FAERE

French Association
of Environmental and Resource Economists

Working papers

Effect of gold mining on income distribution in Ghana

George Adu - Franklin Amuakwa-Mensah - George Marbuah - Justice Tei Mensah

WP 2016.23

Suggested citation:

G. Adu, F. Amuakwa-Mensah, G. Marbuah, J. Tei Mensah *(2016).* Effect of gold mining on income distribution in Ghana. *FAERE Working Paper, 2016.23.*

ISSN number: **2274-5556**

www.faere.fr

Effect of gold mining on income distribution in Ghana

George Adu[a,b]

[a]The Nordic Africa Institute, Uppsala University, Box 1703, SE-75147, Uppsala, Sweden
[b]Department of Economics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. Email: George.Adu@nai.uu.se / gadu.cass@knust.edu.gh

Franklin Amuakwa-Mensah[c], George Marbuah[c], & Justice Tei Mensah[c]
[c]Department of Economics, Swedish University of Agricultural Sciences, Box 1703, SE-75007, Uppsala, Sweden

Abstract

This paper examined the effect of mining on household income and welfare and how such effects are distributed over different quantiles of income and welfare. Using the three most recent rounds of the Ghana Living Standards Surveys together with information on the location of gold mines during the survey years, we estimate effects of living in a mining area on real gross income, employment income, and real per capita household expenditure (a proxy for welfare) using average and quantile treatment effect models. We find robust evidence of negative effect of mining on household income and welfare. Our results also indicate that the income reducing effect of mining activity falls heavily on households at bottom of the income distribution. In the case of household welfare, the interesting revelation from our results is that the negative effect of mining falls largely on both the lower and upper ends of the welfare distribution, with much heavier burden at the lower relative to the upper tail. Our paper, thus, provides ample evidence that mining activity does not only reduce income and welfare, but further increases inequality in the distribution of income and welfare.

Key words: Gold mining, Income and welfare distribution, Quantile treatment effect, Ghana

JEL Classification: C31, O13

1. Introduction

Ghana has a long tradition of gold mining. Over the last two decades or so, the gold mining industry has experienced a boom with large expansion in capital-intensive and industrial-scale production. While the significant contribution of the mining sector to public sector finances has received some acknowledgements, its welfare effects on local population are not well understood. The objective of this paper is to answer two important questions: First, What is the average effect of mining activity on the income and welfare of households living in mining areas? Secondly, how does this effect, if any, vary across different quantiles of income and welfare?

By this we are able to answer the question of whether mining activities have equalizing or non-equalizing effects on income and welfare of households living in mining areas in comparison to similar households living farther away from mining activity. As a way of improving understanding of the effect of mining on local populations, there is growing literature on the social and economic impact of large scale mining activities, applying state-of-the- art econometric techniques (see for instance: World Bank, 2015; Aragon and Rud, 2015, 2013; Kotsadam and Tolonen, 2016; Tolonen, 2015; Chuhan-Pole, et al 2015).

The previous literature on the local economic impact of (gold) mining have addressed the questions of whether mining activities have any impact on local populations and how the impact (if any) is distributed across different dimensions of the population such as gender, sector of employment and migration status. Previous studies, relying on quasi experimental approaches have reported the average effects of living close to a mine (see for instance: Aragon and Rud, 2013; Kotey and Rolfe, 2014; Kotsadam and Tolonen, 2016; Lippert, 2014; Tolonen, 2015, Chuhan-Pole, et al. 2015), but do not provide an answer to the question of how the effect varies across different income dimensions. Nonetheless, the findings from these studies suggest that the benefits and costs of opening of a large scale mine are not evenly distributed across the population. For instance, Kotsadam and Tolonen (2016) concluded that industrial mine openings constitute a mixed blessing for women in sub-Saharan Africa. The authors find that mine opening results in a transition from agriculture to service sector employment which is often female dominated. However, the job losses from the agriculture sector far outweigh the increase in service sector employment; hence the net effect on female employment is negative. The men on the other hand find direct employment in the mining sector (Kotsadam and Tolonen, 2016).

Aragon and Rud (2015) have also reported that pollution from large scale mine openings cause a decline in agricultural productivity and increases poverty in mining areas. Drawing from the findings reported by Aragon and Rud (2015) and Kotsadam and Tolonen (2016), we surmise that mining deepens income gaps between male and females; and between agricultural and non-agricultural workers.

While the studies by Aragon and Rud (2015) and Kotsadam and Tolonen (2016) suggest unequal distribution of the benefits and costs of mining between male and females on one hand, and farmers and non-farmers on the other hand, the distribution of the benefits and costs of mining between "the rich" and "the poor" has also received some attention (see for instance: Bhattacharyya and Williamson, 2015; Fleming and Measham, 2015; Loayza and Rigolini, 2015; Reeson, et al., 2012; Goderis and Malone, 2011; Lay et al., 2008; Ross, 2007). The relationship between mining activity and income inequality for individuals and families living in mining areas has been largely overlooked.

Despite the growing literature on the local economic (and distributional) impact of mining activity, the issue is far less understood in the specific context of Ghana. To the best of our knowledge, we do not know any study on Ghana that has reported evidence on how the effect of mining varies across income groups. While the issue has been addressed elsewhere, as pointed above, the need for country specific evidence cannot be over emphasized. For instance, the effects of mining on local population and how the effect is distributed depend on local institutional setup, social and cultural factors, and the presence of forward and backward linkages of the mining industry to the local economy. These dynamics may be country specific and hence results reported in one country may not carry over to another. It is against this backdrop that we undertake this empirical study on the distributional impact of gold mining in Ghana.

The importance of a study that provides evidence on how the gains (losses) from gold mining in Ghana (on the grounds of her significantcontribution to gold production in the world, and cultural and ethnic diversity of her population) are distributed among different income groups cannot be overestimated. First, unfair distribution of the costs and benefits of mining has a high potential of endangering violent clashes between communities and mining firms. Second, unfair distribution of mineral rents coupled with growing inequality and poverty in mining areas breeds social discontent which can degenerate into large scale civil war, particularly, in fragile states. To overcome these dangers of mining through redistribution and sustainable livelihood policies require a fuller understanding of the effect of mining on household income and welfare and how the effect is distributed over different quantiles of income and welfare.

Using the three most recent household surveys (GLSS 4, GLSS 5, and GLSS 6) on Ghana together with information about open mines during the survey years, we estimated the average treatment effect (ATE) and unconditional quantile treatment effects (QTE) of mining activities on real gross income, employment income, and real per capita household expenditure (a proxy for welfare). We find robust evidence of negative effect of mining on household income and welfare, irrespective of the choice of treatment distance, although the effect is stronger on households living within a 15km radius from an active mine. This confirms the assumption that the effect of mining activity decays with distance.

The results of unconditional quantile regression estimates indicate that the income reducing effect of mining activity falls heavily on households at bottom of the income distribution. The estimated quantile treatment effects were only significant up to the 40th quantile of gross income, and the 50th (median) quantile of employment income. The magnitude of the quantile treatment effects also decreases (in absolute terms) in the quantiles of income, implying that the very poor among the poor suffer the most from the negative effect of mining on income. In the case of household welfare, the interesting revelation of the quantile treatment effect estimates was that the negative effect of mining falls largely on both the "tail" and the "head" of the welfare distribution, with much heavier burden at the tail relative to the head of the distribution. Our paper, thus, provides ample evidence that mining activity is not only harmful to income and welfare, but also has negative effect on inequality in the distribution of income and welfare.

The rest of the paper is organized as follows: Section 2 reviews the relevant literature on the socioeconomic impact of mining in general and the effect of mining on income inequality in particular. Sections 3, describes the econometric techniques applied to data, while data sources and summary statistics on the relevant variables are presented in Section 4. In Section 5, we present and discuss the main results of the paper. Section 6 concludes.

2. Related Literature

Our paper is related to at least two strands of the literature on the linkages between natural resources and economic development. First, this study is related to the general literature on the local socioeconomic impact of mining (see for instance: World Bank 2015; Aragon and Rud, 2015, 2013; Kotsadam and Tolonen, 2016; Tolonen, 2015; Chuhan-Pole et al., 2015). Second, our paper is also related to the literature on the effect of natural resource boom on income inequality (see for instance: Bhattacharyya and Williamson, 2015; Fleming and Measham, 2015; Loayza and Rigolini, 2015; Reeson et al., 2012; Goderis and Malone, 2011; Lay et al., 2008; Ross, 2007).

The literature studying the local economic impact of resource extraction has emphasized the importance of backward linkages as the main means by which the economic benefits of resource extraction are felt by those living close to the resource. This theoretical postulation has been the central idea in many empirical studies on the local economic impact of resource abundance. For instance, Lippert (2014) studies the economic benefits of Copper Belt mine in Zambia on neighboring households. The central result of his study is that an increase in local copper output improves measures of living standards in the respective constituencies through the mines' backward linkages. In particular, Lippert (2014) estimates a 2% increase in real household expenditures is associated with a 10% increase in copper output in the constituency level. The positive effect of natural resource extraction on the local economy contrasts sharply with the enclave thesis that date back to Hirschman (1958).

While the empirical estimates appear to be robust, there is still much less understanding of the distribution of the burden and benefits of resource extraction. Kotsadam and Tolonen (2016), for instance, in a study based on household level data from selected mineral producing countries in Africa, documents evidence of increased female employment from mine opening. They also find evidences of a shift of women into the service sector, although the effect dissipates with distance; and asymmetric effect of mine closure and suspension, with women not fully returning to the agricultural sector, whereas overall employment levels remain low. Aragon and Rud (2013) find evidence of positive effect of large scale mining operations on real income of households of local communities in Peru through the demand for local inputs. Aragon and Rud (2013) also report that local price of non-tradable goods such as housing respond positively to mining. These findings underscore the potential backward linkages from the extractive industries to engender positive spillovers in less developed economies. It however remains a question whether their findings hold for other developing countries. Belanay et al., (2014), have also reported a positive effect of mining activity on income and poverty levels in the Caraga Region in the Philippines. Loayza et al., (2013) also reports a positive impact of mining on producing districts. The authors find evidence that mineral producing districts have better average living standards than otherwise similar districts: larger household consumption, lower poverty rate and higher literacy. However, Loayza et al., (2013) document that the positive impacts of mining dissipates significantly with administrative and geographic distance from the mine, while district level consumption inequality increases in all districts belonging to producing province. They reckoned that the inequalizing effect of mining engender social discontent and violence that is common in most mining areas in the developing world. In a study that explores the relationship between non-oil and gas mining activities on economic growth for nonmetropolitan U.S. counties for the period 2000 to 2007, Deller and Schreiber (2012) found that non-oil and gas mining is associated with lower population growth, and a positive impact on per capita income, but has no impact on employment growth.

In a more recent econometric study on Ghana, Chuhan-Pole et al., (2015) revealed that men are more likely to benefit from direct employment as miners and that women are more likely to gain from indirect employment opportunities in services. They also show that mining improves access to infrastructure (such as electricity and radios) and health outcomes of the children of long established households relative to migrants. Chuhan-Pole et al., (2015) also report that infant mortality rates significantly decrease in mining communities relative to non-mining areas. On the contrary, Aragon and Rud (2015) have reported that farmers located near mines experienced a relative reduction in total factor productivity of almost 40% between 1997 and 2005, with pollution emanating from mining as the most plausible explanation for the agricultural productivity slowdown in mining areas. With agriculture as the backbone of rural economies, Aragon and Rud (2015) finding implies that mining generates negative welfare effects on majority of rural households, which conflicts with the findings reported by Chuhan-Pole et al., (2015) and Aragon and Rud (2013). It is thus very important, at least for policy purposes, to further investigate the distribution of the benefits and costs of mining across different social and economic classes within mining communities.

The literature on the effect of mining (natural resource boom) on inequality (income distribution) has yielded mixed evidence. Within the strand of the local impact of mining on income distribution, Ross (2007) contains detailed discussions on how mineral wealth can affect vertical inequality (inequality between rich and poor households) and horizontal inequality (inequality across districts in a country) and what policy makers can do about both kinds of inequalities. There are studies on natural resource booms and income inequality at both the national and local economies. The national level studies exploit cross-country variations in natural resources wealth to explain equalities in income distribution across countries (see for instance: Goderis and Malone, 2011; Bhattacharyya and Williamson, 2015; Lay et al, 2008). Goderis and Malone (2011) investigated effect of natural resource booms on income inequality from both a two-sector growth theoretic perspective and empirically. From their theoretical analysis, Goderis and Malone reported that under the condition of relatively unskilled labor intensive non-traded sector, inequality falls immediately after a boom, and then increases steadily over time until the initial impact of the boom disappears. Using data on 90 countries between 1965 and 1999, Goderis and Malone (2011) find evidence consistent with the theoretical results, especially for oil and mineral booms. They also found that uncertainty about future commodity prices increases long-run inequality in resource rich economies. Bhattacharyya and Williamson (2015) have also reported evidence that commodity price shocks increases income inequality in Australia. They find that commodity price shock increase the income share of the top 1, 0.05, and 0.01 percent in the short run. They found similar evidence in the long run between commodity prices and top incomes. The findings reported by Bhattacharyya and Williamson (2015) indicate that all top income groups (top 1, 0.05, 0.01) benefits from sustained increases in commodity prices; in particular, the very top end of the income distribution (top 0.05 and 0.01) benefit from commodity booms disproportionately more than the rest of the society.

Studies on the local impact of natural resource booms on both vertical and horizontal inequality in income distribution include: Loayza et al., (2013); Loayza and Rigolini (2015); Reeson et al., (2012); Fleming and Measham (2015). Loayza and Rigolini (2015) find evidence of non-equalizing effect of mining in producing districts in Peru. In particular, they reported evidence of consumption inequality within producing districts is higher than in comparable nonproducing districts. This effect of mining in Peru, according to Loayza and Rigolini (2015), is partly attributed to the attraction of highly skilled (better educated) workers by mining activity. At the same time some local workers, especially farmers, who do not have the skills to work in the mines lost their livelihoods, resulting in widening income gaps. Fleming and Measham (2015), using Gini coefficient as a measure of income inequality, reported a general increase in income inequality across Australian Regions, with income inequality in mining regions being about 4% lower in mining regions relative to non-mining regions. They pointed out, however, that the results show important variations in changes in the Gini coefficient across mining regions, implying that the industry is likely to affect the distribution of local incomes in different ways. Reeson et al., (2012), also found the Gini coefficient of personal income to be significantly associated with levels of mining employment, but nonlinearly. They document that income inequality increases with mining activity, before decreasing at medium and high levels of mining employment. This suggests that the effect of mining on income distribution follows the Kuznets curve pattern. For women, however, Reeson et al., (2012) found that mining activity increases income inequality monotonically. In the case of men, the relationship follows the above described Kuznets pattern. This evidence suggests that mining activity affects men and women differently and the effect varies with the scale of activity and associated levels of employment.

3. Model and empirical strategy

The analysis in this paper follows two sequential steps. First, we estimate the average treatment effect of living close to a mine as has been done in previous studies using the Heckman (1978; 1979) selection model. That is, we examine the mean impact of mining activities on the income and welfare of households living in mining areas. In the second stage, we further investigate the distributional impact by assessing how the effect of mining on income and welfare vary across the different quantiles of income and welfare. This, we achieve by estimating unconditional quantile treatment effect of living close to a mine based on the semiparametric estimation of quantile treatment effects proposed by Firpo (2007).

3.1 Impact of Mining

To estimate the impact of mining activities on households in host communities, we follow an approach similar in spirit to Aragon and Rud (2015), Chuhan-Pole, et. al. (2015), and Kotsadam and Tolonen (2016). We define our treatment (i.e. whether a household resides in a mining area or otherwise) by using geographic distance of the location of the household to the nearest active mining site.

Mining activities in an area is determined largely by presence of mineral deposit (Tolonen, 2015), often in large commercial quantities. However, there is often a time lag between the discovery of deposit and actual extraction of the deposit. Also, given the fact that most of these extractive resources are non-renewable, the extraction of the mineral is limited to a given period of time. On the basis of the above, the exact impact of mining activities will vary across communities with mineral deposits depending on whether the community has an active mine or not. Therefore these factors have to be considered in order to causally identify the impact of mining on host communities. In this paper, we define mining area as an area with the presence of an active/open mine, which implies a mine actively in operation during the time of the survey.

Consequently, we specify our baseline model as follows

$$y_{ij} = \theta_0 + \theta_1 T_{ij} + X_{ij}\delta + \eta_t + \varepsilon_{ij} \tag{1}$$

where y_{ij} is the outcome variable (income and welfare) of household i in community *j*; T_{ij} is the treatment indicator which equals 1 if the household is located in the neighbourhood of an active/open mine and equals zero if otherwise; X_{ij} is a vector of covariates including household characteristics and community attributes; η_t is a year fixed effect; while ε_{ij} is the residual term.

3.2 Identification Strategy

To identify the causal impact of mining activities on the set of outcome variables, requires that both assignment of, and selection to, treatment are both purely random process, making the treatment indicator in equation (1) truly exogenous. Indeed the distribution of mineral deposits can be assumed to be random and hence exogenous (Tolonen, 2015; Alcott and Keniston, 2014). However, the argument of random distribution of mineral deposit is not enough justification for the exogeniety of mining activities (treatment) in a given area. Hence, OLS estimation of equation (1) will not be able to identify the causal impact of mining on the set of outcome variables. The reason is that the treatment indicator is potentially endogenous due to the following reasons. First, the discovery of the deposits and the subsequent extraction of the deposits depend largely on a set of institutional, social, economic and political factors (see, Eggert 2002; Tolonen, 2015). This suggests that the presence of mining activities in a given location may not be entirely exogenous. That notwithstanding, there are instances where communities develop following the establishment of mines in a given area. Secondly, given that our unit of analysis is at the household level, issues of selection bias cannot be completely ignored. Under the assumption of free mobility of labour within a given country, households can self-select into and out of mining communities either in search of employment opportunities or perhaps as a result of loss of livelihood due to the expansion in mining activities (e.g., loss of farm lands to mining activities). To the extent that the above mechanisms are operative, the assumption of independence between the error terms and the treatment indicator is no longer valid. In other words our treatment indicator is endogenous in the model, i.e., $E\left(T_{ij}, \varepsilon_{ij}\right) \neq 0$, hence addressing these issues will be required in order to identify the causal impact of living close to an active mine. Admittedly, the first argument may be difficult to resolve at least within the context and limits of the data used in this study, and also may have little impact on the results. However, the issue of self-selection is crucial and requires sufficient attention in order to identify the impact of interest.

To address selection bias in our model, we rely on the Heckman (1978, 1979) two-step selection model which estimates the following

$$y_{ij} = \alpha_0 + \alpha_1 T_{ij} + X_{ij}\beta + \varepsilon_{ij} \qquad (2a)$$

$$T_{ij} = \gamma_0 + \gamma_1 Z_{ij} + X_{ij}\phi + u_{ij} \qquad (2b)$$

$$\varepsilon_{ij} = \delta Q_{ij} + e_{ij} \qquad (2c)$$

$$u_{ij} = \rho Q_{ij} + \xi_{ij} \qquad (2d)$$

Equations (2a) and (2b) are the outcome and selection models respectively; α_1 is the average treatment effect (ATE); X is a vector of common control variables that affect both outcome and selection into treatment; Q is a common unobservable component[1]; e and ξ are two exogenous shocks with zero unconditional mean. Here, the selection into treatment depends on observable household and community characteristics (X) and an instrumental variable (Z). We use the probability of the individual (household head) being born in a particular locality (Z) as an instrument for the probability of the person to live close to an open mine. Thus, the exclusion restriction assumption being advanced here is that, even though there is the likelihood for an individual born close to an open mine may choose to live in the same area it does not guarantee that the individual will necessarily have higher income or welfare. From the above argument, if our assumption holds then we are confident that the relevance and excludability (exogeneity) properties of a good instrument is justified, hence $\hat{\alpha}_1$ can be regarded as the average treatment effect of living close to an open/active mine on the outcome variables.

Further, we assume heterogeneity across households in terms of age, year of education and marital status of household heads. This implies that there is possibility for households to differ in terms of the years of education, age, and marital status of the household head.

3.3 Heterogeneity in the Impact of Mining

Contrary to the preceding section, which estimates the average impact of mining activities, in this section, we relax the assumption of a homogenous impact and examine how the impact of mining activities on households' income and welfare vary across the income and welfare distribution using an unconditional quantile regression approach. Our choice of unconditional quantile approach to estimating quantile treatment effect is motivated by its advantages over the conditional quantile regressions. First, the definition of the unconditional QTE does not change when we change the set of covariates, X. While conditional $QTEs$ are defined conditionally on the value of the regressors, unconditional effects summarize the causal effect of a treatment for the entire population (Frolich and Melly, 2010).. Second, unconditional effects can be estimated consistently at the $\sqrt{n}$ rate without any functional form (parametric) restrictions, which is not

[1] Since Q is unobservable, it is part of both the error terms ε (in the outcome equation) and u (in the selection equation).

possible for conditional effects (Frolich and Melly, 2010). Last but not least, conditional and unconditional $QTEs$ coincide in the absence of covariates.

We estimate unconditional quantile treatment effect of living close to an active mine by following the efficient semiparametric estimation of quantile treatment effect proposed by Firpo (2007). The unconditional quantile treatment effect for quantile τ is given by

$$\Theta^{\tau} = y^{\tau}_{ij,T=1} - y^{\tau}_{ij,T=0} \tag{3}$$

Θ^{τ} is the unconditional quantile treatment effect of the τth quantile of the outcome variable of interest, income or welfare. Thus the QTE is estimated as the different between the outcome variable between the treatment and control households at each quantile of the distribution of the outcome variable. Note however, that although our objective is to estimate the unconditional effect, we make use of the individual and community characteristics (vector X) for two reasons. The inclusion of the covariates help us to correct potential selection biases in estimating the treatment effects, as well as increasing the efficiency of the estimation of the impact estimates. However, the definition of the treatment effects is not a function of the covariates. This is an advantage over the conditional quantile treatment effect, QTE, which changes with the set of conditioning variables even if the covariates are not needed to satisfy the selection on observables assumption (Frolich and Melly, 2010). For further details on the specification of the unconditional quantile treatment effect model, interested readers may refer to ……..

4. Data and descriptive statistics

This study combines data from a geo-referenced household survey data with geo-data on the location of gold mines in Ghana. Specifically, we utilize the three most recent waves of Ghana Living Standards Survey (GLSS 4 (1998/99), GLSS 5 (2005/06), & GLSS 6 (2012/2013)). This dataset is a nationally representative repeated cross-sectional survey of households in Ghana, and conducted by the Ghana Statistical Service with support from the World Bank and other agencies. The sampling frame for the survey was the population living in private households in Ghana. The above sample frame was divided into two sampling units, a primary and secondary sampling unit. The primary sampling unit was defined as the census enumerated areas (EAs) that are stratified into the ten administrative regions of Ghana based on proportional allocation using the population in each of the ten regions. The second sampling unit on the other hand was defined as the households living in each of the EAs. The sampling design for the survey was that of two-stage stratified random sampling approach, where in the first stage 550 EAs was considered, while in the second stage, 15 households per EA was considered. All the data in the three waves used in the study are geo-referenced, at the enumeration level or clusters. In other words, they contain the GPS coordinates of the communities within which households are located rather than the exact location of the households. This is mainly due to the privacy issues related to the households interviewed. Thus our implicit assumption is that households in the same cluster share the same location. Nonetheless, this does not pose any serious limitation to our study in the sense that the use of the community location suffices in determining whether a household is located in a mining community or otherwise. The same approach has been widely used by studies such as Tolonen (2015), Aragon and Rud (2015), Chuhan-Pole, et al. (2015) and Kotsadam and Tolonen (2016).

Using the above survey design, data was collected on the following key variables: household real gross income (*realgrossincom*) and real employment income (*realtotempincom*)), welfare (real per capita household expenditure (*realwelfare*)), demographic characteristics such as education, gender of the head of the household (*sexhead*: a binary variable, which is equal to one if the household head is a male, and zero otherwise), age of the household head, size of the household, and mining sector employment (*mining;* a binary variable which takes the value of 1 if the household head is employed in the mining sub-sector, and zero otherwise). The education variable is measured in terms of years of schooling.

Using geo-data of gold mines in Ghana obtained from Aragon and Rud (2015) and Chuhan-Pole et al., (2015) we match the location of the mines (see Table A2 in appendix for mines) to the household data and then compute the geographic distances of each enumeration area to the mines. Our definition of the mining areas is based on various distance thresholds, from a minimum of 15 km radius to a maximum of 90 km radius. As argued by Tolonen (2015) and Kotsadam and Tolonen (2016), there is no universally accepted distance threshold for the classification of mining communities and non-mining communities. As a result even though we use a distance bandwidth of 15 km from a mine, we complement our analysis by estimating the model with 20km distance from a mine, which is the maximum treatment distance used in the existing literature (see: Aragon and Rud, 2015; Chuhan-Pole, et.al., 2015; Kotsadam and Tolonen, 2016). Thus, for instance, OpenXdist_15km is a treatment indicator variable which takes the value of one if the household lives within a 15 kilometre radius to an open mine and zero, if otherwise. After data cleaning, we arrive at total of 31,457 households (see Table A1 in appendix for distribution across waves). Following earlier studies (such as Chuhan-Pole, et al., 2015; World Bank, 2015) we restrict our sample to households leaving within 100 kilometre radius of an open mine. As such, our total sample for this study is about 17,925 households. In all the estimations in this paper, we pooled data across the three nationwide surveys described above. Table 1 presents the summary of descriptive statistics of our pooled data from GLSS 4, GLSS 5 and GLSS 6. From Table 1, household size in our sample ranges between 1 and 22, with average household size being about 4 with a standard deviation of 2.48, indicating significant variation in household size in the sample. Of the 16,174 respondents for which information on the sector of employment is available, just about 2% (mean of 0.0201) of the active household heads are employed in the mining sub-sector. Regarding the gender of the household head, of the 17,925 household heads in our sample, approximately 67% are males. The age of the household head measured in years ranges between 15 and 99 years. The mean age of the household head in years is 45.27 (45 year and about 3 months) years with a standard deviation of 15.72 years, suggesting much wider dispersion in the age of the household heads.

By restricting treatment to living within 15km radius to open mine (openXdist_15), only approximately 8.4% of the 17,925 households living within 100km radius to a mine fall into the treated category, with the remaining 91.6% serving as control. When we extended the treatment sample of households living within a 20km radius of an open mine (openXdist_20), the share of the treated in the total sample of 17,925, increases to about 12%.

Table 1: Descriptive Statistics

	Pooled sample		15km distance			20km distance		
Variables	N	mean	Treatment	Control	Difference	Treatment	Control	Difference
Household size	17,925	3.917	4.01	3.91	-0.104 (0.067)	4.01	3.91	-0.0995 (0.056)*
Mining sector employment	16,174	0.0201	0.090	0.014	-0.077 (0.004)***	0.078	0.012	-0.067 (0.003)***
Married	17,923	0.508	0.518	0.507	-0.011 (0.013)	0.511	0.507	-0.004 (0.011)
Sex of head	17,925	0.673	0.692	0.6714	-0.021 (0.013)*	0.682	0.672	-0.01 (0.011)
Born here	17,870	0.494	0.503	0.493	-0.010(0.013)	0.509	0.492	-0.017(0.012)
Age of head	17,925	45.27	45.22	45.28	0.063 (0.423)	45.41	45.26	-0.15 (0.362)
Education of head	13,736	9.429	9.245	9.446	0.202 (0.122)*	9.27	9.45	0.18 (0.104)*
Log of real gross income	17,424	7.112	7.562	7.071	-0.491 (0.047)***	7.52	7.06	-0.466 (0.04)***
Log of real total employment income	8,801	7.187	7.486	7.154	-0.333 (0.052)***	7.54	7.13	-0.414 (0.044)***
Log of real welfare	17,925	6.759	6.963	6.74	-0.223 (0.028)***	6.97	6.73	-0.241 (0.024)***
OpenXdist_15km	17,925	0.0838						
OpenXdist_20km	17,925	0.120						

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Using the openXdist_15km treatment indicator, we observe a significant difference between the treated and control groups with regard to sector of employment, gender, years of education, income and welfare (see Table 1). With the exception of years of education, the treatment group has higher mean values for these variables earlier mentioned. Similar situation exist if living close to_20km radius of an open mine is used as the treatment indicator, except for gender which is relatively equal for the treatment and control groups. In addition, households living close to 20km radius of an open mine have significantly higher household members than those living farther away.

5. Results and discussions

In this section of the paper, we present and discuss the main results and implications thereof. We undertake the analysis in two steps. First, we estimate the effect of living in vicinity of an open mine on household income and welfare. In doing so, we use two measures of household income: real gross income and real employment income. Following previous studies on household welfare, we measure household welfare by real household expenditure per capita. Second, we investigate the distributional impact of mining on household income and welfare by estimating the effect of living in the vicinity of an open mine on different quantiles of income and welfare at the household level. The results[2] of the first part of our investigation (estimation of average treatment effects) are presented in Table 2, and Table A4 and A5 of the appendix to this paper.

In columns 1 – 3 of table 2 restrict the treatment group to households living within 15 kilometre radius to an open mine (with the rest serving as controls), while in columns 4 – 6, we expand the treatment group to households living within 20 kilometre radius of an open mine. The full sample for the analysis in this paper is restricted to households living within 100 kilometre radius to an open mine. All regressions include additional controls and year fixed effects. A crucial assumption about the Heckman's two-step selection model is that of joint normality of the errors (residuals) of the selection and outcome equations. The combined effect of the selectivity bias variable (inverse Mill's ratios) on unconditional income or welfare is as expected. The combined truncation effect is positive, meaning that the process of self-selection serves to enhance the unconditional expected income or welfare. The combined truncation effect of inverse Mill's ratios is given as the difference between the coefficient of the selectivity bias variables (inverse Mill's ratio) for the treatment group (wL_1) and that of control (wL_0). All the alternative specifications estimated using the selection model satisfied this

[2] The first stage estimation is shown in Table A2 in appendix. The result shows a significant and positive effect of the probability of an individual born in a locality on the treatment indicator.

condition since $wL_1 - wL_0 > 0$ for all the cases considered (see tables A4 & A5 of the appendix). This implies that accounting for selection biases using the heckit model improves the results.

5.1 Income and welfare effects of gold mining in Ghana

The estimates based on the Heckman's two-step selection (heckit) model indicate that income and welfare are lower among households living in the vicinity (within 15km and 20km radius) of open mine relative to households farther away from an active mine. On the basis of the estimates, the *ATEs* suggest that gross income, and household welfare, on average, are approximately 1.378%; and 1.651% respectively lower for households living within 15km radius of an active mine, relative to households farther away. Effect on employment income, however, is not statistically significant; indicating that there is no difference in employment income for those living close to open mine and those living farther away, once we control for direct employment in mining related activity. The fact that the effect of mining on employment income is not significant after controlling for employment in mining sector is not surprising. The reason is that formal sector wages, particularly in the public sector, has virtually no spatial variation. However, the change in sectoral composition in employment (shift from agricultural employment to service sector jobs as documented in Kotsadam and Tolonen, 2016) has significant effect on household income and welfare, with likely effect on their distribution across gender, age, sector of employment, income level, among others. Our results provide strong evidence that average income and welfare of households are lower if the household lives in a vicinity of open mine than when the household is farther away from the mine.

By extending the treated group to include households living within 20km radius, the estimates remained very stable in terms of the sign, and the level of statistical significance on the coefficients. However, the size of the effects of living closer to an active mine is consistently lower relative to the case when the treated group are restricted to within 15km radius, indicating the distance decay hypothesis of the impact of living close to mine has some validity. We take these as robust evidence that mining has negative effect on income and welfare on households living in vicinity of open mine. Interestingly, we find that while mining reduces income and welfare of the households living close to the mine, income and welfare of households whose heads are employed in the mining sub-sectors are higher than those who have not direct employment in mining. This result is very plausible given that the compensation of labour in mining is higher (on average) relative to a worker with similar skills employed in a sector other than mining. However, due to the limited employment opportunities in large scale mining because of their capital intensive nature, this benefit of mining employment accrues to just about 2% of households in our sample (see Table 1).

In addition to the sector of employment of the active household head, other personal and household characteristics such as gender, age, and years of schooling, of the household head and household size have significant effects on household income and welfare. See tables A4 and A5 of the appendix for the details on the effects of household and individual characteristics of the household head on income and welfare. Also, we find that gross income of the household is not independent of the gender of the household head, but employment income is about 28% higher for male headed households. This confirms the popular view that labour market participation and earnings are higher for males relative to females. Interestingly, welfare is lower for male headed households. The reason why expenditure per person for male headed households being lower relative to female headed households maybe that men are more likely to have other spending commitments outside the "home" relative to females. The estimates indicate that welfare levels are about 15% lower when the active household head is male relative to female headed households (irrespective of whether treatment is restricted to within 15km or 20km radius). The age of the household head had positive and statistically significant effect on gross income of the household. The effects of age on employment income and welfare are both insignificant. The marital status of the household head also matters for household income and welfare levels. We find gross income, employment income, and welfare are approximately 29%, 22% and 11% respectively higher when the household head is married relative to when he is not. This may be due to complementarity of resources by couples. Education of the household head (measured in years of schooling) also increases gross income, employment income and welfare of the household, an indication of higher returns to schooling. Specifically, the results show that each additional year of schooling raises gross income, employment income and welfare by approximately 6.4%, 9.4%, and 5%, respectively. Household size also has beneficial effect on both gross income and employment income, but a negative effect on household welfare. Per the estimates based on the heckit model, increase in household size by one member raises gross income by about 13%, employment income by about 6%, but reduces household welfare by about 12.8%.

Table 2: Conditional endogenous treatment effect using 15km and 20km distances from mine

	Gross Income	Employment Income	Welfare	Gross Income	Employment Income	Welfare
Variables	(1)	(2)	(3)	(4)	(5)	(6)
OpenXdist_15km	-1.378**	-0.822	-1.651***			
	(0.562)	(0.561)	(0.292)			
Mining sector employt	0.970***	1.170***	0.917***	0.714***	0.953***	0.774***
	(0.233)	(0.238)	(0.122)	(0.201)	(0.205)	(0.100)
OpenXdist_20km				-1.061**	-0.480	-1.365***
				(0.450)	(0.456)	(0.226)
Constant	5.718***	5.228***	7.159***	5.625***	5.179***	7.128***
	(0.0786)	(0.0899)	(0.0391)	(0.0736)	(0.0848)	(0.0370)
Observations	12,216	7,140	12,538	12,216	7,140	12,538
R-squared	0.552	0.438	0.699	0.551	0.438	0.699
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

5.2 The Distribution of income and welfare effects of mining

The results presented in the previous subsection indicated that the "average treatment effect" of open mine is negative. This has the implication that mining activity reduces household income and welfare for households living within at least 20km radius to open mines. In this section of the paper, we investigate how the burden (negative effect on income and welfare) of mining is distributed among "low income" households and "high income" households using quantile regressions. We estimate the same specification as in section 5.1, but on different quantiles (0.1; 0.2; 0.3; 0.4; 0.5; …, 0.9) of real income (gross and employment) and welfare. We estimate these using unconditional quantile regression technique proposed by Firpo (2007). The results are presented in tables 3 and 4 for 15km radius and 20km radius treatments respectively.

Using living within 15km radius to an open mine as a treatment, the results reported in Table 3 show that the effect of mine is negative at all income quantiles for gross income, but the magnitude of the treatment effect diminishes as we move from the bottom to the top of the income ladder. Moreover, the treatment effect becomes insignificant after the 40th quantile of gross income. The estimated treatment effect for the 10th quantile of gross income is -3.365 which is statistically significant at 5% error margin. This means that real gross income of households living within 15km of an open mine is about 3.36% lower compared to households farther away, who occupy the bottom 10% of income distribution at the household level. In the case of the 20th quantile, the estimated treatment effect is -3.316 for real gross incomes, which is statistically significant at 5% error level. It follows that real gross income of households living within 15km of an open mine is about 3.32% lower relative to households farther away, who occupy the bottom 20% of income distribution at the household level. The estimated treatment effects for the 30^{th} and 40^{th} quantiles of real gross incomes are -3.165 and -2.53, both of which are statistically significant at 1% and 10% margins of error respectively. Thus, at the bottom 30% (40%) of the distribution of household real gross incomes, households living within 15km radius to an open mine have incomes that are 3.17% (2.53%) lower in comparison to households farther away from the mine. The estimated treatment effects above the 40^{th} quantile are all statistically insignificant at all conventional margins of error permissible.

Table 3: Unconditional endogenous quantile treatment effect using 15km distance from mine

	Using optimal smoothing parameters			Robustness check with under smoothing parameters		
	(1)	(2)	(3)	(4)	(5)	(6)
	Gross Income	Employment Income	Welfare	Gross Income	Employment Income	Welfare
Quantile						
0.1	-3.365**	-4.098	-3.954**	-3.538*	-3.570	-3.957**
	(1.343)	(6.137)	(1.628)	(1.911)	(3.762)	(1.632)
0.2	-3.316**	-3.929*	-3.665***	-3.443**	-3.464	-3.747***
	(1.479)	(2.383)	(1.304)	(1.506)	(5.445)	(1.250)
0.3	-3.165***	-3.480**	-3.372*	-3.307***	-3.150	-3.446**
	(1.219)	(1.728)	(1.754)	(1.194)	(4.865)	(1.641)
0.4	-2.533*	-3.160**	-2.965**	-2.961**	-3.064	-3.085**
	(1.380)	(1.382)	(1.416)	(1.466)	(2.369)	(1.408)
0.5	-2.165	-3.044**	-2.244	-2.389	-3.052*	-2.324
	(1.528)	(1.322)	(1.554)	(1.805)	(1.767)	(1.884)
0.6	-1.505	-2.674	-1.645*	-2.258	-2.674	-1.696*
	(1.950)	(1.719)	(0.856)	(2.971)	(1.965)	(0.964)
0.7	-0.958	-2.383	-1.516*	-1.790	-2.317	-1.453
	(2.682)	(1.918)	(0.790)	(3.848)	(2.013)	(0.916)
0.8	-0.660	-2.151	-1.538**	-1.400	-2.077	-1.493*
	(2.915)	(1.633)	(0.687)	(2.107)	(1.709)	(0.794)
0.9	-0.393	-2.098	-2.011**	-1.163	-2.054	-1.943*
	(2.872)	(1.656)	(0.969)	(2.207)	(1.760)	(1.167)
Observations	12,216	7,140	12,538	12,216	7,140	12,538
Bandwidth	2	2	2	1	1	1
Lambda	0.800	0.800	0.800	0.400	0.400	0.400

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4: Unconditional endogenous quantile treatment effect using 20km distance from mine

	Using optimal smoothing parameters			Robustness check with under smoothing parameters		
	(1) Gross Income	(2) Employment Income	(3) Welfare	(4) Gross Income	(5) Employment Income	(6) Welfare
Quantile						
0.1	-3.388***	-3.677**	-3.502***	-3.742***	-3.423**	-3.598***
	(1.003)	(1.644)	(0.634)	(1.157)	(1.331)	(0.668)
0.2	-3.243***	-3.364	-3.154***	-3.298***	-2.934**	-3.202***
	(0.884)	(4.039)	(0.548)	(1.032)	(1.223)	(0.573)
0.3	-2.976***	-3.137*	-2.978***	-3.165**	-2.405	-3.052***
	(0.991)	(1.837)	(0.597)	(1.248)	(1.697)	(0.607)
0.4	-2.940***	-2.807**	-2.384*	-3.094**	-2.623	-2.620*
	(0.994)	(1.241)	(1.291)	(1.208)	(2.954)	(1.427)
0.5	-2.251	-2.410*	-1.879	-2.619	-2.459	-1.999
	(1.617)	(1.286)	(1.246)	(2.211)	(1.505)	(1.330)
0.6	-1.069	-1.874	-1.128*	-2.036	-1.993	-1.316*
	(2.659)	(1.695)	(0.627)	(3.541)	(1.974)	(0.786)
0.7	-0.580	-1.339	-1.222**	-1.274	-1.305	-1.233**
	(2.351)	(1.286)	(0.535)	(2.429)	(1.464)	(0.566)
0.8	0.378	-1.243	-1.215***	-0.670	-1.189	-1.213***
	(1.734)	(1.137)	(0.409)	(1.725)	(1.313)	(0.445)
0.9	0.300	-1.158	-1.346***	0.224	-1.040	-1.290**
	(1.017)	(1.438)	(0.490)	(0.998)	(1.508)	(0.525)
Observations	12,216	7,140	12,538	12,216	7,140	12,538
Bandwidth	2	2	2	1	1	1
Lambda	0.800	0.800	0.800	0.400	0.400	0.400

Standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Using a distance of 20km to open mine as a treatment, the results were very similar to the 15km treatment case. In particular, the quantile treatment effects were -3.39, -3.24, -2.98, and -2.94 for the 10th, 20th, 30th, and 40th quantiles respectively. This means that real gross income for households living within 20km radius to open mine are 3.39%, 3.24%, 2.98%, and 2.94% lower for the households at the bottom 10%, 20%, 30%, and 40% of the income distribution respectively, in comparison to households living farther than 20km to an open mine. Similar to the 15km radius treatment case, the estimated quantile treatment effects after the 40th quantile are all statistically insignificant.

With reference to real income from employment, the estimated quantile treatment effects were negative and statistically significant for the 20th, 30th, 40th and 50th (median) quantiles, with the estimates on the remaining quantiles being negative but insignificant, for 15km radius treatment. The estimated quantile treatment effects were -3.93, -3.48, -3.16, and -3.04 for the 20th, 30th, 40th and median quantiles respectively. The implication from these estimates is that real incomes from employment for the bottom 20%, 30%, 40% and 50% of household income distribution are 3.93%, 3.48%, 3.16%, and 3.04%, respectively, lower in comparison to households living farther than 15km to open mine but occupy similar positions on the income distribution. The results remained quite stable when we used a distance of 20km to open mine as treatment. Under this scenario, the quantile treatment effects were negative and statistically significant for the 10th, 30th, 40th and 50th quantiles. The estimates on the remaining quantiles are negative but statistically insignificant. Regarding the size of the impact of "treatment", our estimates indicate that for households living within 20km radius to open mines, real income from employment is approximately 3.68%, 3.14%, 2.81%, and 2.41% lower relative to households farther than 20km, for the bottom 10%, 30%, 40%, and 50% of the income distribution respectively.

The results here (both real gross income and employment income) indicate that the negative effect of mining on income reported in tables 2, 3, and 4 above falls heavily on households at the bottom of the income distribution. The quantile treatment effects are only significant up to at most the 50th (40th) quantile for employment (real gross) income, with the absolute size of the effect decreasing as one move up on the income ladder. This means that "the poor" gets "poorer" in mining areas while having no significant effect on "the rich" in these areas, when they are both compared with similar income groups outside mining areas. This evidence is taken to mean that mining activity aggravate inequality in mining areas. The reasons why the negative effect of mining falls on the very poor in mining areas may include the following. First, the poor are less able to adapt to the new conditions after mines opening by way of switching from say supply of agricultural labour to service sector employment. Second, most poor farmers do not own the land on which they farm and hence may not receive income compensation when the mines take over farmlands. Third, the very poor in the mining communities are more likely to be those who only supply agricultural labour on daily basis. Hence, reduction in agricultural employment due to competition between mining and agriculture for land reduces the demand for hired labour in agriculture, which is a source of negative income shock.

In terms of household welfare, the trend in the effect of living close to open mining on households is different from its effect on real gross and employment income. In the case of welfare, the estimated quantile treatment effect is insignificant only at the median quantile, regardless of whether treatment is restricted to 15km or 20km radius. Interestingly, the estimated quantile treatment effects suggest that the effect of mining on welfare follows a U-shape in the quantiles, in the sense that the negative effect of mining activity on household welfare is stronger at the lower and upper quantiles. Using a distance of 15km to an open mine as treatment, the estimated treatment effect for the 10^{th} quantile of welfare was -3.95. The treatment effect reduced to -3.67, -3.37, and -2.97 at the 20^{th}, 30^{th}, and 40^{th} quantiles respectively. Technically, the effect dropped to zero at the median quantile since the estimated quantile treatment effect is not statistically significant at all acceptable margins of error.

From the 60^{th} to 90^{th} quantiles, all the estimated treatment effects were negative and statistically significant, with the absolute size of the effects increasing in quantiles. The estimated quantile treatment effects for the 60^{th}, 70^{th}, 80^{th}, and 90^{th} quantiles of household welfare were -1.65, -1.52, -1.54, and -2.01 respectively. The results for the 20km treatment follows the same trend, although, the magnitudes of the effects were consistently lower compared to the 15km case. This is consistent with the assumption that the effect of mining activity decreases with distance. Again, although effect on welfare of mining activity is significant at the lower and upper quantiles of welfare distribution, the absolute size of the effect is stronger at the lower quantile. This means that the poor bears the lion's share of the burden of mining on household welfare. The reason why the effect of mining on welfare is large at the bottom may follow similar reasons as that of income. What is difficult to account for is why the effect is also strong at the upper quantiles of the welfare distribution. A possible reason may be that, the growing incidence of poverty and inequality may induce the relatively high income households to engage in precautionary savings by cutting down current expenditure. This precautionary saving serve as insurance for household welfare against any potential negative income shocks in the future.

6. Conclusion

This paper examined the effect of mining on income and welfare distribution in mining areas using quasi-experimental approach. Specifically, we considered living within 15th km radius to an active mine as baseline treatment indicator and check the robustness of the estimates with using a treatment distance of 20km radius to an open mine. A key assumption to identification of the effect of mining is that the effect (whether positive of negative) decays with distance. Given this background, we set out to answer the following two questions. What is the average effect of mining activity on the income and welfare of households living in the mining areas? How does this effect, if any, vary across different quantiles of income and welfare?

Using three most recent household surveys (GLSS 4, GLSS 5, and GLSS 6) together with information about open mines during the survey years we estimated the average treatment effect (ATE) and unconditional quantile treatment effects (QTE) for real gross income, employment income, and real per capita household expenditure (a proxy for welfare). We find robust evidence of negative effect of mining on household income and welfare is irrespective of the choice of treatment distance, although the effect is stronger for the 15km treatment distance, which confirms the assumption that the effect of mining activity decays with distance.

The results of unconditional quantile regression estimates indicate that the income reducing effect of mining activity falls heavily on households at bottom of the income distribution. The estimated quantile treatment effects were only significant up to the 40th quantile of gross income, and the 50th (median) quantile of employment income. The magnitude of the quantile treatment effects also decreases (in absolute terms) in the quantiles of income, implying that the very poor among the poor suffer most from the negative effect of mining on income. In the case of household welfare, the interesting revelation of the quantile treatment effect estimates was the negative effect of mining on falls largely at both the "tail" and the "head" of the welfare distribution, with much heavier burden at the tail relative to the head of the distribution. Our paper, thus, provides ample evidence that mining activity is not only harmful to income and welfare, but also has negative effect on inequality in the distribution of income and welfare. Mining makes society more unequal in terms of income and welfare than they would have otherwise been in the absence of mining. There is therefore the need for polices that targets the very poor in mining areas in a way of internalizing the negative effect of mining on the poor. In this regard, the 20% (10% to mining development fund in producing districts and 10% to the stools) of mining royalties transferred to mining areas should be channelled into pro-poor programmes. Further, alternative livelihood programmes and corporate social responsibility activities of the mining companies should specifically target the poor in mining areas.

Appendix

Table A1: Sample selection

Survey wave	Households	Sampled Households (within 100km)
GLSS 4	5,998	4,818
GLSS 5	8,687	6,244
GLSS 6	16,772	6,863
Total	31,457	17,925

Table A2: Mines and their activeness status

Mine Name	**1999**	**2005**	**2012**	**Mine Name**	**1999**	**2005**	**2012**
Adom	0	0	1	Kubi	0	0	1
Ahafo-Ntotoroso	0	1	1	Kwahu Praso	0	0	0
Akrokeri Property	0	1	1	North Ashanti	1	0	1
Akyem	0	0	1	Noyem	0	1	1
Apapam (Kibi)	0	0	1	Nzema Gold Project	0	0	1
Asumura	0	0	1	Obuasi	1	1	1
Banso/Muoso	0	0	0	Ochinso (Akwatia)	0	0	0
Benso (Hbb)	0	0	1	Safric	0	0	0
Bibiani	1	1	1	Safric Extension	0	0	0
Bogoso/Prestea	1	1	1	Sian	0	0	0
Cape Three Points	0	0	1	Tarkwa	1	1	1
Central Ashanti	1	1	0	Tinga	0	0	0
Chichiwere	1	0	0	Tumentu	0	0	0
Chirano	0	1	1	Wassa	1	1	1
Cluster/Chert Ridge	0	0	1	Obotan	1	1	0
Damang Mine	1	1	1	Essase Placer	1	1	0
Dunkwa (Mampon)	1	1	1	Dunkwa Palcer	0	0	0
Esaase Gold Project	1	1	1	Konong/Obenamasi	1	1	1
Hwiden-Ahafo Project	0	0	1	Edikan-Aynfuri	1	1	1
Iduapriem /Teberebie	1	1	1	Jeni-Bonte	1	0	0
Kade	0	0	1	Prestea Sankofa	1	0	0
Kanyankaw	0	0	1				

Note: Entries in the columns "1999", "2005" and "2012" are binary indicator where 1 means the mine is opened and active in that year, 0 otherwise.

Table A3: First stage probit estimation

	OpenXdist_15km	OpenXdist_20km
VARIABLES	(1)	(2)
Born here	0.0897***	0.108***
	(0.0330)	(0.0300)
Household size	0.0164**	0.0206***
	(0.00755)	(0.00688)
Mining sector employ.	h1.083***	1.105***
	(0.0782)	(0.0769)
Sex of head	-0.0337	-0.0407
	(0.0412)	(0.0374)
Age of head	0.00130	0.00198
	(0.00135)	(0.00123)
Married	0.0255	-0.00700
	(0.0390)	(0.0356)
Education of head	-0.00819*	-0.00809**
	(0.00424)	(0.00385)
Constant	-1.737***	-1.649***
	(0.0846)	(0.0771)
Observations	12,538	12,538
Pseudo R2	0.0583	0.0705
LR chi2(12)	436.8	669.6
Year FE	Yes	Yes

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A4: Full results of mean effects with 15km treatment distance

	Gross Income	Employment Income	Welfare
Variables	(1)	(2)	(3)
Open*Dist_15km	-1.378**	-0.822	-1.651***
	(0.562)	(0.561)	(0.292)
Household size	0.129***	0.0596***	-0.128***
	(0.00518)	(0.00653)	(0.00297)
Mining sector employt	0.970***	1.170***	0.917***
	(0.233)	(0.238)	(0.122)
Sex	0.0284	0.283***	-0.150***
	(0.0268)	(0.0375)	(0.0132)
Age	0.00374***	-0.000256	0.000464
	(0.000921)	(0.00136)	(0.000478)
Married	0.292***	0.226***	0.113***
	(0.0251)	(0.0331)	(0.0128)
Education	0.0637***	0.0937***	0.0502***
	(0.00293)	(0.00342)	(0.00158)
ws_Age	0.00344	-0.00358	0.00250*
	(0.00283)	(0.00399)	(0.00142)
ws_Married	-0.00250	0.0509	-0.0944***
	(0.0761)	(0.0998)	(0.0363)
ws_Education	-0.0181**	-0.0122	-0.0105**
	(0.00879)	(0.0113)	(0.00449)
wL1	0.532**	0.248	0.692***
	(0.257)	(0.263)	(0.135)
wL0	-1.763***	-1.289**	-1.884***
	(0.562)	(0.514)	(0.281)
Constant	5.718***	5.228***	7.159***
	(0.0786)	(0.0899)	(0.0391)
Observations	12,216	7,140	12,538
R-squared	0.552	0.438	0.699
Year FE	Yes	Yes	Yes

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A5: Full results of mean effects with 20km treatment distance

	Gross Income	Employment Income	Welfare
Variables	(1)	(2)	(3)
Open*Dist_20km	-1.061**	-0.480	-1.365***
	(0.450)	(0.456)	(0.226)
Household size	0.127***	0.0607***	-0.128***
	(0.00541)	(0.00662)	(0.00305)
Mining sector employt	0.714***	0.953***	0.774***
	(0.201)	(0.205)	(0.100)
Sex	0.0338	0.277***	-0.151***
	(0.0271)	(0.0378)	(0.0132)
Age	0.00357***	-0.000233	0.000619
	(0.000954)	(0.00141)	(0.000498)
Married	0.284***	0.217***	0.0997***
	(0.0253)	(0.0335)	(0.0129)
Education	0.0657***	0.0943***	0.0507***
	(0.00294)	(0.00341)	(0.00158)
ws_Age	0.00372	-0.00260	0.00244**
	(0.00258)	(0.00357)	(0.00123)
ws_Married	-0.0228	0.00520	-0.0795***
	(0.0673)	(0.0861)	(0.0308)
ws_Education	-0.0135*	-0.000265	-0.00551
	(0.00753)	(0.00946)	(0.00390)
wL1	0.487**	0.139	0.603***
	(0.212)	(0.223)	(0.108)
wL0	-0.793*	-0.718*	-1.344***
	(0.436)	(0.403)	(0.213)
Constant	5.625***	5.179***	7.128***
	(0.0736)	(0.0848)	(0.0370)
Observations	12,216	7,140	12,538
R-squared	0.551	0.438	0.699
Year FE	Yes	Yes	Yes

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

References

Aragón, F. M. & Rud, J. P. (2015), Polluting Industries and Agricultural Productivity: Evidence from Mining in Ghana. *The Economic Journal*. doi: 10.1111/ecoj.12244

Aragona, F. M., Chuhan-Pole, P., & Land, B. C. (2015). The Local Economic Impacts of Resource Abundance: What Have We Learned?. World Bank Policy Research Working Paper, (7263).

Bhattacharyya, S., & Williamson, J. G. (2015). Distributional Consequences of Commodity Price Shocks: Australia Over a Century. *Review of Income and Wealth*. DOI: 10.1111/roiw.12167

Chuhan-Pole, P., Dabalen, A. L., Kotsadam, A., Sanoh, A., & Tolonen, A. (2015). The Local Socioeconomic Effects of Gold Mining: Evidence from Ghana, The World Bank, Policy Research Working Paper 7250

Cust, J., & Poelhekke, S. (2015). The Local Economic Impacts of Natural Resource Extraction. *Annual Review of Resource Economics*, 7, 251 – 268

Firpo, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, *75*(1), 259-276.

Fleming, D. A., & Measham, T. G. (2015). Income Inequality across Australian Regions during the Mining Boom: 2001–11. *Australian Geographer*, 46(2), 203-216.

Frölich, M., & Melly, B. (2010). Estimation of quantile treatment effects with Stata. *Stata Journal*, *10*(3), 423.

Goderis, B., & Malone, S. W. (2011). Natural Resource Booms and Inequality: Theory and Evidence. *The Scandinavian Journal of Economics*,*113*(2), 388-417.

Hajkowicz, S. A., Heyenga, S., & Moffat, K. (2011). The relationship between mining and socio-economic well-being in Australia's regions. *Resources Policy*, *36*(1), 30-38.

Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, 46, 931 – 959

Heckman, J. J. (1979). Sample selection bias as a specification error, *Econometrica,* 47, 153 – 161

Hirschman, A. O. (1958). *The Strategy of Economic Development*. Yale University Press.

Kitula, A. G. N. (2006). The environmental and socio-economic impacts of mining on local livelihoods in Tanzania: A case study of Geita District. *Journal of Cleaner Production*, *14*(3), 405-414.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica:* 46, 33-50.

Kotey, B., & Rolfe, J. (2014). Demographic and economic impact of mining on remote communities in Australia. *Resources Policy*, *42*, 65-72.

Kotsadam, A. & Tolonen, A. (2016), African Mining, Gender and Local Employment. *World Development*, 83, 325 - 339

Lay, J., Thiele, R., & Wiebelt, M. (2008). Resource booms, inequality, and poverty: The case of gas in Bolivia. *Review of Income and Wealth*, *54*(3), 407-437.

Lippert, A. (2014). Spill-Overs of a Resource Boom: Evidence from Zambian Copper Mines (No. 131). Oxford Centre for the Analysis of Resource Rich Economies, University of Oxford.

Loayza, N., Mier y Teran, A., & Rigolini, J. (2013). Poverty, inequality, and the local natural resource curse. World Bank Policy Research Working Paper, (6366).

Loayza, N., & Rigolini, J. (2015). The Local Impact of Mining on Poverty and Inequality: Evidence from the Commodity Boom in Peru (No. 2015-33).

Reeson, A. F., Measham, T. G., & Hosking, K. (2012). Mining activity, income inequality and gender in regional Australia*. *Australian Journal of Agricultural and Resource Economics*, *56*(2), 302-313.

Ross, M. (2007), How Can Mineral Rich States Reduce Inequality?, in J. D. Sachs, J. E. Stiglitz, and M. Humphreys (eds), *Escaping the Resource Curse,* Columbia University Press, New York

Tolonen, A. (2014). Local Industrial Shocks, Female Empowerment and Infant Health: Evidence from Africa's Gold Mining Industry. Mimeo. University of Gothenburg, Gothenburg

World Bank (2015). Socioeconomic Impact of Mining on Local Communities in Africa, Report No. ASC14621