, have more children and are less likely migrants. In contrast, women are of similar age, have similar education and occupation (but slightly more likely earning cash). Note that these are raw mean values not controlling for any regional and individual differences. Overall, these differences are in line with previous research finding that large-scale mines tend to open in more rural and less developed communities (Benshaul-Tolonen, 2018; Kotsadam and Tolonen, 2016).

In active mining communities, women are still less likely to live in urban areas (although the gap between mining and non-mining areas may be smaller) than in non-mining communities, but more likely to have some education. The difference-in-difference estimation strategy assumes similar trends over time across the treatment (close to mines) and control group (far away from mines), in absence of the gold mining expansion. While this assumption cannot be tested using our dataset, previous analyses have found evidence for parallel pre-trends in infant mortality and night lights (Benshaul-Tolonen, 2019) for gold mining countries in West and East Africa (including Ghana). The baseline differences in observable characteristics – in particular, lower levels of economic development preceding the mine opening - indicate that a cross-sectional approach using only the post-period may not be sufficient to understand the impact of gold mining on socio-economic variables.

**Table 2 Summary statistics for women's survey**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Before mining | | During Mining | |
| | >20 km | <20 km | >20 km | <20 km |
| | Mean | Coefficient | Mean | Coefficient |
| *Woman Characteristics* | | | | |
| Age | 28.79 | 0.836 | 28.95 | -0.352 |
| Total children | 2.18 | 0.417* | 2.56 | -0.035 |
| Wealth | 3.85 | -0.619** | 3.33 | -0.028 |
| Nonmigrant | 0.32 | 0.123** | 0.33 | -0.028 |
| Urban | 0.62 | -0.300** | 0.49 | -0.150** |
| No education | 0.17 | -0.045 | 0.20 | -0.042** |
| <3 years education | 0.77 | 0.035 | 0.74 | 0.045** |

| | | | | |
|---|---|---|---|---|
| *Woman occupation* | | | | |
| Earns cash | 0.90 | 0.059** | 0.89 | 0.007 |
| Works all year | 0.88 | -0.047 | 0.88 | 0.023 |
| Not working | 0.25 | -0.021 | 0.24 | -0.015 |
| Agriculture | 0.19 | 0.055 | 0.25 | 0.011 |
| Service & sales | 0.39 | 0.057 | 0.35 | 0.016 |
| Professional | 0.05 | -0.028 | 0.04 | -0.010 |
| Manual | 0.11 | -0.063*** | 0.12 | -0.003 |

*Note:* Column (1) is a sample at 20 to 100 km from a nonactive mine.
Column (2) difference for sample at 0-20 km from an nonactive, compared with column (1)
Column (3) is a sample within 20 to 100 km of an active mine.
Column (4) difference for sample at 0-20 km of an active mine, compared with Column (3)
*** p<0.01, **p<0.05, *p<0.1. Univariate regression model.

Appendix Table A2 also shows selected child health outcomes as summary statistics across the four treatment groups. We note that, once again, the sample looks quite balanced in the first three columns, although children seem to be worse off in communities close to mines that have not started producing, evidenced by the fact that infant mortality is 8 percent compared to 7 percent farther away, and 6 percent in communities with active mines. The anthropometrics height-for-age (stunting or chronic malnutrition), weight-for-age (wasting or acute malnutrition), and weight-for-height (underweight) show that the children living in mining communities before the mine started operating have the lowest scores of all four groups. The outcomes seem to improve with mining, although not enough to offset the initial adverse situation.

To test for exogeneity, we run regressions using baseline individual-level data to explore changes in observable characteristics among women (the main part of the sample). Table 3 shows that there are no significant effects of the mine opening on the age structure, migration history, marital status, fertility, or education, using the difference-in-difference specification with a full set of controls. If anything, it seems that women in active mining communities are marginally older, more likely to never have moved, and more likely to be or have been in a cohabiting relationship or married. Given the women's slightly higher age, it is not surprising to find that they have higher fertility and lower schooling (assuming that schooling has increased over time in Ghana). All these estimates are, however, insignificant.

**Table 3 Observable characteristics in the DHS individual data**

| | non- | ever | currently | ever | total | any schooling |
|---|---|---|---|---|---|---|

| | age | migrant | married | cohabiting | divorced | fertility | woman | partner |
|---|---|---|---|---|---|---|---|---|
| active*mine | 0.263 | 0.028 | 0.025 | 0.018 | -0.003 | 0.030 | -0.036 | -0.003 |
| | (0.510) | (0.042) | (0.027) | (0.029) | (0.017) | (0.115) | (0.031) | (0.030) |

*Note:* Robust standard errors clustered at the DHS cluster level in parentheses. All regressions control for year and district fixed effects, urban dummy, age (not column 1), and years of education (not columns 6 and 7). Active is active status of mine in the survey year. *** p<0.01, **p<0.05, *p<0.1.

### 5.1 Employment outcomes

Using the difference-in-differences approach (equation 1), we estimate results on occupation, child health, and inequality. First, panel A of table 4 indicates that women in active mining communities (*active*mine*) are insignificantly more likely to work in service and sales and less in agriculture, and 1.7 percentage points less likely to work as professionals (statistically significant). There is no change in the likelihood that she is not working. These 5 categories stem from the same occupational variable in the DHS data, and are mutually exclusive. The surveyed individual is told to report their main occupation. The coefficients can therefore be interpreted as relative increases of each specific sector. Women are more likely to earn cash for work, and the likelihood increases by 5.4 percentage points, which is equal to a 6 percent increase.

While the directionality of the occupational outcomes is broadly in line with previous results (Kotsadam and Tolonen, 2016, for 29 African countries, and Benshaul-Tolonen, 2018, for 8 African gold-producing countries), the estimates are largely insignificant, potentially due to a limited sample size. Two categories have positive, albeit insignificant, coefficients: services and manual labor. The (insignificant) estimate for service jobs[9] is equivalent to 6.7% increased employment, and manual labor 10.2%, alongside which the likelihood that a woman earns cash for her work increases with 6%.

For men (panel B of table 4), the estimates point toward an increase in agriculture, services, and professional (all statistically insignificant estimates), but a decreased likelihood of working in manual labor. Results for men in panel B are largely insignificant, and it is worthwhile noting that the sample size is only slightly above 50% than the women sample size due to DHS sampling frame.

**Table 4 OLS estimates women's and men's occupation in the DHS individual-level analysis**

Occupation

[9] Service sector jobs in the wake of structural transformation has been found important in increasing women's work hours and reduce the gender wage gap (Ngai and Petrongolo, 2017).

| | agri-culture | service sales | profess-ional | manual labor | not working | earns cash | works all year |
|---|---|---|---|---|---|---|---|
| PANEL A: Women | | | | | | | |
| active*mine | -0.025 | 0.024 | -0.017* | 0.012 | 0.006 | 0.054** | -0.013 |
| | (0.039) | (0.031) | (0.009) | (0.021) | (0.023) | (0.026) | (0.033) |
| mine | -0.025 | 0.056* | -0.001 | -0.012 | -0.018 | -0.069*** | -0.012 |
| | (0.031) | (0.029) | (0.008) | (0.018) | (0.020) | (0.022) | (0.024) |
| active | 0.014 | -0.000 | -0.006 | 0.009 | -0.016 | -0.037** | -0.007 |
| | (0.015) | (0.016) | (0.006) | (0.011) | (0.012) | (0.015) | (0.016) |
| Observations | 12,176 | 12,176 | 12,176 | 12,176 | 12,176 | 9,262 | 7,085 |
| R-squared | 0.350 | 0.103 | 0.124 | 0.024 | 0.234 | 0.095 | 0.042 |
| Mean of dep var. | 0.237 | 0.358 | 0.045 | 0.117 | 0.739 | 0.891 | 0.877 |
| PANEL B: Men | | | | | | | |
| active*mine | 0.050 | 0.020 | 0.027 | -0.069* | 0.006 | -0.013 | -0.015 |
| | (0.051) | (0.020) | (0.026) | (0.036) | (0.023) | (0.028) | (0.051) |
| mine | -0.060 | 0.002 | 0.000 | 0.041 | -0.018 | -0.009 | 0.066* |
| | (0.042) | (0.016) | (0.020) | (0.030) | (0.020) | (0.028) | (0.039) |
| active | 0.000 | 0.002 | -0.001 | -0.029 | -0.016 | -0.107*** | -0.025 |
| | (0.021) | (0.014) | (0.015) | (0.020) | (0.012) | (0.039) | (0.028) |
| Observations | 7,157 | 7,157 | 7,157 | 7,157 | 7,157 | 4,374 | 2,794 |
| R-squared | 0.290 | 0.415 | 0.084 | 0.183 | 0.076 | 0.107 | 0.104 |
| Mean of dep var. | 0.328 | 0.111 | 0.137 | 0.214 | 0.209 | 0.928 | 0.841 |

*Note:* Robust standard errors clustered at the DHS cluster level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. Active is active status of mine in the survey year. *** p<0.01, **p<0.05, *p<0.1. Results for women's partners also available upon request. OLS = ordinary least squares. The women and men samples have different sampling frames and therefore we are not analyzing these results jointly.

Two further caveats should be noted: (i) with many variables tested, such evidence should not be given too much weight, (ii) number of sampled men in the treatment category is quite small. Out of 7,157 sampled med, only 484 men live close to active mines. A bigger treated sample would provide more reliable results.[10]

### 5.2 Child health

We explore effects on child health, such as size at birth, infant mortality, anthropometrics, and incidence of cough, diarrhea, and fever. Panel A and Panel B use different variable definitions

[10] Results for sampled women's partners are similar.

because of the nature of the data. Panel A uses three variables that reflect conditions around the year of birth of the child. Note that women report their birth history, in what year the child was born, and the baby's health. She reports this information in the survey year, but retroactively, recollecting the year of birth. Therefore, we deem that the birth year mining activity is more relevant than the survey year mining activity for the outcomes in Panel A. In particular, we test if an active mine nearby in the birth year of the child influences the baby's size at birth, infant mortality, and the number of antenatal visits. In Panel B we use the main specification of mining activity in the survey year, as these variables are more reflective of current conditions in the survey year than past conditions around the time of the birth.

Panel A of table 5 shows that infants in active mine communities are less likely to be born large, and that the mother had insignificantly fewer prenatal visits. However, infant mortality decreases by 4 percentage points. Splitting the sample by gender, we note that this decrease is only statistically significant for boys at an effect size of 6.6 percentage points.

**Table 5 OLS estimates of birth outcomes, infant survival, and child health in the DHS individual-level analysis**

| PANEL A | size at birth | | | infant mortality (<12months) | | | antenatal visits | |
|---|---|---|---|---|---|---|---|---|
| | small | average | large | all | boys | girls | # visits | at least 1 |
| active*mine | 0.022 | 0.053 | -0.075* | -0.041* | -0.066** | -0.020 | -0.151 | -0.007 |
| | (0.028) | (0.041) | (0.041) | (0.022) | (0.030) | (0.035) | (0.331) | (0.028) |
| mine | -0.010 | 0.071** | -0.061** | 0.004 | 0.008 | 0.001 | 0.153 | 0.000 |
| | (0.019) | (0.028) | (0.030) | (0.015) | (0.020) | (0.024) | (0.241) | (0.019) |
| active | -0.010 | 0.054** | -0.044 | 0.002 | 0.014 | -0.012 | 0.012 | 0.002 |
| | (0.016) | (0.026) | (0.027) | (0.014) | (0.022) | (0.018) | (0.209) | (0.012) |
| Observations | 6,771 | 6,771 | 6,771 | 5,356 | 2,718 | 2,638 | 5,704 | 5,704 |
| R-squared | 0.031 | 0.054 | 0.059 | 0.135 | 0.160 | 0.152 | 0.186 | 0.062 |
| Mean of dep var. | 0.136 | 0.359 | 0.505 | 0.073 | 0.08 | 0.066 | 5.79 | 0.941 |

| PANEL B | in the last 2 weeks, had: | | | anthropometrics (WHO) in sd | | | has |
|---|---|---|---|---|---|---|---|
| | fever | cough | diarrhea | ht/a | wt/a | wt/ht | health card |
| active*mine | -0.035 | -0.061* | 0.042 | -3.532 | -5.208 | -0.641 | 0.014 |
| | (0.037) | (0.033) | (0.027) | (11.472) | (9.283) | (8.948) | (0.027) |
| mine | -0.002 | -0.006 | -0.038 | -0.828 | 3.481 | 3.853 | -0.006 |
| | (0.031) | (0.028) | (0.024) | (10.385) | (8.574) | (7.468) | (0.022) |
| active | 0.023 | -0.003 | -0.033** | -1.904 | 5.265 | 9.433* | 0.009 |
| | (0.020) | (0.020) | (0.016) | (5.942) | (5.304) | (5.183) | (0.012) |
| Observations | 6,246 | 6,257 | 6,262 | 5,627 | 5,627 | 5,727 | 6,378 |
| R-squared | 0.024 | 0.043 | 0.024 | 0.136 | 0.080 | 0.036 | 0.084 |
| Mean of dep var. | 0.211 | 0.221 | 0.164 | -101.6 | -60.3 | -16.7 | 0.913 |

*Note:* In panel A, active is status of mine in birth year; in panel B, active is active status of mine in survey year. Robust standard errors clustered at the DHS cluster level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. *** p<0.01, **p<0.05, *p<0.1. OLS = ordinary least squares.

An active mine is associated with a decrease in cough among children under age five (panel B of table 5), and children are also (insignificantly) more likely to have a health card. An active mine is associated with insignificant decreases in the anthropometrics measures (World Health Organization measures in standard deviations), such as height-for-age and weight-for-age. However, the standard errors for these coefficients are very large relative to the estimated coefficients, which is why the effects are imprecisely estimated.

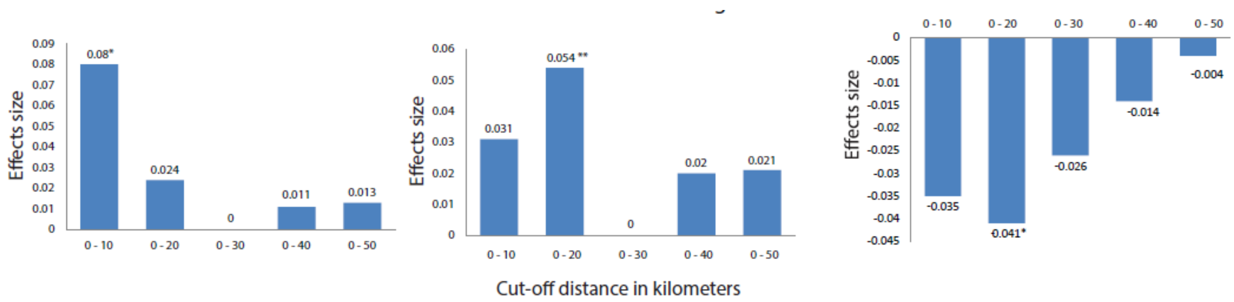
### 5.3 Spatial heterogeneity of results

Thus far we have used a cutoff distance of 20 km. Panel A of figure 4 shows that the largest treatment effect for services for women is found within 10 km of a mine, with an 8 percentage point increase in the probability that a woman works in the service sector. This is equivalent to a 22 percent increase in service sector participation. However, this effect is only statistically significant at the 10 percent level, possibly due to the small sample size within that distance. This is in contrast to the dummy for 0-20 km which is insignificant, pointing highly localized effects on service sector employment for women. Using distance bin of 30 km, we estimate zero treatment effect on the probability of service sector employment. Panel B of figure 4 shows the results for cash-earning opportunities, and similarly, we estimate positive treatment effects within 20 km.

In panel C of Figure 4, infant mortality is re-estimated using different distance spans. The decrease in child mortality decreases almost linearly with the increase in distance bins, an indication that the effects are only found close to a mine. The largest drop, and the only significantly negative drop, is found for the distance bin 0–20 km.

**Figure 4 Varying the cutoff distance: Service sector employment, cash earnings, and infant mortality**

**Panel A Service and sales** **Panel B Cash earnings** **Panel C Infant mortality**




*Note:* Figure 4 shows the main treatment coefficients using the baseline estimation strategy (with DHS individual-level data; see table 4 for more information), but with different distance cutoffs (10 km, 20 km, 30 km, 40 km, and 50 km). *** p<0.01, **p<0.05, *p<0.1.

The results above suggest substantial heterogeneity in outcomes and highlights the importance to explore spatial heterogeneity in the results. In the following sections we explore plausible explanations for these outcomes.

### 5.4 Difference-in-differences at the district level

The results for female employment in the district-level analysis are shown in table 6. Agricultural work decreases for women in mining districts and manual work increases. Following from this, the likelihood that a woman is working year-round increases.[11] This is similar to what we saw in the individual-level regressions, but the results are now statistically significant.

**Table 6 Effects of mine opening at the district level on female employment**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | not working | agri-culture | service or sales | profess-ional | manual work | earns cash | works all year |
| Active district | 0.019 | -0.085** | 0.034 | -0.018** | 0.050** | -0.021 | 0.054* |
| | (0.027) | (0.042) | (0.030) | (0.008) | (0.020) | (0.049) | (0.032) |
| Observations | 19,226 | 19,226 | 19,226 | 19,226 | 19,226 | 19,270 | 15,991 |
| R-squared | 0.207 | 0.327 | 0.128 | 0.137 | 0.037 | 0.213 | 0.278 |

*Note:* Robust standard errors clustered at the district level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. Active is active status of mine in the survey year. *** p<0.01, **p<0.05, *p<0.1.

[11] Working year-round is derived from a question if the woman works occasionally, all year or seasonally. Agricultural work also decreases for the partners of the women (results are available upon request).

Investigating the district-level effects on children's health and birth outcomes in table 7, we note a higher number of prenatal visits and an increase in attendance of a midwife in panel A. These results are highly statistically significant and the effects are economically significant. A mine opening increases the number of prenatal visits by 0.76 and increases the probability that the birth was preceded by a prenatal visit supervised by a midwife by 12.5 percentage points. In column 6 of panel A, we see that mine openings in a district reduce child mortality. The probability of an infant dying before 12 months of age is reduced by 8.5 percentage points. Given the importance of child mortality for human welfare, we strongly encourage future research to investigate the mechanisms behind these striking results. Since the share of prenatal visits supervised by a midwife also increases with mine openings, the results potentially speak to the importance of midwives for reducing infant mortality.

**Table 7 Effects of mine opening at the district level on birth outcomes and child health**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| PANEL A. | antenatal # visits | at least 1 antenatal | doctor attended | midwife attended | has health card | died 12 months |
| active district | 0.759*** | 0.026 | 0.055 | 0.125*** | 0.039 | -0.085*** |
| | (0.244) | (0.022) | (0.115) | (0.033) | (0.059) | (0.031) |
| N | 9,245 | 9,245 | 9,462 | 9,462 | 11,047 | 9,270 |
| R-square | 0.242 | 0.121 | 0.160 | 0.154 | 0.161 | 0.138 |
| | | at birth, the child was | | | | |
| PANELB. | small size | average size | large size | height for age | weight for age | weight for height |
| active district | 0.066 | 0.078 | -0.148 | -6.333 | -23.676** | -20.080 |
| | (0.057) | (0.085) | (0.090) | (18.753) | (9.364) | (13.428) |
| N | 11,837 | 11,007 | 11,007 | 9,646 | 9,646 | 9,851 |
| R-square | 0.041 | 0.061 | 0.060 | 0.199 | 0.163 | 0.073 |
| | | in the last 2 weeks, had | | | | |
| PANEL C. | fever | cough | diarrhea | | | |
| active district | 0.016 | 0.010 | 0.058 | | | |
| | (0.057) | (0.035) | (0.036) | | | |
| N | 10,849 | 10,883 | 10,887 | | | |
| R-square | 0.052 | 0.046 | 0.055 | | | |

*Note:* Robust standard errors clustered at the district level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. Active is active status of mine in the survey year. Panel b, columns 1, 2, and 3 show size at birth. Panel B, columns 4, 5, and 6 show anthropometrics (new WHO) in standard deviations. *** p<0.01, **p<0.05, *p<0.1.

The effects on child health are, however, not all positive. We note, particularly, a statistically significant decrease in weight-for-age, but the other two measures are also negative (panel B of table 7). Low weight-for-age is an indicator for acute malnutrition, whereas height-for-age is an indicator for chronic malnutrition. This could indicate that mining districts are less food secure.[12] Table 7 shows that there are no effects on illness in the last two weeks.

## 6. Distributional effects, mechanisms and robustness

### 6.1 Decomposing results by migration status

We argue that one source of heterogeneity is to consider when exploring socio-economic impacts and distributional effects of large-scale mining is migration status. First because mining may cause inward migration of individuals that are different from the previous local population. While it has its limitations, disaggregating the effects between nonmigrants and migrants may shed some light on the effect on the initial population. Second, to understand the distributional effects of mining we argue that migration status may be an important factor.

In the analysis, we distinguish between nonmigrants (where the woman respondent report being born in the locality) and migrants (born elsewhere). We note several caveats with this analysis, the first being that we cannot follow migrant households before the migration decision. Therefore, we cannot make any causal claims on changes in this group over time. We compare migrant households in mining communities with migrant households elsewhere, and the null hypothesis would be similar trajectory over time. If we reject the null, we cannot distinguish between selective migration to mining communities and the impact of the mining. The nonmigrant analysis can plausibly reflect similar households over time, with the limitation of selective outward migration. We believe inward migration to mining areas to be more common than outward migration (in line with Fafchamps et al., 2016).

Diarrhea is a major concern in many developing countries. Diarrheal diseases are, in part, a matter of infrastructure, where access to clean water and proper sanitation are important determinants. To further understand the effects on diarrhea, we look at the difference between migrants and nonmigrants and the effects by distance (Figure 5). There are, in fact, large differences between the migrant and the nonmigrant populations. Among nonmigrants, a mine opening is associated with large decreases in incidence, whereas for migrants, the opposite is true. Considering all children between 0 km and 20 km of an active mine, children born to

[12] In table 5 we saw very small insignificant changes in nutritional status.

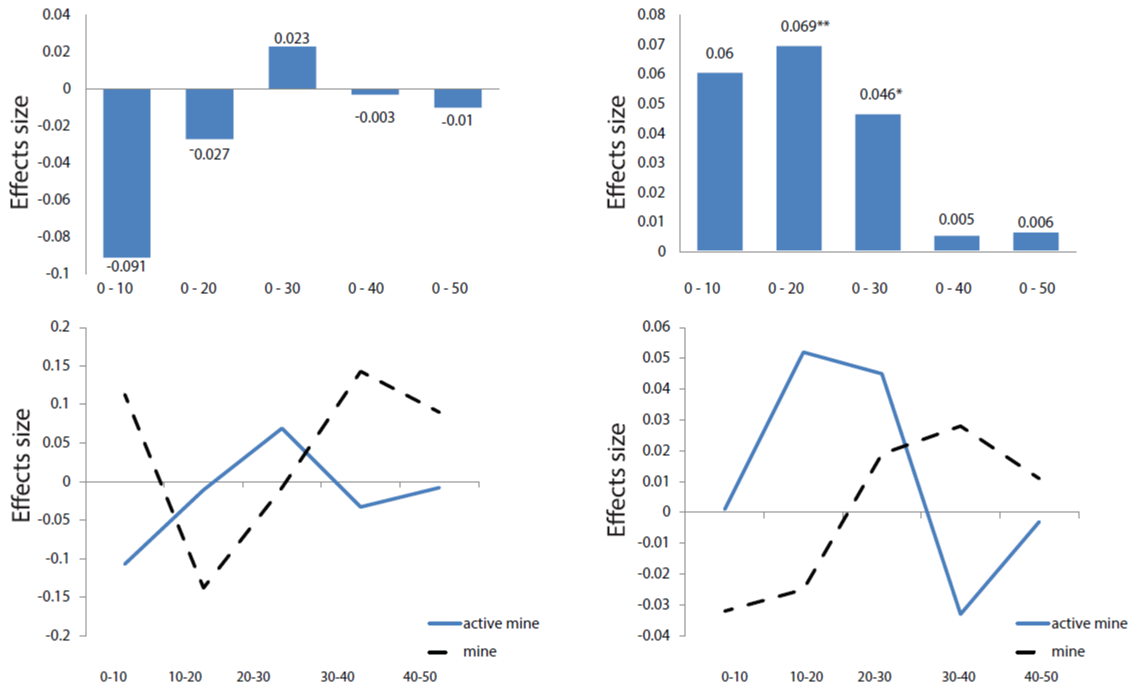
migrant mothers are 6.9 percentage points more likely to have suffered from diarrheal diseases in the two weeks prior to the start of the survey.

To further understand these effects, we decompose them by distance bins in a spatial lag model (bottom two graphs in figure 5). It becomes evident that, from a high-level incidence (dashed line) among the original local population (panel A of figure 5), the mine has brought substantial reductions in diarrheal incidence (as shown by the blue line). In the migrant population, the incidence is actually higher after mine opening than before, and the likelihood increases by 6.9 percentage points. The spatial lag model in panel B of figure 5 reveals that much of the effect is driven by a spike in incidence 10 to 20 km away from the mine center point. If more migrants move to the area because of the mine, they will be less settled, and health outcomes can deteriorate, on average, within that population. Nevertheless, we should be careful in interpreting the effects this way. The mine-induced migration, which we partly capture here, could be different from the migration happening further away. The deteriorating status of migrants can thus in part be because a less-well-off part of the population chooses to migrate to mining areas, not that they are made worse off because of the mine activities.

We also explored a decomposition of the anthropometric results along the migration division, but we found no important differences.

**Figure 5 Diarrheal incidence among children under 5 by migration status**

**Panel A Nonmigrants** **Panel B Migrants**




*Note:* Figure 5 shows the main treatment coefficients (*active*mine*) using the baseline estimation strategy (with DHS individual-level data; see table 4 for more information) in the top panel, but with different cutoffs (10 km, 20 km, 30 km, 40 km, and 50 km). *** p<0.01, **p<0.05, *p<0.1. The bottom panel shows the result using a spatial lag model that divided the plane into different treatment bins (0–10, 10–20, 20–30, 40–50) and compares them with farther away. Panel A shows the result for nonmigrants, and panel B shows the result for migrants.

### 6.2 Access to infrastructure and health care

Another source of heterogeneity is asset ownership and access to infrastructure. Table 8 shows that fewer households have electricity in active mining communities, but they spend less time fetching water and are more likely to own a radio (all estimates are statistically insignificant, however). There is no change in the likelihood of having a flush toilet. Moreover, it seems that households are just as likely to have access to a pit toilet as not having a toilet (and instead use a bucket, bush, and so forth).

**Table 8 OLS estimates for ownership of assets and access to infrastructure**

| | water access | | household has | | | | |
|---|---|---|---|---|---|---|---|
| | in minutes | less 10 min away | electricity | radio | flush toilet | pit toilet | no toilet |
| active*mine | -1.485 | 0.039 | -0.095* | 0.054 | 0.005 | -0.015 | 0.010 |
| | (1.933) | (0.048) | (0.056) | (0.036) | (0.023) | (0.033) | (0.027) |
| mine | -0.134 | -0.013 | 0.099* | 0.005 | 0.010 | -0.012 | 0.002 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1.805) | (0.039) | (0.054) | (0.029) | (0.021) | (0.029) | (0.021) |
| Active | 0.007 | 0.001 | 0.050** | 0.034** | -0.032 | 0.054** | -0.023 |
| | (1.012) | (0.026) | (0.024) | (0.017) | (0.021) | (0.026) | (0.022) |
| | | | | | | | |
| Observations | 9,790 | 9,790 | 12,226 | 12,216 | 12,227 | 12,227 | 12,227 |
| R-squared | 0.128 | 0.180 | 0.453 | 0.148 | 0.208 | 0.171 | 0.095 |
| mean of dep var | 0.407 | 14.84 | 0.565 | 0.652 | 0.151 | 0.732 | 11.6 |

*Note:* Robust standard errors clustered at the DHS cluster level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. Active is active status of mine in the survey year. *** p<0.01, **p<0.05, *p<0.1. OLS = ordinary least squares.

Overall, we estimated in table 8 that a household's access to electricity decreased when a mine became active. This finding is surprising if we assume that electricity infrastructure is seldom destroyed, so that with time, access has generally been increasing.[13]

Figure A1 (Annex) further decomposes the effect, and panel Aa of the figure confirms that the coefficient for electricity access is negative using the treatment distance 20 km. However, with a treatment distance of 10 km, the effect is marginally positive and insignificant. If we use a treatment distance of 50 km, we no longer see a significant effect. In panel Ab, the results are replicated using a spatial lag model, meaning that we allow for nonlinear effects with distance. In reality, it seems like the electricity rate is much higher before a mine (dashed line) than with an active mine (the blue line). However, when the results are decomposed by migrant status in panel Ac of figure A1 (Annex) we find that migrants are driving the lower electricity rate. In fact, among nonmigrants, the electricity rate is higher 0–10 km from an active mine, although it is slightly lower 10–20 km away.[14]

### 6.3 Distributional effects on wealth and inequality

Table 9 presents the effects of mining on asset wealth and on asset wealth inequality. Wealth data are available in the form of a wealth index, but only for the two last DHS surveys. Following Fenske (2015) and Flatø and Kotsadam (2014), we calculate inequality by means of a Gini coefficient (recoding the wealth variable to be positive only, and using the command

[13] It is also possible that mining companies compete with households for electricity if supply cannot be increased in the short run.

[14] In panels Ba, Bb, and Bc of Appendix figure A1, we analyze access to radio. We learn that access to radio is higher close to active mines, and that this seems true according to both the first method (Ba), according to the spatial lag model (Bb), and for both migrants and nonmigrants (Bc). The difference in effects between electricity and radio access might be due to electricity being more dependent on public infrastructure, and that electricity access may come with a time lag to other development indicators such as employment and access to radio, since a battery radio can be bought and used instantly, and easily moved.

fastgini[15] in STATA). We do this for both the cluster and district level. None of the effects of mining are statistically significant, but they point to increased asset wealth.

**Table 9 OLS estimates for wealth and inequality in the DHS individual-level analysis**

| | | Wealth Gini | |
|---|---|---|---|
| | wealth index | cluster level | district level |
| active*mine | 7,290 | -0.004 | 0.004 |
| | (12,849) | (0.013) | (0.018) |
| mine | 9,922 | 0.011 | 0.006 |
| | (8,676) | (0.013) | (0.016) |
| active | 7,854 | -0.006 | 0.034** |
| | (9,016) | (0.010) | (0.017) |
| Observations | 4,909 | 4,909 | 4,909 |
| R-squared | 0.613 | 0.227 | 0.548 |

*Note:* Robust standard errors clustered at the DHS cluster level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. Active is active status of mine in the survey year. *** p<0.01, **p<0.05, *p<0.1. OLS = ordinary least squares.

### 6.4 Bottom 40% of the population

To understand the welfare effects of the bottom 40 percent of the population in the income scale, we split the sample according to the wealth score provided by DHS. Given the data structure, which is repeated cross-section, we cannot follow a particular household that was identified as belonging to the bottom 40 percent in the initial time period. Instead, we identify the bottom 40 percent in four groups: far away, before mine or during mine, and close to mine, before mine or during mine. The summary statistics for selected main outcomes are presented in table 10. As the table shows, the bottom 40 percent in mining communities are more likely engaging in agriculture than the bottom 40 percent elsewhere. This could illustrate that agricultural workers are overrepresented among the less well-off in mining communities. However, women in this group still more often work in services than women did before in the same communities.

[15] Fastgini is a user-written command in STATA that helps calculate the gini coefficient.

**Table 10 Summary statistics for bottom 40% of women**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | far from a mine | | close to a mine | |
| | before | during | before | during |
| not working | 0.188 | 0.183 | 0.200 | 0.162 |
| service & sales | 0.340 | 0.203 | 0.179 | 0.222 |
| professional | 0.018 | 0.005 | 0.014 | 0.006 |
| agriculture | 0.362 | 0.530 | 0.490 | 0.539 |
| manual labor | 0.091 | 0.079 | 0.117 | 0.072 |
| earning cash | 0.876 | 0.855 | 0.879 | 0.901 |
| work all year | 0.852 | 0.859 | 0.879 | 0.838 |

*Note:* Column (1) is bottom 40% of sample at 20 to 100 km from a nonactive mine.
Column (2) is bottom 40% of sample at 20 to 100 km from an active mine.
Column (3) is bottom 40% of sample within 20 km of a nonactive mine.
Column (4) is bottom 40% of sample within 20 km of an active mine.

Regression results comparing these four groups are presented in panel B (urban) of Annex table 3. The results suggest that women in the bottom 40 percent are more likely agricultural workers in mining communities than elsewhere, but also more often service sector workers. They are less likely to work in manual labor, less likely to work all year, but more likely to earn cash for work. This indicates, possibly, that the economy becomes more reliant on cash as a mine starts producing. It is possible that the difference from the main results presented in table 4 indicates that agricultural workers are overrepresented among the bottom 40 percent in mining communities. However, given the issues associated with doing this analysis with repeated cross-section, we should be cautious in interpreting these results.

### 6.5 Heterogeneous results, sensitivity and intensity of mining

In panel B (urban) of table A3 (Annex), we interact our treatment variables (*active*mine*, *mine*, *active*) with an indicator variable for whether the locality is urban. This allows us to pick up potential differential effects across urban compared to rural localities. None of the treatment effects are statistically significantly different between rural and urban areas.[16] In panel C, we have constructed a new treatment variable *#active*mines* that counts the number of actively producing mines within 20 km. Women are sampled within 20 km of one mine (593 women), within 20 km of two mines (137 women), and within 20 km of three mines (64 women). The

[16] Few of the other interaction coefficients are also statistically significant. The interaction between *urban*mine* is significant, and women in urban localities with a future mine are 12 percentage points less likely to be working in agriculture.

mean value of the independent variable is 0.085 - that is, on average women are close to 0.085 mines. Conditional on being close to a mine, the main independent variable is 1.33 - that is, a woman sampled close to a mine is close to 1.33 mines on average. Panel C of table A3 shows the effects on women's labor market participation. We note that the estimates are similar in direction as before, where mines are positively associated with service and sales jobs and with cash earnings, but negatively associated with agriculture and professional jobs.

Panel D of table A3 shows the results if we drop the part of the sample that lives 20 km to 40 km away from a mine, and if we drop those that are sampled two years before mine opening. The rationale for this is to have a cleaner control group, since those that live just outside our 20 km cutoff distance may also be "treated" by the mine, and the investment phase of the mine that precedes initial production can generate substantial employment. Overall, the effects do not change much except making the cash earnings coefficient larger and more significant. The increase in cash earning opportunities is estimated at 7.5 to 7.8 percentage points compared with 5.4 in the baseline estimation.

### 6.6 Employment and wages using the GLSS

The DHS data do not provide detailed information regarding how much an individual earns for work, or her wage rate, but the GLSS does collect such data. First, we try to replicate the results estimated with the DHS data. Panel A of table 11 indicates that agriculture becomes less important in mining communities for women (statistically insignificant), who mainly shift into services and sales (statistically insignificant, except for strategy 2). Men are more likely to work as miners (statistically significant across all strategies).

**Table 11 Using GLSS: Employment on extensive and intensive margin and wages**

| | (1) worked last year | (2) work 7 days | (3) hours worked per week | (4) agri-culture | (5) service and sales | (6) miner |
|---|---|---|---|---|---|---|
| Panel A: Women | | | | | | |
| *1. baseline* | | | | | | |
| active*mine | -0.067* | -0.032 | 3.565 | -0.075 | 0.074 | 0.025 |
| | (0.040) | (0.038) | (3.140) | (0.064) | (0.054) | (0.016) |
| *2. drop 20-40 km* | | | | | | |
| active*mine | -0.062 | -0.039 | 3.849 | -0.076 | 0.094* | 0.026* |
| | (0.040) | (0.039) | (3.359) | (0.064) | (0.057) | (0.015) |
| *3. drop 2 years before* | | | | | | |
| active*mine | -0.067* | -0.031 | 3.565 | -0.087 | 0.080 | 0.024 |
| | (0.040) | (0.038) | (3.140) | (0.065) | (0.055) | (0.016) |
| *4. mine FE* | | | | | | |
| active*mine | -0.067 | -0.012 | 8.560* | -0.084 | 0.104 | 0.025* |
| | (0.051) | (0.048) | (5.125) | (0.075) | (0.065) | (0.015) |
| *5. mine clustering* | | | | | | |
| active*mine | -0.067* | -0.032 | 3.565 | -0.075 | 0.074 | 0.025 |
| | (0.032) | (0.036) | (3.521) | (0.081) | (0.080) | (0.022) |
| Mean dep var. | 0.727 | 0.673 | 40.39 | 42.32 | 0.391 | 0.005 |
| Panel B: Men | | | | | | |
| *1. baseline* | -0.086** | -0.055 | 3.705 | -0.058 | -0.032 | 0.125*** |
| active*mine | (0.041) | (0.039) | (3.460) | (0.066) | (0.036) | (0.043) |
| *2. drop 20-40 km* | | | | | | |
| active*mine | -0.094** | -0.062 | 3.893 | -0.064 | -0.031 | 0.126*** |
| | (0.042) | (0.040) | (3.842) | (0.066) | (0.038) | (0.042) |
| *3. drop 2 years before* | | | | | | |
| active*mine | -0.094** | -0.062 | 3.708 | -0.071 | -0.026 | 0.125*** |
| | (0.041) | (0.039) | (3.459) | (0.067) | (0.036) | (0.043) |
| *4. mine FE* | | | | | | |
| active*mine | -0.123** | -0.094* | 8.233 | -0.068 | -0.049 | 0.113** |
| | (0.057) | (0.051) | (5.425) | (0.075) | (0.044) | (0.045) |
| *5. mine clustering* | | | | | | |
| active*mine | -0.086*** | -0.055** | 3.705 | -0.058 | -0.032 | 0.125** |
| | (0.025) | (0.025) | (2.898) | (0.086) | (0.032) | (0.051) |
| Mean dep var | 0.715 | 0.705 | 45.71 | 0.491 | 0.259 | 0.028 |

*Note:* The table uses GLSS data for Ghana for the survey years 1998, 2005, 2012. The sample is restricted to women and men aged 15–49. Robust standard errors clustered at the village or neighborhood level in parentheses (except if otherwise stated). All regressions control for year and district fixed effects, urban dummy, age, and years of education. Active is active status of mine in the survey year. The treatment distance is defined to 20 km. Rows 2 drop sample between 20 to 40 km of a mine, and rows 3 drop sample that was surveyed two years before mine opening. *** p<0.01, **p<0.05, *p<0.1. FE = fixed effects.

To show the robustness of these results, we have tried three alternative strategies for each of the outcomes. In row 2, we drop the sample that lives 20 to 40 km away, since they might be affected by the mine; in row 3, we drop the sample that was surveyed two years prior to mine opening; in row 4, we add closest mine fixed effects; and in row 5 we cluster on the closest mine. The coefficients do not change much, even if some magnitudes become bigger and the estimates more significant. However, as in the results using DHS data, these estimates are not precisely measured – few are statistically significant because the standard errors appear large.

Women are 7.4 to 10.4 percentage points more likely to work in service or sales if they live close to a mine (depending on the estimations in panel A columns 4, only one statistically significant estimate). Women close to mines are 2.5 to 2.6 percentage points more likely to work in mining (only one statistically significant estimate).

Men, on the other hand, (results shown in panel B of table 12), are significantly more likely to work in mining, and insignificantly less in agriculture or service and sales. The likelihood that a man works in mining increases by 11.3 to 12.6 percentage points, which is more than a 400 percent increase in likelihood from the mean value which is 2.8%. For both men and women, the results are indicative of changes in labor force participation on the extensive and intensive margin. Fewer people work, as indicated by columns (1) and (2) (significant for men), but those who work, work more hours than before (column 3, albeit insignificant). It should be noted that the sample sizes are limited and these estimates may suffer from lack of power.

Annex figure 2 presents the results graphically and shows the spatial structure for a subset of the variables. The likelihood of a woman working in services decreases with distance from mine, and log wages are higher within 10 to 20 km of an active mine. Men are, intuitively, more likely to work as miners if they reside close to an active mine, and the correlation decreases with distance. Wages for men are also higher close to active mines (panel D). Beyond 40 km, the estimated effects are close to zero.

Table 12 shows that log annual wages are higher close to mines (column 1), and that most of the increase is driven by the increase in wage rates for women (column 2). Women, however, have lower wages before the mine, and a smaller share of women earn wages. Globally, it is considered that the historic expansion in service sector employment (which in this context increased significantly within 10 km) has played a pivotal role in reducing the gender wage and hour gap (Ngai and Petrongolo, 2017).

Despite the possible gains in wages for wage earners, we note a decrease in the regionally deflated total household expenditure (column 5), and a decrease in per capita expenditure on food and nonfood items (column 4). The increase in wages but decrease in total expenditure can possibly be explained by rising prices and wages in mining communities, where everyone has to pay the higher prices but only some (those who earn wages), benefit from a rise in wage rate.

Columns 6 through 9 of table 12 look at nondeflated expenditure measures for food, housing, health and education, and household energy.[17] We confirm that total household expenditure on food decreases (compared with the per capita deflated measure in column (4)), but find that households spend more money on housing, transport, and communication, and household energy, such as electricity and gas. The electricity and gas expenditure is only for those who have any positive expenditure on these, and we saw earlier that electricity access changes with the mine. This confirms that, among those who spend anything on electricity, they spend more on it in mining communities.

**Table 12 Using GLSS: Household income and expenditure**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| | ln | ln | ln | ln | household level ln expenditure | | | | |
| | wages | wages | wages | pc | total | | | health | hh |
| | all | women | men | exp. | exp. | food | housing | education | energy |
| active*mine | 0.520** | 0.694*** | 0.391 | -0.178* | -0.126 | -0.069 | 0.316** | -0.168 | 0.297** |
| | (0.226) | (0.241) | (0.238) | (0.093) | (0.089) | (0.095) | (0.139) | (0.199) | (0.119) |
| Observations | 6,226 | 2,914 | 3,312 | 7,522 | 7,522 | 7,396 | 7,420 | 6,541 | 4,752 |
| R-squared | 0.121 | 0.128 | 0.118 | 0.959 | 0.964 | 0.963 | 0.933 | 0.837 | 0.950 |
| *controls* | | | | | | | | | |
| individual | Y | Y | Y | | | | | | |
| hh head | | | | Y | Y | Y | Y | Y | Y |
| hh size | | | | | Y | Y | Y | Y | Y |
| district fe | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| year fe | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| deflated | N | N | N | Y | Y | N | N | N | N |
| mean (ln) | 15.30 | 15.29 | 15.31 | 13.04 | 14.19 | 13.42 | 10.88 | 10.74 | 9.52 |

*Note:* (1) Annual wages and salaries for individuals in all ages (nondeflated).
(2) Annual wages and salaries for women in all ages (nondeflated).
(3) Annual wages and salaries for men in all ages (nondeflated).
(4) Real per capita annual food and nonfood expenditure (regionally deflated).

[17] Additional results for recreation and transport and communication are available upon request. The expenditure on the three measures increased in mining communities.

(5) Total annual regionally adjusted household expenditure (local currency, regionally deflated).
(6) Total food expenditure (nondeflated).
(7) Total housing expenditure (nondeflated).
(8) Total health and education expenditure (nondeflated).
(9) Total household energy expenditure (gas and electricity) (nondeflated).
*** p<0.01, **p<0.05, *p<0.1. All regressions control for year and district fixed effects, urban dummy, age, and years of education.

## 7. Robustness for district-level estimations

### 7.2 Using production levels

We continue by exploring the effects of mining intensity as proxied by district-level production volumes. The estimation will be similar to equation (2), but we replace the indicator variable for being an active mining district with annual gold production in the district:

$$Y_{idt} = \beta_1 gold\_production_{dt} + \propto_d + g_t + \lambda X_{it} + \varepsilon_{idt} \qquad (4)$$

The measure of gold production is in 10 tons of gold produced, and $gold\_production_{dt}$ is either *gold_year_district*, which equals the total production of all mines in a district in the different survey years, or *gold_period_district*, which equals total production for the years before the survey. For the 1993 survey, the period is 1990–93, for 1998 it is 1994–98, and so on.

Using production levels instead of an indicator of having any production in the district has the advantage of capturing the intensity of mining production. Since it is somewhat unclear when mining production spills over to other types of employment, we use two measures of mining production. Panel A of table 13 shows the results of mining production in the period before the survey, including the survey year, on female employment, and we see that mining production leads to less agricultural employment but more employment in services and sales, as well as in professional work. Panel B shows that the effects are larger but not as precisely estimated for the yearly measure. That they are larger is not surprising, since a 10-ton increase one year is much more than a 10-ton increase over a longer time period. The precision is also probably lower since it is unclear what year the production spills over to other activities. In any case, we see that the effects are similar across these two specifications.[18]

[18] The effects for infant health and infant mortality are also stronger when we add production levels (results available upon request).

**Table 13 Effects of gold production at the district level on employment**

| Panel A: Using production in the previous period | | | | | | | |
|---|---|---|---|---|---|---|---|
| | (3) not working | (4) agri-culture | (5) service or sales | (6) profess-ional | (7) manual work | (1) earns cash | (2) works all year |
| gold period district | 0.003 | -0.009** | 0.003* | 0.004*** | -0.002 | -0.001 | 0.008** |
| | (0.004) | (0.004) | (0.002) | (0.002) | (0.004) | (0.002) | (0.003) |
| observations | 19,175 | 19,175 | 19,175 | 19,175 | 19,175 | 19,270 | 15,991 |
| R-squared | 0.207 | 0.327 | 0.127 | 0.137 | 0.037 | 0.213 | 0.278 |
| PANEL B. Using production in the same year | | | | | | | |
| | (3) not working | (4) agri-culture | (5) service or sales | (6) profess-ional | (7) manual work | (1) earns cash | (2) works all year |
| gold year district | 0.012 | -0.033 | 0.020 | 0.019* | -0.018 | -0.010 | 0.041*** |
| | (0.022) | (0.025) | (0.013) | (0.011) | (0.015) | (0.009) | (0.008) |
| observations | 19,175 | 19,175 | 19,175 | 19,175 | 19,175 | 19,270 | 15,991 |
| R-squared | 0.207 | 0.327 | 0.128 | 0.137 | 0.037 | 0.213 | 0.278 |

*Note:* Robust standard errors clustered at the district level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. *** p<0.01, **p<0.05, *p<0.1.

### 7.3 Investigating spillovers

The districts are small, and some mines are located in border regions. We thus expect there to be spillovers across district borders. For example, a mine can change demand for labor, agricultural produce, and services across the district border, and induce reallocation of work across districts. We explore neighbor spillovers by estimating the following equation:

$$Y_{idt} = \beta_1 gold\_prod_{dt} + \beta_2 neighbor\_gold\_prod_{dt} + \propto_d + g_t + \lambda X_{it} + \varepsilon_{idt} \qquad (5)$$

That is, we add gold production for the mining districts to their neighbors, and $\beta_2$ measures the effects of gold production in these districts, as well. That is, if $\beta_2$ is statistically and economically significant, it would imply that increased production in a neighboring district has spillover effects on the district in question.

In Annex table 4, we have added the gold production of the mining areas to their neighbors and we estimate the spillover effects of mining production in a district on employment in adjacent districts. As before, panel A shows effects of increasing production in the previous period, and panel B shows the effects of increased production in the survey year. There is no evidence of spillovers in the sense that there is a similar effect in neighboring districts. In fact, most coefficients point in the opposite direction for mining and neighboring districts indicating, if anything, a shift in employment from neighboring districts to the producing ones.

## 8 Conclusions

Ghana has a long history of gold production and has recently been experiencing its third gold rush, during which annual gold production skyrocketed. It was the first gold rush the country has experienced as an independent nation, and it brings hope of improving the lives of its citizens. Natural resource extraction is often argued to have detrimental effects on countries, however, and the so-called natural resource curse may imply that resource wealth is harmful to social development and inclusive growth. We use rich geocoded data with information on households and mining production over time to evaluate the gold boom at the local and district levels in difference-in-differences analyses.

Men benefit from direct job creation within the mining sector, and women seem to benefit from indirectly generated jobs in the service sector (statistically significant within 10 km from a mine). Women are more likely to earn cash and less likely to work in agriculture after mine openings. We find similar results when we analyze the effects at the district level and when we use production levels instead of openings and closings of mines. We interpret this as there being additional effects of being very close to a mine (within 20 km), beyond the effects from being in a mining district. No spillovers into neighboring districts are detected.

The results are in accordance with the results in Kotsadam and Tolonen (2016), who find similar effects on occupation in mining communities across the whole of Sub-Saharan Africa, and with Aragón and Rud (2013), who find that agricultural productivity in Ghana is reduced by mining production nearby. We find no statistically significant results on wealth and inequality, although the results point toward increases in both. The effects on infrastructure are ambiguous; we cannot detect any better access to flush toilets and radios, and the effects on electricity access are negative. Further decomposing these effects, we learn that migrant households are less likely to have access to electricity (compared with the change among migrant households living further away), whereas nonmigrant households that never moved might gain better access to electricity (compared with the change among nonmigrant households living further away).

Applying the same strategies to analyze child health and birth outcomes, we find both positive and negative effects of mining activity. Mining activity appears to marginally reduce the anthropometric status (short-term malnutrition) of children in mining districts, which could point to less food security. These results are in sharp contrast to the improvements in birth attendance and the decrease in infant mortality observed in mining communities and mining districts. A child in a mining district born after a mine has become active has had more prenatal visits and is less likely to die as an infant. This result is similar to what Benshaul-Tolonen (2019) finds for a larger sample of gold-producing countries in Africa. Despite substantial reductions in diarrheal diseases, the analysis highlights that migrant children are more likely to suffer from diarrheal diseases. The effects on the migrant community should be interpreted with care, however, since it may be that less-well-off people choose to migrate to mining communities and that the mine activities do not make them any less or better off. In addition, mine closure or downscaling could lead to deterioration in local employment conditions and health care access, as has been observed in Tanzania (Rhee et al., 2018).

The analysis shows that mining has created structural shifts in labor markets, and that it has reduced infant mortality rates. However, along with increased wage rates, we find that household level expenditure on housing and energy increases. In addition, the migrant population may have lower living standards with less electricity and a higher disease burden among children. We have no information where the migrant population moved from, and we cannot tell whether they have migrated to the area to benefit from the industry, or whether they were part of a relocation program due to the mining. One caveat is that these observed differences among migrant households in mining communities and non-mining communities could stem from untestable selection, as we do not observe the migrant households before the migration. Regardless of the motivation behind the migration decision, the policy recommendation is to ensure policies are in place to ensure sustainable living conditions in this group.

These district level findings should be placed in the context of seminal work by Caselli and Michaels (2010) who found weak increases in living standards in Brazilian municipalities after increase in off-shore oil revenue accruing to municipalities, alongside increased illegal activity by mayors (Caselli and Michaels, 2010). We estimate district level effects on living standards (in mining districts, but no spillovers to adjacent districts, in line with Mamo et al, 2019), but do not have further information on public spending by sector to put these effects in context to expected changes. Political outcomes such as clientelism, corruption and reelection of local politicians that have been linked to mining activities in other countries such as Peru (Maldonado, 2017) and India (Asher and Novosad, 2018), such as clientelism, corruption and reelection of local politicians were not analyzed within the context of Ghana. We encourage future analysis along similar lines.

A few caveats should be noted. As the gold mining industry in Ghana matures further, it will be important to determine the long sustainability of these economic effects. This paper does not tease out the effect of mine closure on local socio-economic conditions, an aspect that warrants future focus. Moreover, lack of clearly estimated effects both in the individual level and district level analysis could stem from limited sample sizes. We encourage future analysis to use more rounds of data to ensure consistent results. Lastly, for the health and employment effects that we observe, we cannot determine if they stem from changes in the market-based economy, from corporate social responsibility policies or public spending. Future studies should try to carefully disentangle the mechanisms at play.

**References**

Aragón, F. M., P. Chuhan-Pole, and B. C. Land. 2015. "The Local Economic Impacts of Resource Abundance: What Have We Learned?" Mimeo. World Bank, Washington, DC.

Aragón, F. M., and J. P. Rud. 2015. "Polluting Industries and Agricultural Productivity: Evidence from Mining in Ghana." The Economic Journal.

Aragón, F. M., and J. P. Rud. 2013. "Natural Resources and Local Communities: Evidence from a Peruvian Gold Mine." *American Economic Journal, Economic Policy* 5 (2): 1–25.

Aryee, B. N. 2001. "Ghana's mining sector: its contribution to the national economy." *Resources Policy* 27 (2): 61–75.

Asher, S., & Novosad, P. (2018). *Rent-seeking and criminal politicians: Evidence from mining booms*. Working Paper.

Axbard, S., Poulsen, J., & Benshaul-Tolonen, A. (2016). Extractive industries, production shocks and criminality: Evidence from a middle-income country. CDEP-CGEG WP, (30).

Benshaul-Tolonen, A. 2019. "Local industrial shocks and infant mortality." *The Economic Journal*.

———.2018. Endogenous Gender Norms: Evidence From Africa's Gold Mining Industry. SSRN 3284519.

Bermudez-Lugo, O. 2012. "The mineral industry of Ghana." In *Minerals Yearbook: Area Reports International Review of 2010 Africa and the Middle East*. United States Geological Survey (USGS), pp. 19–1.

———. 2013. "The mineral industry of Ghana." In *Minerals Yearbook: Area Reports International Review of 2011 Africa and the Middle East*. United States Geological Survey (USGS), pp. 19–1.

Bloch, R. and Owusu, G., 2012. *"Linkages in Ghana's gold mining industry: Challenging the enclave thesis."* Resources Policy, *3.*

Caselli, F., and G. Michaels. 2013. "Do Oil Windfalls Improve Living Standards? Evidence from Brazil." *American Economic Journal: Applied Economics* 5 (1): 208–238.

Corno, L., and D. de Walque. 2012. "Mines, Migration and HIV/AIDS in Southern Africa." *Journal of African Economies* 21 (3): 465–498.

Fafchamps, M., M. Koelle, and S. Forhad. 2016. "Gold Mining and Proto-Urbanization: Recent evidence from Ghana." Journal of Economic Geography: lbw015.

Fenske, J. 2015. "African polygamy: Past and present." Journal of Development Economics 117, 58-73.

Flatø, M., and A. Kotsadam. 2014. "Droughts and Gender Bias in Infant Mortality in Sub-Saharan Africa." University of Oslo (UiO) Working Paper No. 02/2014, Oslo.

Frankel, J. A. 2012. "The Natural Resource Curse: A Survey of Diagnoses and Some Prescriptions." HKS Faculty Research Working Paper Series RWP12-014, John F. Kennedy School of Government, Harvard University.

Gajigo, O., E. Mutambatsere, and G. Mdiaye. 2012. "Gold Mining in Africa: Maximizing Economic Returns for Countries." African Development Bank, Abidjan, Côte d'Ivoire.

Garvin, T., T. K. McGee, K. E. Smoyer-Tomic, and E. A. Aubynn. 2009. "*Community--company relations in gold mining in Ghana*". Journal of Environmental Management.

Hausermann, H., Ferring, D., Atosona, B., Mentz, G., Amankwah, R., Chang, A., Hartfield, K., Effah, E., Asuamah, G.Y., Mansell, C. and Sastri, N., 2018. *"Land-grabbing, land-use transformation and social differentiation: Deconstructing "small-scale" in Ghana's recent gold rush". World Development*, *108*, pp.103-114.

Hilson. G. 2009. "*Small-scale mining, poverty and economic development in sub-Saharan Africa: An overview*". Resources Policy.

Hilson. G. 2002. "*Harvesting mineral riches: 1000 years of gold mining in Ghana.*" Resources Policy, Elsevier, 2002, 28, 13-26

Hilson, G., and N. Yakoleva. 2007. "Strained relations: A critical analysis of the mining conflict in Prestea, Ghana." *Political Geography* 26 (1): 1–22.

ILO (International Labour Organization). 1999. Social and Labour Issues in Small-scale Mines: Report for Discussion at the Tripartite Meeting on Social and Labour Issues in Small-scale Mines. Geneva.

InterraRMG. 2013. http://www.rmg.se/Products/RawMaterialsData.aspx.

Kotsadam, A., and A. Tolonen. 2016. "African Mining, Gender and Local Employment." World Development 83, 325-339.

Loayza, N., A. Mier y Teran, and J. Rigolini. 2013. "Poverty, inequality, and the local natural resource curse." Discussion Paper Series, Forschungsinstitut zur Zujunft der Arbeit, 7226, Bonn.

Maldonado, S. (2017). The non-monotonic political effects of resource booms.

Mamo, N., Bhattacharyya, S., & Moradi, A. (2019). Intensive and extensive margins of mining and development: evidence from Sub-Saharan Africa. *Journal of Development Economics*, *139*, 28-49.

Michaels, G. 2011. "The Long Term Consequences of Resource-Based Specialisation." *The Economic Journal* 121: 31–57.

MineAtlas. 2013. https://mining-atlas.com/

Rhee, S., Nyanza, E. C., Condon, M., Fisher, J., Maduka, T., & Benshaul-Tolonen, A. (2018). Understanding environmental, health and economic activity interactions following transition of ownership in gold mining areas in Tanzania: A case of private to public. *Land use policy*, *79*, 650-658.

Standing, A., and G. Hilson. 2013. "Distributing mining wealth to communities in Ghana: Addressing problems of elite capture and political corruption." *U4* issue 2013 (5).

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

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

Wilson, N. 2012. "Economic Booms and Risky Sexual Behavior: Evidence from Zambian Copper Mining Cities." *Journal of Health Economics* 31 (6): 797–812.

World Bank. 2011. "Global Economic Prospects June 2011. Regional Annex." World Bank, Washington, DC.

## Annex

### Table A1. Variable definitions

| | |
|---|---|
| *Characteristics* | |
| total children | Total lifetime fertility |
| wealth | Household wealth index score |
| non-migrant | Respondent was born in the location and has never moved |
| migrant | Any respondent who has ever moved in their life |
| urban | The household lives in urban area |
| | |
| *Woman's occupation* | |
| not working | Was not working in last 12 months |
| service & sales | Works in services or sales |
| professional | Works as a professional |
| agriculture | Works in agriculture |
| manual labor | Works in manual labor |
| earning cash | Earns cash for work (0= not paid, in kind) |
| works all year | Works all year (0= seasonally, occasionally) |
| | |
| *Woman's education* | |
| 3 years education | At least 3 years of education |
| no education | No education |
| | |
| *Child health* | |
| first 12 months | Child died within 12 months from birth |
| diarrhea | Child had diarrhea in last 2 weeks |
| cough | Child had cough in last 2 weeks |
| fever | Child had fever in last 2 weeks |
| | |
| *Child anthropometrics* | |
| ht/age (st dev.) | Height for age (standard deviation) |
| wt/age (st dev.) | Weight for age (standard deviation) |
| wh/ht (st dev.) | Weight for height (standard deviation) |

**Table A2 Summary statistics for children's surveys**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | far from a mine | | close to a mine | |
| | before | during | before | during |
| *infant mortality* | | | | |
| first 12 months | 0.07 | 0.07 | 0.08 | 0.06 |
| *child health* | | | | |
| diarrhea | 0.17 | 0.17 | 0.13 | 0.17 |
| cough | 0.24 | 0.22 | 0.22 | 0.18 |
| fever | 0.20 | 0.21 | 0.24 | 0.20 |
| *child anthropometrics* | | | | |
| ht/age (st dev.) | -94.43 | -104.88 | -127.04 | -115.76 |
| wt/age (st dev.) | -90.80 | -100.16 | -114.28 | -103.48 |
| wh/ht (st dev.) | -40.29 | -45.24 | -47.816 | -40.52 |
| Sample size (child at birth) | 3709 | 2204 | 661 | 314 |

*Note:* Column (1) is a sample at 20 to 100 km from a nonactive mine.
Column (2) is a sample at 20 to 100 km from an active mine.
Column (3) is a sample within 20 km of a nonactive mine.
Column (4) is a sample within 20 km of an active mine.
Infant mortality considers mine active status in birth year.
ht/age = height-to-age; wt/age = weight-to-age; wh/ht = weight to height; st. dev. = standard deviation.

**Table A3. Heterogeneous effects for bottom 40%, with urban locality interactions, intensity of mining, and timing of opening**

| | Woman's occupation | | | | | | |
|---|---|---|---|---|---|---|---|
| | agri-culture | service sales | profess-ional | manual labor | not working | earns cash | works all year |
| PANEL A: Bottom 40% | | | | | | | |
| active*mine | 0.033 | 0.029 | -0.004 | -0.078** | 0.020 | 0.089*** | -0.083* |
| | (0.068) | (0.048) | (0.019) | (0.038) | (0.059) | (0.034) | (0.044) |
| Mine | -0.009 | 0.044 | 0.006 | 0.018 | -0.058 | -0.065** | 0.012 |
| | (0.057) | (0.040) | (0.010) | (0.038) | (0.046) | (0.032) | (0.036) |
| Active | 0.068 | -0.071* | -0.008 | 0.014 | -0.003 | -0.052 | -0.064* |
| | (0.041) | (0.038) | (0.007) | (0.025) | (0.031) | (0.048) | (0.036) |
| Observations | 2,536 | 2,536 | 2,536 | 2,536 | 2,082 | 2,083 | 2,536 |
| PANEL B: Urban | | | | | | | |
| active*mine | -0.037 | 0.022 | -0.013 | 0.019 | 0.009 | 0.062** | -0.014 |
| | (0.044) | (0.034) | (0.010) | (0.023) | (0.025) | (0.028) | (0.034) |
| Mine | -0.005 | 0.046 | -0.002 | -0.018 | -0.022 | -0.068*** | -0.011 |
| | (0.033) | (0.031) | (0.008) | (0.020) | (0.022) | (0.023) | (0.026) |
| Active | 0.007 | -0.005 | -0.008 | 0.012 | -0.007 | -0.046** | 0.004 |
| | (0.024) | (0.021) | (0.005) | (0.015) | (0.015) | (0.021) | (0.022) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| active*mine*urban | 0.074 | 0.004 | -0.022 | -0.038 | -0.018 | -0.041 | -0.002 |
| | (0.054) | (0.058) | (0.024) | (0.037) | (0.042) | (0.052) | (0.052) |
| active*urban | 0.011 | 0.009 | 0.002 | -0.006 | -0.015 | 0.015 | -0.019 |
| | (0.025) | (0.025) | (0.010) | (0.017) | (0.019) | (0.021) | (0.023) |
| mine*urban | -0.121*** | 0.052 | 0.006 | 0.040 | 0.024 | -0.004 | 0.001 |
| | (0.043) | (0.045) | (0.017) | (0.029) | (0.033) | (0.043) | (0.045) |
| Urban | -0.240*** | 0.130*** | 0.009 | 0.033** | 0.068*** | 0.013 | 0.026 |
| | (0.023) | (0.023) | (0.008) | (0.016) | (0.018) | (0.018) | (0.022) |
| Observations | 12,176 | 12,176 | 12,176 | 12,176 | 12,176 | 9,262 | 7,085 |
| PANEL C. Intensity #active*mines | -0.026 | 0.039* | -0.020*** | 0.005 | 0.001 | 0.038* | -0.010 |
| | (0.028) | (0.024) | (0.007) | (0.017) | (0.020) | (0.021) | (0.021) |
| Mines | -0.023 | 0.047 | 0.001 | -0.009 | -0.016 | -0.063*** | -0.012 |
| | (0.030) | (0.028) | (0.008) | (0.018) | (0.020) | (0.021) | (0.023) |
| Active | 0.014 | -0.001 | -0.006 | 0.009 | -0.016 | -0.037** | -0.007 |
| | (0.015) | (0.016) | (0.006) | (0.011) | (0.012) | (0.015) | (0.016) |
| Observations | 12,176 | 12,176 | 12,176 | 12,176 | 12,176 | 9,262 | 7,085 |
| PANEL D. Robustness | | | | | | | |
| *1. Drop 20-40 km* | | | | | | | |
| active*mine | -0.040 | 0.020 | -0.024** | 0.017 | 0.026 | 0.078*** | 0.023 |
| | (0.043) | (0.030) | (0.009) | (0.022) | (0.024) | (0.028) | (0.040) |
| *2. Drop 2 years before* | | | | | | | |
| active*mines | -0.013 | 0.025 | -0.018* | 0.002 | 0.003 | 0.075*** | -0.028 |
| | (0.040) | (0.030) | (0.009) | (0.021) | (0.024) | (0.028) | (0.037) |

*Note:* Robust standard errors clustered at the DHS cluster level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. *** p<0.01, **p<0.05, *p<0.1. Panel A is limited to bottom 40% in the income distribution, panel B uses urban interaction, and panel C has a count variable for active mines. Panel D1 drops sample between 20 and 40 km away, and D2 drops individual samples two years before mine opening. 151 women are sampled within 20 km from an active mine and in an urban area, and 246 women are sampled within 20 km from a mine regardless of its activity status and in an urban area.

**Table A4 Spillovers on employment across districts**

| Panel A: Using production in the previous period | | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| VARIABLES | not working | agri-culture | service or sales | profess-ional | manual work | earns cash | works all year |
| gold period District | 0.004<br>(0.004) | -0.009**<br>(0.004) | 0.003*<br>(0.002) | 0.004***<br>(0.002) | -0.002<br>(0.004) | 0.001<br>(0.003) | 0.006<br>(0.004) |
| neighbor gold production | -0.004<br>(0.004) | 0.005<br>(0.004) | -0.001<br>(0.004) | -0.002***<br>(0.001) | 0.001<br>(0.003) | 0.008*<br>(0.004) | -0.002<br>(0.004) |
| observations | 19,175 | 19,175 | 19,175 | 19,175 | 19,175 | 14,852 | 11,568 |
| R-squared | 0.207 | 0.327 | 0.127 | 0.137 | 0.037 | 0.146 | 0.255 |

Panel B: Using production in the same year

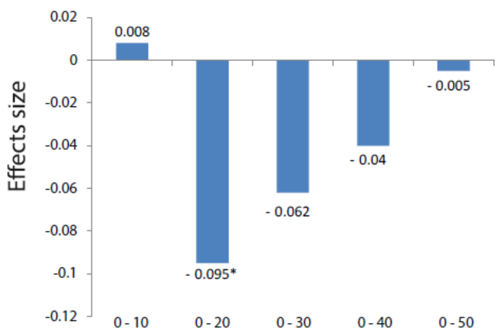
| VARIABLES | (1) not working | (2) agri-culture | (3) service or sales | (4) profess-ional | (5) manual work | (6) earns cash | (7) works all year |
|---|---|---|---|---|---|---|---|
| gold period | 0.012 | -0.033 | 0.020 | 0.019* | -0.018 | -0.001 | 0.028 |
| District | (0.022) | (0.024) | (0.013) | (0.011) | (0.015) | (0.015) | (0.022) |
| | | | | | | | |
| neighbor | -0.042** | 0.036 | 0.007 | -0.009** | 0.008 | 0.020 | 0.013 |
| gold production | (0.017) | (0.025) | (0.021) | (0.004) | (0.010) | (0.025) | (0.019) |
| | | | | | | | |
| observations | 19,175 | 19,175 | 19,175 | 19,175 | 19,175 | 14,852 | 11,568 |
| R-squared | 0.207 | 0.327 | 0.128 | 0.137 | 0.037 | 0.146 | 0.255 |

*Note:* Robust standard errors clustered at the district level in parentheses. All regressions control for year and district fixed effects, urban dummy, age, and years of education. *** p<0.01, **p<0.05, *p<0.1.

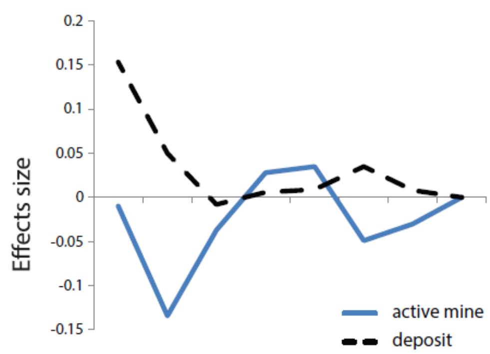
## Figure A1. Access to infrastructure: Varying the cutoff and spatial lag model

### Panel A Household has electricity

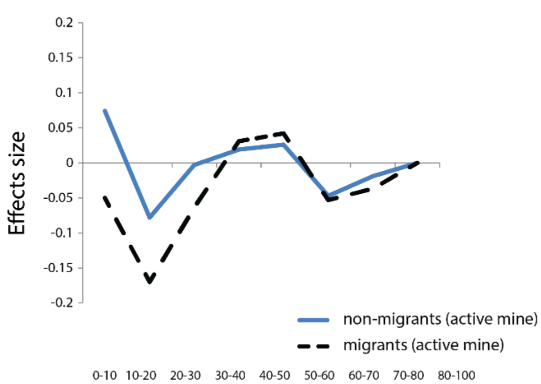
**a. Varying cutoff**




**b. Spatial lag model**

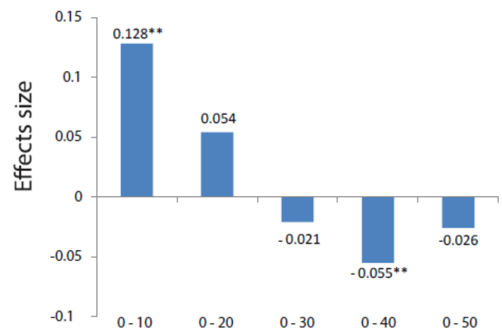



**c. Spatial lag model by migration status**

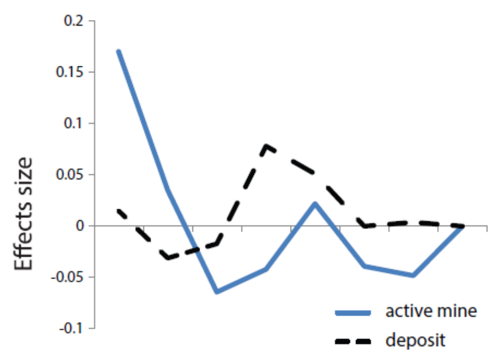



### Panel B Household has radio

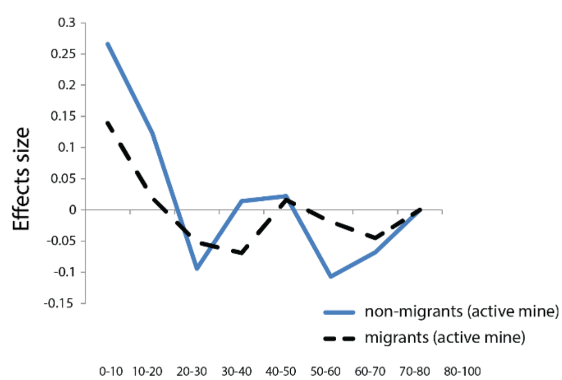
**a. Varying cutoff**




**b. Spatial lag model**



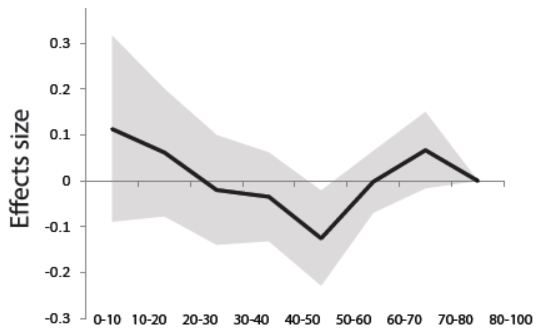

**c. Spatial lag model by migration status**



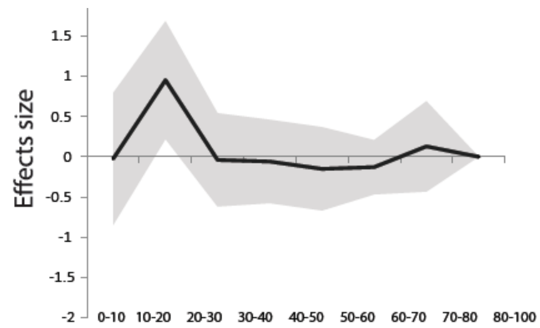

*Note:* The figure shows the main treatment coefficients (*active*mine*) using the baseline estimation strategy (with DHS individual-level data; see table 4 for more information) in panel A, but with different distance cutoffs (10 km, 20 km, 30 km, 40 km, or 50 km). *** p<0.01, **p<0.05, *p<0.1. Panels Ab, Ac, Bb, and Bc show the result using spatial lag models, which divided the plane into different treatment bins (0–10, 10–20, 20–30, 40–50) and compares them with farther away distances. Panel B shows the result for all individuals, and panel C shows the main treatment result (active mine) when the sample has been split into migrants and nonmigrants.

**Figure A2 Using GLSS: Employment and wages**

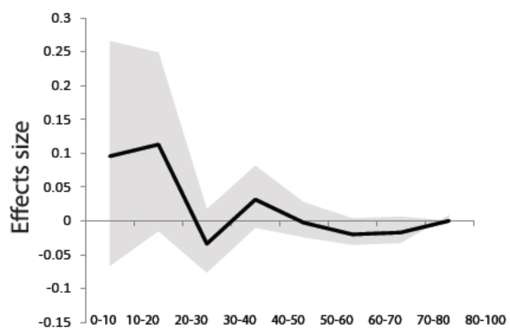
**Panel A Woman working in services**



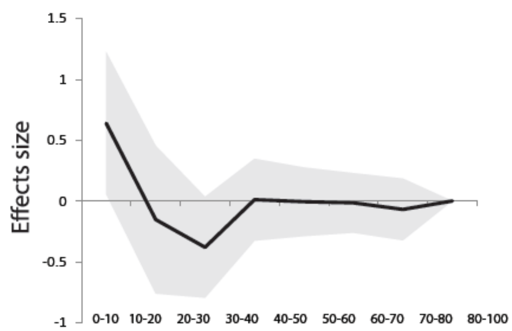

**Panel B Log wages (woman)**




**Panel C Man working in mining**




**Panel D Log wages (man)**




Distance in kilometers

*Note:* The four panels show the regression results from four spatial lag models using the GLSS sample. The sample is restricted to women (top two) and men (bottom two) aged 15–49. The solid lines are the coefficient for *active*mine* for 7 distance bins (0–10 km, 10–20 km, … 60–70 km) compared with a control group (80–100 km) away. The regressions also control for mine location at the same distances. See table 11 for control variables.