Modern Industries, Pollution and Agricultural Productivity: Evidence from Mining in Ghana*

Fernando M. Aragón† Juan Pablo Rud‡

Draft version: February 2013

Abstract

The development of modern sectors has long been linked to the crowding out of traditional agriculture. The economic literature has focused on explanations associated with reallocation of inputs but has neglected other possible mechanisms, such as environmental pollution. To explore this issue, we examine the case of modern gold mining in Ghana and estimate an agricultural production function to tell apart mechanisms. Consistent with a crowding out effect driven by pollution, we find that the expansion of mining production has reduced agricultural productivity by almost 40%, but is not associated with changes in the use or price of agricultural inputs. We provide evidence of air pollution and increased rural poverty near mining areas.

Keywords: industrialization, agriculture, natural resources, mining, pollution.

*This project was funded by the International Growth Centre, grant RA-2010-12-005. We thank Douglas Gollin, Matt Neidell, Krishna Pendakur, Francis Teal, Chris Udry and seminar participants at Calgary, Novafrica, Galway, Gothenburg, NEUDC. Godwin Awuah and Sarah Khan provided excellent research assistance. We also thank Phylomena Nyarko at Ghana Statistical Service and Henry Telly at IGC Ghana for invaluable help with this project.

†Department of Economics, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada; Tel: +1 778 782 9107; Email: faragons@sfu.ca

‡Department of Economics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, United Kingdom; Tel: +44 (0)1784 27 6392; Email: juan.rud@rhul.ac.uk

1 Introduction

The process of development is often understood as a phenomenon of structural transformation in which productivity gains shift resources towards the modern sector[1]. In particular, there is a large literature that deals with labor reallocations (Lewis, 1954; Matsuyama, 1992; Caselli and Coleman, 2001; Hansen and Prescott, 2002; Matsuyama, 2008). More recently, an emerging literature has been exploring conflicting interests between traditional agriculture and modern industries around the use of valuable resources, such as land or water (see Ghatak and Mookherjee (2013) and Keskin (2009), and references therein). In densely populated rural areas –such as vasts regions of South Asia and Sub-Saharan Africa– this competition for inputs directly impacts the main source of livelihood of farming households.

The discussion of other negative spillover effects that are independent of input use, such as environmental degradation and loss of agricultural output, is less prominent in the academic and policy debate. This dimension has been neglected despite the existing biological evidence linking pollution to reduction in crop yields (Emberson et al., 2001; Maggs et al., 1995; Marshall et al., 1997). In this paper we explore this channel by looking at the effects of modern mining activities on agricultural production. The economic literature has recently started to provide evidence of pollution's negative externalities on productive outcomes, such as labor productivity and labor supply (see Graff Zivin and Neidell (2011) and Hanna and Oliva (2011), for example). To the best of our knowledge, this paper is the first to show that this mechanism is associated with the crowding out of other productive activities, in this case through a reduction in agricultural output and productivity.

We focus on the case of gold mining in Ghana. We consider this to be an ideal testing ground to study the effect of modern industries on agriculture for a variety of reasons. First, gold mining is mostly a modern, capital-intensive, and heavily mechanized industry. Second, the industry has experienced a boom since the late 1990s –mostly driven by the expansion and opening of large-scale operations. It has become the most important extractive industry in Ghana, but has failed to keep a good environmental record.[2] Finally, gold mines mines are located in the vicinity of densely populated areas, where land is very fertile and agriculture is

[1] This could be due to push or pull factors, according to whether the productivity shocks affect the backward or the modern sector, respectively. See Matsuyama (2008) for a review.

[2] See for example Human Rights Clinic (2010), Akabzaa (2009), Aryeetey et al. (2007), and Hilson and Yakovleva (2007).

the main source of livelihood.

From a methodological standpoint, we use micro-data from repeated cross-sections to estimate an agricultural production function. The estimation of input coefficients through different specifications provide us with estimates of residual farmers' productivity. Then, we compare the evolution of total factor productivity in areas near mines to areas farther away. The main identification assumption is that the change in productivity in both areas would be similar in the absence of mining production. This allows us to isolate changes in agricultural output induced by input adjustments from those produced by pollution, that would affect residual productivity. To implement this approach, we use a household survey available for 1998/99 and 2005, and detailed information on the geographical location of gold mines and households. We also allow for treatment intensity to vary across mines, by using cumulative production, under the presumption that pollution emissions increase with production.

A non-trivial empirical challenge is the endogeneity of input use. This problem has long been recognized in the empirical literature on production functions (Blundell and Bond, 2000; Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2006). We are, however, unable to implement the standard solutions due to the lack of panel data. Instead, we address this issue controlling for farmer's observable characteristics and district fixed effects. We complement this strategy with an instrumental variables approach. We first show that, in the presence of imperfect input markets, endowments are a good predictor of input use. Consequently, we use farmers' input endowments, such as land holdings and household size, as instruments. The validity of the exclusion restriction might be, however, questioned. To address this concern, we use a partial identification approach proposed by Nevo and Rosen (2012) that allows for some correlation between the (imperfect) instruments and the error term. The validity of this method relies on two assumptions: (1) the instrument and the endogenous variables have the same direction of correlation with the error term, and (2) the instrument is less correlated to the error term than the endogenous variable. These assumptions are weaker than the exclusion restriction required in a standard IV, and, as we discuss below, are more likely to be met in the case we study.

We find evidence of a significant reduction in agricultural productivity. Our estimates suggest that an increase of one standard deviation in gold production is associated with a 30 percent decline in productivity in areas closer to mines. Given the increase in mining activity between 1998/99 and 2005, this implies that the average productivity in mining areas decreased 40 percent relative to areas farther away. The negative effects decline with distance and extend to areas within 20 km from mine sites. On the contrary, we find no evidence of change in input use or prices.

We interpret these results as evidence that pollution from mining activities has induced a reduction in agricultural productivity. To further explore this interpretation, we would need measures of water and air pollutants. These data, however, are unavailable in the Ghanaian case. Instead, we rely on a novel approach using satellite imagery to obtain local measures of nitrogen dioxide (NO_2), a key indicator of air pollution. We find that concentrations of NO_2 are higher in mining areas and decline with distance, in a way that parallels the reduction in agricultural productivity.

Our second set of results move beyond agricultural productivity and focus instead on local living standards. This is a natural extension given the importance of agriculture in the local economy. We find that rural poverty in mining areas shows a relative increase of almost 18%. The effects are present not only among agricultural producers, but extend to other residents in rural areas.

This paper contributes to the economic literature studying the effect of environmental degradation on living standards. This literature has focused mostly on examining the effect of pollution on health outcomes. For example, Chay and Greenstone (2003) find that reduction in air pollution, associated with an economic slump in early 1980s in the US, has reduced infant mortality. Currie et al. (2009) use U.S. school level data and find that air pollution increases school absence, a proxy for worse children health. In the context of developing countries, Jayachandran (2009) shows that exposure to pre-natal air pollution generated by wildfires in Indonesia in 1997 has increased child mortality, while Ebenstein (2012) links water pollution to incidence of cancer in China. In contrast, Greenstone and Hanna (2011) find that air regulation in India were effective on reducing air pollution, but did not have significant knock-on effects on infant mortality.

Others have explored the long-term effects of environmental disasters such as soil erosion (Hornbeck, 2012) and climate change (Dell et al., 2008; Guiteras, 2009). Recent papers have also started to explore the link between pollution, workers' health, and labor market outcomes. Hanna and Oliva (2011) use the closure of a refinery in Mexico as a natural experiment and document an increase in labor supply associated to reductions in air pollution in the vicinity of the emissions source. In a closely related paper, Graff Zivin and Neidell (2011) find a negative effect of air pollution on productivity of piece-rate farm workers in California's central valley. They, however, cannot estimate the effect of pollution on total factor productivity that may occur, for instance, if land becomes less productive or if crop yields decline. Our approach accounts for that and find much larger effects. The effects we uncover are closer in magnitude to the reduction in crop yields due to pollution documented in the natural sciences literature. Our paper contributes to this literature by documenting another, non-health related, channel through which pollution may affect living standards in rural setups: reduction in agricultural productivity and the subsequent increase in poverty.

This paper also contributes to the literature studying the effect of natural resources on development. Using country level data, this literature finds that resource abundance may hinder economic performance, specially in the presence of bad institutions (Sachs and Warner, 1995; Sachs and Warner, 2001; Mehlum et al., 2006). Departing from these cross-country comparisons, a growing literature is exploiting within-country variation to study other complementary channels which may be more relevant at local level. For example using the Brazilian case, Caselli and Michaels (2009) find that the revenue windfall from oil wells has not improved local income. In the same setup, Brollo et al. (2010) document a political resource curse: the revenue windfall has increased corruption and deteriorated political selection. Vicente (2010) also finds an increase in corruption in Sao Tome and Principe in anticipation to oil production. On a more positive side, Aragon and Rud (2013) document how the expansion of a mine's backward linkages can improve the income of local populations. Our results highlight the importance of considering potential loss of agricultural productivity and rural income as part of the social costs of extractive industries. So far, this dimension is absent in the policy debate. Instead, both environmental regulators and opponents of the industry have focused mostly on other aspects such as risk of environmental degradation, health hazards, and social change. This omission may overestimate the contribution of extractive industries to local economies and lead to insufficient compensation and mitigation policies.

The next section provides an overview of mining in Ghana and discusses the link between mining, pollution and agricultural productivity. Section 3 describes the empirical strategy and the data. Section 4 presents the main results. Section 5 concludes.

2 Background

Our empirical analysis uses the case of gold mining in Ghana. Gold is the most important export product in Ghana, ahead of more traditional commodities such as cocoa, diamond, manganese and bauxite, and a major source of fiscal revenue and foreign direct investment. Since the late 1990s, gold production in Ghana has increased significantly. For example, in the period 1988-1997, annual production was 32.1 metric tons (MT)[3] In contrast, between 1998-2004, annual production increased to 65.2 MT due to the expansion of existing operations and opening of new mines (see Table 1).

Table 1: Average annual production of gold by mine, in MT

		Average annual production		
Mine name	Type	1988-1997	1998-2004	Diff.
Bibiani	open pit	0.0	6.8	6.80
Bogoso/Prestea	open pit, underground and tailings	3.2	4.4	1.19
Central Ashanti	open pit	0.5	0.6	0.08
Damang	open pit	0.0	9.4	9.41
Dunkwa placer	placer	0.1	0.0	-0.12
Essase placer	placer	0.3	1.4	1.08
Iduapriem/Teberebie	open pit	6.2	7.5	1.25
Konong/Obenamasi	open pit	0.1	0.0	-0.15
Obotan	open pit	0.2	2.5	2.25
Obuasi	open pit, and underground	20.4	18.5	-1.89
Tarkwa	open pit, and underground	0.9	12.9	12.00
Wassa	open pit	0.0	1.2	1.17
TOTAL		32.1	65.2	33.08

Source: U.S. Geological Service, *The Mineral Industry of Ghana*, several years.
Note: the first year of available data is 1988. MT=Metric tonne.

Most of the gold (around 97%) is produced by modern, large-scale mines.[4] These mines,

[3] These figures include only production from large mines studied in this paper.

[4] The rest is produced by small artisanal mines, and informal miners also called *galamseys*. Both share similar labor-intensive, small-scale technology and are usually own by locals.

similar to other modern mines in the world, are capital intensive, highly mechanized operations. They are located in rural areas, amidst fertile agricultural land, and have little interaction with local economies: they hire few local workers, buy few local products, their profits are not distributed among local residents, and only a small fraction of the fiscal revenue is allocated to local authorities (Aryeetey et al., 2007). More importantly, large-scale mines, as other modern industries, have the potential to pollute the environment and affect quality of soil, water and air.

These features of modern mining provide an ideal setup to study how the expansion of a modern sector (mining) can crowd out traditional economic activities, such as agriculture. The economic literature has focused mostly on the channel of input competition: modern industries may displace traditional activities by competing for inputs such as labor (e.g. Lewis (1954)), land (e.g. Ghatak and Mookherjee (2013)), or water (e.g. Keskin (2009)).

In this paper we explore an alternative channel: the posible negative effect of environmental pollution on agricultural productivity, i.e., a reduction in output controlling for input use. This channel has been disregarded in the economic literature even though it has been explored by other disciplines such as natural and environmental sciences. These studies document the effect of (mostly) airborne pollutants on crop yields in controlled environments (Emberson et al., 2001; Maggs et al., 1995; Marshall et al., 1997).[5] They find drastic reductions in yields of main crops -e.g. rice, wheat, and beans- coming from the exposure to air pollutants associated to the burning of fossil fuels, such as nitrogen oxides and ozone.[6] Depending of the type of crop, the yield reductions can be as high as 30 to 60%.

There are various channels through which pollution can affect agricultural productivity. First, it can deteriorate health of crops either directly (e.g. leaf tissue injury or plant growth) or indirectly (e.g. reducing resistance to pests and diseases). Second, it can affect quality of inputs (such as land or labor). As a result, in the presence of pollution, agricultural product would fall even if there is no change in the quantity of inputs used.

It is important to note that mining activities produce several air pollutants such as nitrogen

[5]Most of the available evidence comes from experiments in developed countries. The above mentioned studies, however, document the effect of pollution in developing countries such as India, Pakistan and Mexico.

[6]Tropospheric ozone is generated at low altitude by a combination of nitrogen oxides, hydrocarbons and sunlight, and can be spread to ground level several kilometers around polluting sources. In contrast, the ozone layer is located in the stratosphere and plays a vital role filtering ultraviolet rays.

oxides, sulphur oxides and particulate matter.[7] These air pollutants are contributors to smog and acid rain. Most importantly for the external validity of our exercise, these emissions are akin to any fuel-intensive technology and similar to the ones associated to industrial sites, power plants, and motor vehicles. The main direct sources of air emissions are diesel engines for haulage, drilling, heating and cooling, among others. Additionally, the process of blasting, crushing and fragmenting the rocks, followed by smelting and refining generate substantial aerial emissions in large-scale, open pit, mining.

The potential harmful effect of pollution on agriculture from mining activities has been raised by environmental agencies. For example, Environment Canada states that "Mining activity may also contaminate terrestrial plants. Metals may be transported into terrestrial ecosystems adjacent to mine sites as a result of releases of airborne particulate matter and seepage of groundwater or surface water. In some cases, the uptake of contaminants from the soil in mining areas can lead to stressed vegetation. In such cases, the vegetation could be stunted or dwarfed." (Environment Canada, 2009, p. 39)

In the case of Ghana, there is substantial evidence, ranging from anecdotal to scientific, that gold mining is associated with high levels of pollution and loss of agricultural livelihoods[8]. (Human Rights Clinic, 2010; Akabzaa, 2009; Aryeetey et al., 2007; Hilson and Yakovleva, 2007) Most studies focus on gold mining areas in the Western Region such as Tarkwa, Obuasi, Wassa West and Prestea.

Armah et al. (2010) and Akabzaa and Darimani (2001) document heavy metal pollution in surface and groundwater near Tarkwa. The levels of pollutants decrease with distance to mining sites. The authors also document levels of particulate matter, an air pollutant, near or above international admissible levels. Similarly, Tetteh et al. (2010) find high levels of mercury and zinc content in the topsoil of towns in Wassa West. The levels of concentration decrease with distance to mining sites, and extend beyond mining areas, probably due to the aerial dispersion of metals from mining areas [9].

[7] This is in addition to other industry-specific pollutants such as cyanide, heavy metals, or acid mine drainage.

[8] Reports also suggest an increase in social conflict and human rights abuse in mining areas.

[9] Other papers, such as Serfor-Armah et al. (2006) and WACAM (2010) also report measures of water contamination near mining sites.

3 Methods

3.1 A consumer-producer household

In this section we lay down a simple analytical framework based on the standard model of consumer-producer decision (see Bardhan and Udry (2012) for a review). This framework has been used to analyze farmer's decisions, and allows us to study how the expansion of mining could affect input use and agricultural output, and guide the empirical strategy.

We assume that households (farmers) are both consumers and producers of an agricultural good with price $p = 1$. Households have an idiosyncratic productivity A and use labor (L) and land (M) to produce the agricultural good $Q = F(A, L, M)$, where F is a well-behaved production function.

Households have endowments of labor and land (E^L, E^M). They can use these endowments as inputs in their farms (L^f, M^f) or sell them in local input markets (L^s, M^s) at prices w and r, respectively. As producers, households can also buy additional labor and land (L^b, M^b).

Households maximize utility $U(c)$ over consumption c, subject to the endowment constraints and agricultural technology. In particular, the household's problem becomes:

$$
\begin{aligned}
& \max U(c) \text{ subject to} \\
c &\leq F(A, L, M) - w(L^b - L^s) - r(M^b - M^s) \\
L &= E^L + L^b - L^s \\
M &= E^M + M^b - M^s.
\end{aligned}
$$

We assume households are heterogeneous in their access to markets for inputs. In particular, there are two types of farmers: unconstrained farmers, who operate as in perfectly competitive input markets, and fully-constrained farmers, who cannot buy nor sell inputs.[10] The assumption of imperfect input markets is reasonable in the context of weak property rights of rural Ghana. Besley (1995), for example, documents the co-existence of traditional and modern property right systems in West Ghana. Some farmers have limited rights to transfer property of land, and in many cases require approval from the community while others do not face this constraint. Botchway (1998) also discusses the customary framework that rules the right to trade land in

[10] Results would not change qualitatively if we allow for partially constrained farmers.

Ghana. Similar arguments can be made about labor markets (see Bardhan and Udry (2012)).[11]

In the case of unconstrained farmers, the maximisation problem follows the separation property: the household chooses the optimal amount of inputs to maximise profits and, separately, chooses consumption levels, given the optimal profit. From standard procedures, the optimal levels of inputs and output, $L^*(A, w, r)$, $M^*(A, w, r)$ and $Q^*(A, w, r)$, depend only on productivity and input prices. Consumption depends on these factors plus on the level of the initial endowments.

In the case of fully-constrained farmers, i.e., unable to sell or buy inputs, the optimal input decisions are shaped by their endowments. Since the opportunity cost of inputs is zero, they will use all their available land and labor. For this subset of farmers, input use will not depend on productivity A nor on input prices, but instead will be equal to a household's endowment. In this case, consumption is a function of endowment levels and A.

At this point it is important to note that, for our purposes, input market imperfections simply capture the proportion of constrained farmers. The larger this proportion, the greater the correlation between input use and endowments.

In this framework, we can now introduce two possible channels for mining to affect agricultural output, and households' consumption. First, mines could increase demand for local inputs (input competition). This may lead to increase in w and r and, through that channel, reduce input use and agricultural output. Similar effects would occur if mines reduce supply of inputs due to land grabbings or population displacement, for example. Note, however, that the effect on consumption depends on the relative size of endowments. If endowments are small, so that a household is a net purchaser of inputs, then the effect would be negative. This mechanism is similar in flavor to the Dutch disease and has been favored as an explanation for the perceived reduction in agricultural activity, and increase in poverty, in mining areas (Akabzaa, 2009; Aryeetey et al., 2007).[12].

Second, mining-related pollution may affect crop yields, as discussed in Section 3.1. This would imply a reduction in output even if input use remains unchanged. In terms of the model,

[11]Data also show that inputs markets are thin: in the area of study, around 8% of available land is rented,and only 1.4% of the total farm labor (in number of hours) is hired.

[12]For example, Duncan et al. (2009) suggests a reduction of around 15% i agricultural land use associated with the expansion of mining in the Bogoso-Prestea area. The conflict over resources seems to have exacerbated due to weak property rights (i.e., customary property rights) and poor compensation schemes for displaced farmers (Human Rights Clinic, 2010).

this represents a drop in productivity, A. This would, unambiguously, have a negative effect on agricultural output and household's consumption. Note that, if markets are flexible enough so that input prices reflect factors' marginal productivity, then mining would also reduce w or r.

This simple framework highlights three issues relevant for the empirical analysis:

1. If the main channel is through pollution, then mining would: (i) reduce agricultural output and productivity, A, (ii) depending of the flexibility of input markets, decrease input use and prices, and (iii) have a negative effect on farmer's consumption.

2. If the main channel is through input competition, then mining would: (i) reduce agricultural output, but have no effect on A, (ii) increase input prices and reduce input use, and (iii) depending of the relative size of endowments, decrease or increase farmers' consumption.

3. In the presence of imperfect input markets, household endowments are a good predictor of input use.

3.2 Empirical implementation

The aim of the empirical analysis is to explore the importance of mining-related pollution on agricultural activity. To do so, our main approach is to estimate the production function, i.e., output conditional on input, and evaluate the effect of mining on residual productivity, A. We complement this approach by also studying the effect of mining on input use, prices and poverty. As previously mentioned, the effect of mining on these outcomes can also be informative of the main mechanisms at play.

We start by assuming the following agricultural production function:[13]

$$Y_{ivt} = A_{ivt} M_{it}^{\alpha} L_{it}^{\beta} e^{\epsilon_{it}}, \tag{1}$$

where Y is actual output, A is total factor productivity, M and L are land and labor, and ϵ_{it} captures unanticipated shocks and is, by definition, uncorrelated to input decisions. All these variables vary for farmer i in locality v at time t.

[13] We assume a Cobb-Douglas technology for simplicity. In the empirical section, we check the robustness of the results to alternative specifications such as translog and CES.

We assume that A is composed of three factors: farmers' heterogeneity (η_i), time-invariant local economic and environmental conditions (ρ_v) and time-varying factors, potentially related to the presence of local mining activity (S_{vt}). In particular, $A_{ivt} = \exp(\eta_i + \rho_v + \gamma S_{vt})$. Note that if mining affects input availability or prices (input competition channel), it will change input use but would not affect productivity A so $\gamma = 0$. In contrast, if the pollution mechanism is at play, we should observe $\gamma < 0$.

We can anticipate two main empirical challenges. The first one is related to the fact that mining and non-mining areas may have systematic differences in productivity. This omitted variable problem may lead to endogeneity issues when estimating the coefficients of interest. To address this issue, we exploit the time variation in the repeated cross section to compare the evolution of productivity in mining areas relative to non-mining areas. In particular, we define S_{vt} as the cumulative gold production near a mining area.[14] To control for differences in initial cumulative production by mine and overall changes in productivity, we also include dummies of proximity to each mine ($mine_v$) and time fixed effects (ψ_t). This approach is basically a difference in difference with a continuous treatment. In this case, proximity to a mine defines the treated and control group, while the intensity of the treatment is the amount of gold produced in a nearby mine.[15] The validity of this approach relies on the assumption that the evolution of productivity in both areas would have been similar in the absence of mining.[16]

The second problem arises because, for unconstrained farmers, both output and choice of inputs are affected by productivity, and hence are simultaneously determined. Thus, unobserved heterogeneity in A would go into the error term and create an endogeneity problem in the estimation of the input coefficients.

We address these concern in several ways. First, we proxy for farmer heterogeneity η_i using farmer observable characteristics and, in addition to $mine_v$ we also include district fixed effects

[14]In the baseline specification, we define a mining area as localities within 20 km of a mine. For those areas, S_{vt} is equal to gold production in nearby mines from 1988 to $t-1$. For areas farther than 20 km, $S_{vt} = 0$.

[15]We also use a simpler specification replacing S_{vt} by $(mining\ area_v) \times GLSS5_t$ where $mining\ area_v$ is an indicator of being close to any mine and T_t is a time trend. The results using this discrete treatment are, however, similar (see Table 11 in the Appendix).

[16]In the Appendix, we explore the evolution of average agricultural output in areas closer and farther from mines for three years with available data: 1988, 1998 and 2005. Figure 6 shows that the evolution of output is remarkably similar in the first period (1988-1998), when gold production is relatively low, but there is a trend change in mining areas in the period when gold production increases (1998-2005). Table 10 formally tests the similarity of trends, and subsequent change, by regressing agricultural output on $(mining\ area_v) \times T_t$ for both periods. Note that the similarity of trends prior to the expansion of mining is a necessary, though not sufficient, condition for the identification assumption to be valid.

that would capture differences in average product due to local characteristics[17]. With these modifications, and taking logs, the model we estimate becomes:

$$y_{ivdt} = \alpha m_{it} + \beta l_{it} + \gamma S_{vt} + \phi Z_i + \delta_d + \psi_t + \theta mine_v + \xi_{ivt}, \tag{2}$$

where y, l and m represent the logs of observed output, labor and land, respectively. Z_i is a set of farmer's controls, and S_{vt} is the cumulative gold production in the proximity of a locality. δ_d and ψ_t represent district and time fixed effects, while $mine_v$ is a set of indicators of proximity to each mine. ξ_{ivt} is an error term that includes ϵ_{it} and the unaccounted heterogeneity of η_i and ρ_v.

Under the assumption that use of inputs is uncorrelated to residual unobserved heterogeneity, ξ_{ivt}, we can estimate the parameters of (2) using an OLS regression. This assumption would be satisfied if farmer heterogeneity is fully captured by the controls included in the regression or if all farmers are fully constrained, i.e., if input use is unaffected by A.

Second, we relax the previous identification assumption and exploit the presence of some constrained farmers. In particular, we estimate a standard IV model using endowments as instruments for input use. Recall from the model that the higher the fraction of constrained households, the greater the correlation between input use and household endowments. This approach would be valid if the correlation is strong enough and if endowments affect output only through its effect on input use, i.e., endowments are not conditionally correlated to unobserved heterogeneity, ξ_{ivt}.[18].

Finally, we consider the possibility that endowments are correlated to ξ_{ivt}.[19] This would invalidate the exclusion restriction of the IV strategy. We can make, however, further progress by using a partial identification strategy proposed by Nevo and Rosen (2012). This methodology uses imperfect instrumental variables (IIV) to identify parameter bounds.[20] Nevo and Rosen (2012) show that if (i) the correlation between the instrument and the error term has the same sign as the correlation between the endogenous variable and the error term, and (ii) the instru-

[17] Districts are larger geographical areas than localities v. We cannot use locality fixed effects because we use repeated cross-section data and not all localities are in all samples.

[18] The interpretation of this IV strategy would be as a local average treatment effect, since the coefficients would be identified from constrained farmers only.

[19] This could happen, for example, if more productive farmers have systematically larger landholdings or household size (measures of input endowments).

[20] In contrast, the standard IV approach focuses on point identification.

ment is *less* correlated to the error than the endogenous variable, then it is possible to derive analytical bounds for the parameters. The parameter set could be a two- or one-sided bound depending on the observable correlation between endogenous variables and instruments[21]. These (set) identification assumptions are weaker than the exogeneity assumption in the standard IV and OLS approaches. Because, we are not interested in the point estimates of the input coefficients *per se*, we can impose weaker exogeneity assumptions in order to obtain a more robust inference of the evolution of residual productivity, i.e., not driven by the assumed exclusion restriction.

3.3 Data

Our main results use household data from the rounds 4 and 5 of the Ghana Living Standards Survey (GLSS) These surveys were collected by the Ghana Statistical Service (GSS) in 1998/99 and 2005, respectively. [22]

The survey contains several levels of geographical information of the interviewees. The higher levels are district and region. The district is the lower sub-national administrative jurisdiction, while the region is the highest.[23] The survey also distinguishes between urban and rural areas, as well as ecological zones (coastal, savannah and forest). The finer level is the enumeration area, which roughly corresponds to villages (in rural areas) and neighborhoods (in urban areas). For each enumeration area we obtain its geographical coordinates from the GSS.[24]

We are mainly interested on two set of variables: measures of mining activity, and measures of agricultural inputs and output.

Mining activity Our main measure of mining activity is the cumulative production of gold in the proximity of a household, the empirical counterpart of S_{vt}. To construct this variable, we

[21] In particular, denoting X as the endogenous variable, Z as the imperfect instrument, and W other additional regressors, there is a two-sided bound if, in addition to the (set) identification assumptions, $(\sigma_{\tilde{x}x}\sigma_z - \sigma_x\sigma_{\tilde{x}z})\sigma_{\tilde{x}z} < 0$, where $\tilde{x}$ is the projection of X on W. In the complementary case, there is a one-sided bound. In the empirical section we do check that this expression has a negative value. We refer the reader to Nevo and Rosen (2012) for a detailed exposition of the estimation method.

[22] We also use the GLSS 2, taken in 1988, for evaluating pre-trends in agricultural output between mining and non-mining areas. We do not use this data, however, in the estimation of the production function since it does not contain comparable information on input use. In addition, we do not use the GLSS 3 (1993/94) because there is not available information on the geographical location of the interviewees.

[23] In 2005, there were 10 regions and 138 districts.

[24] The GSS does not have location of enumeration areas for the GLSS 2. In this case, we extracted the location using printed maps of enumeration areas in previous survey reports.

first identify mines active during the period 1988 to 2004, and aggregate the annual production of each mine since 1988 to a year prior to the survey. This information comes from reports prepared by the U.S. Geological Service.[25] Second, we obtain geographical coordinates of each mine site.[26] Using a geographical information system (ArcGIS), we identify the enumeration areas within different distance brackets of each mine site. For reasons that will be clearer later, we define the enumeration areas within 20 km of mine sites as mining areas. Finally, we assign the cumulative production of each mine to its surrounding mining area, and zero for areas farther away.

Figure 1 displays a map of Ghana with the location of active gold mines between 1988 and 2004. Note that all mines are located in three regions: Western, Ashanti and Central. In the empirical section, we restrict the sample to these regions.[27] Figure 2 zooms in these three regions and depicts the enumeration areas and a buffer of 20 km around each mine. The areas within each buffer correspond to the mining areas (treated group), while the rest correspond to the non-mining areas (comparison group).

We restrict attention to medium and large-scale gold mines. We do not consider artisanal and informal gold mines for two reasons. First, the magnitude of their operations is relatively small (they represent around 4% of total gold production). Second, there is no information on their location, though anecdotal evidence suggests they are located in the vicinity of established mines. For similar reasons, we do not consider mines of other minerals (such as diamonds, bauxite and manganese). These minerals are less important than gold in Ghana's mining output. Moreover, their mine sites are concentrated in fewer locations that overlap with existing gold operations. For example bauxite and diamonds are mined in Awaso (south of Bibiani gold mine), while manganese is extracted at the Nsuta-Wassaw mine near Tarkwa. Note that the omission of these other mines would, if anything, attenuate the estimates of the effect of large scale gold mining.

[25]See the annual editions of *The Mineral Industry in Ghana* from 1994 to 2004 available at `http://minerals.usgs.gov/minerals/pubs/country/africa.html`. There is, however, no information available for the period prior to 1988.

[26]This information comes from industry reports prepared by Infomine, see `http://www.infomine.com/minesite/`.

[27]The results, however, are robust to using a broader sample.

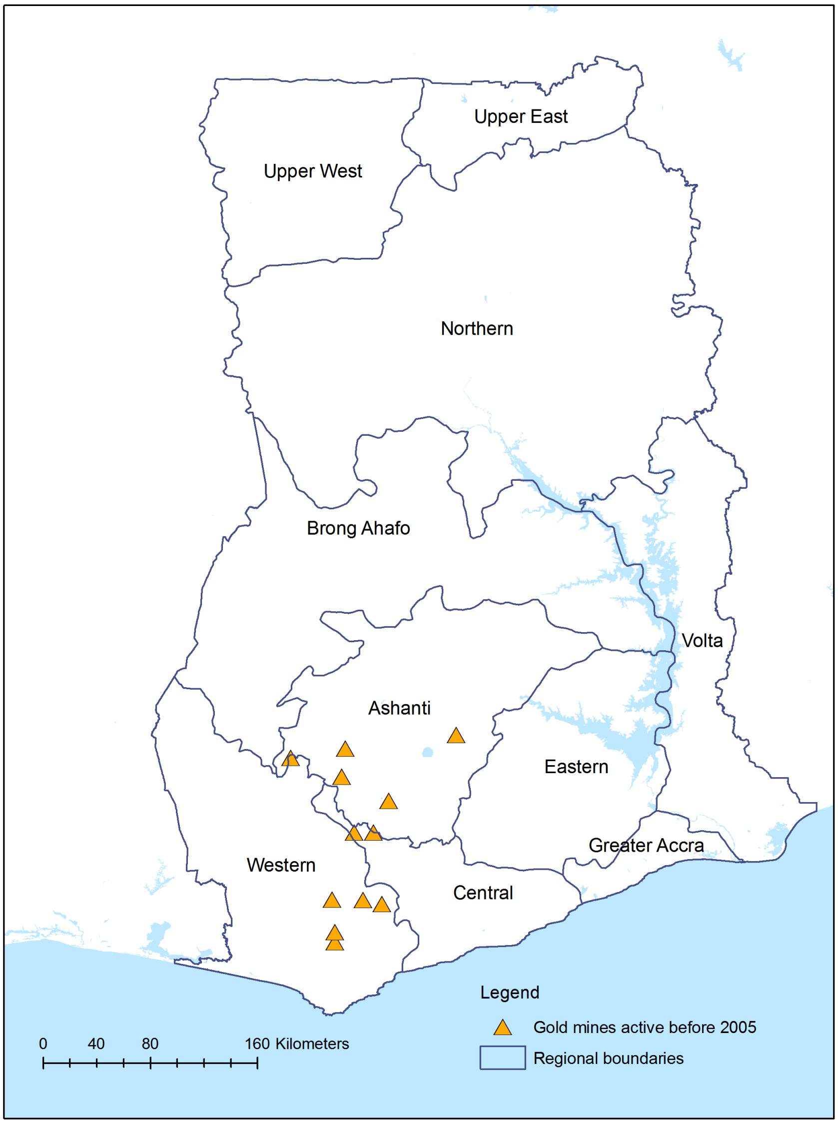

Figure 1: Location of active gold mines

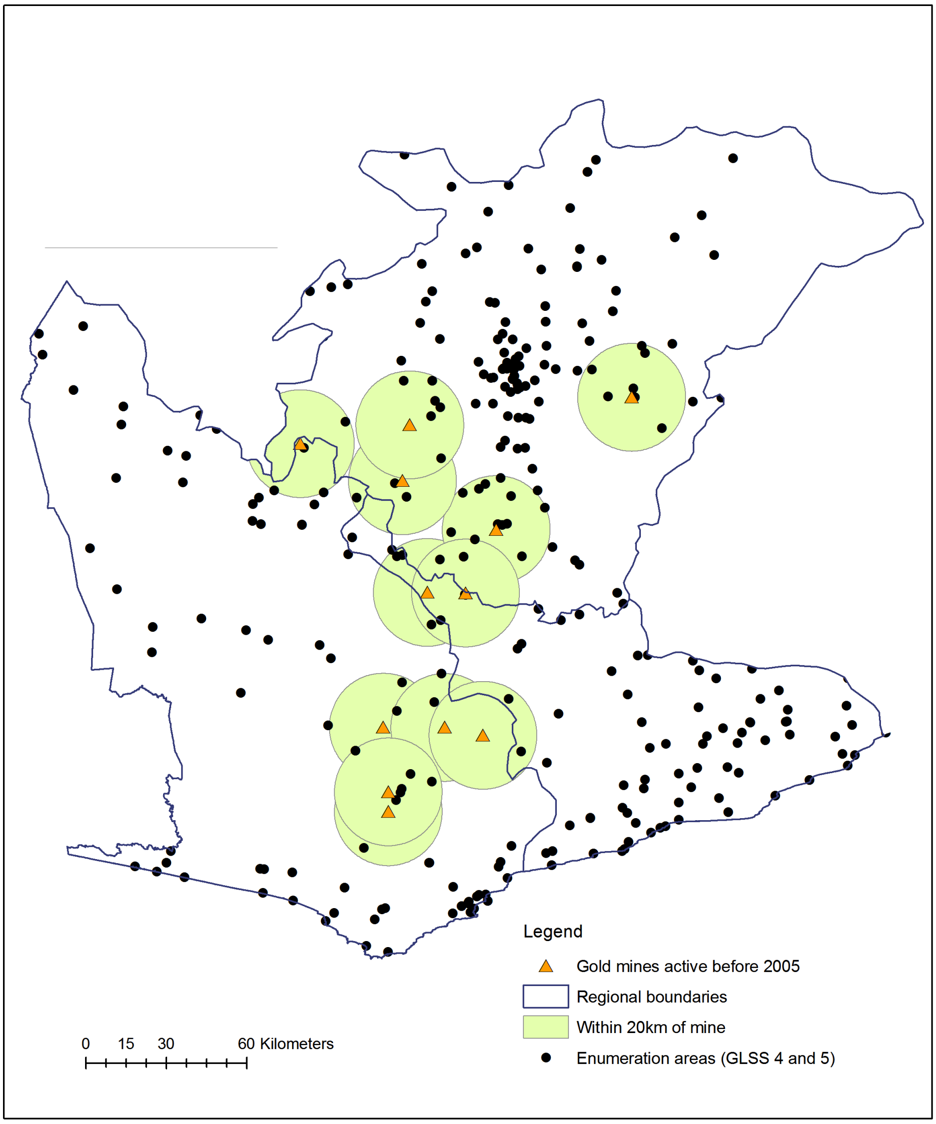

Figure 2: Area of study and enumeration areas

Agricultural output and inputs To measure agricultural output Y, we first obtain an estimate of the nominal value of agricultural output. To do so, we add the reported value of annual production of main crops. These category includes cash crops, staple grains and field crops such as cocoa, maize, coffee, rice, sorghum, sugar cane, beans, peanuts, etc. Then, we divide the nominal value of agricultural output by an index of agricultural prices.[28] This price index uses data from agricultural producers and varies by region and by mining and non-mining areas.[29]

We also construct estimates of the two most important agricultural inputs: land and labor. The measure of land simply adds the area of plots cultivated with major crops in the previous 12 months. To measure labor, we add the number of hired worker-days to the number of days each household member spends working in the household farm. Finally, we measure land endowment as the area of the land owned by the farmer, while the labor endowment is the number of equivalent adults in the household.

The resulting dataset contains information on agricultural inputs and output for 1,627 farmers in years 1998/99 and 2005. The farmers are located in 42 districts in three regions of south west Ghana: Western, Ashanti and Central. Table 2 presents the mean of the main variables.

[28] The results are similar using a consumer price index reported by the GSS, which varies by ecological zone and by urban and rural areas (see Table 12 in the Appendix).

[29] In particular, we obtain data from individual farmers on unit values of cocoa and maize, the two main crops in the area of study, and their relative share in the value of agricultural output in 1998. Then, we take the average of prices and weights by region and by mining and non-mining area, i.e., six different values every survey, and calculate a Laspeyre price index.

Table 2: Mean of main variables

Variable	GLSS 4	GLSS 5
Within 20 km of mine (%)	21.6	23.9
Cumul. prod. within 20 km (00s' MT)	5.0	17.5
ln(real agricultural output)	6.5	6.5
Land (acres)	8.1	11.4
Labor (days)	350.5	364.5
Land owned (acres)	11.9	15.4
Nr adults equivalents	3.8	3.5
ln(relative land price)	14.0	14.1
ln(real wage)	8.4	8.8
Age (years)	46.9	47.5
Literate (%)	54.2	45.6
Born in village (%)	52.3	46.4
Owns a farm plot (%)	57.5	84.3
Poverty headcount (%)	29.8	19.6
Nr. Observations	713	914

Note: Means are estimated using sample weights.

4 Main results

4.1 Mining and agricultural productivity

Table 3 presents the main results. In column 1, we start by exploring the relation between agricultural output and the measure of mining activity (i.e,. the amount of cumulative production in nearby mines) without controlling for input use. Note that this relation is negative and significant. As previously discussed, this negative effect is consistent with mining affecting agriculture through pollution or input competition.

To explore the likely channels driving this relation, we proceed to estimate the agricultural production function laid out in equation (2). Column 2 provides OLS estimates, while column 3 estimates a 2SLS using input endowments (such as area of land owned and the number of adults equivalents in the household) as instruments for actual input use.[30]. As a reference, column 4 estimates the 2SLS regression using as proxy of S_{vt} the interaction between a dummy

[30]The first stage of the 2SLS reveals a positive and significant correlation between input endowments and input use. This is consistent with imperfect input markets as discussed in Section 3.1. See Table 13 in the appendix for the first stage regressions.

of proximity to a mine and a time trend, so the estimate of γ represents the average change in output, conditional on inputs, of mining areas relative to non-mining areas. All regressions include a set of farmer controls, district and time fixed effects. We also use sample weights and cluster errors at district level to account for sampling design and spatial correlation of shocks.

Both approaches suggest a large negative relation between mining and output, after controlling for input use.[31] Under the identification assumptions discussed above, we interpret this as evidence that mining has reduced agricultural productivity. This result is consistent with mining-related pollution negatively affecting agriculture.

The magnitude of the effect is relevant: an increase of one standard deviation in the measure of mining activity is associated to a reduction of almost 30% in productivity.[32] Given the increase in production between 1998 and 2005, this implies that average agricultural productivity in areas closer to mines decreased around 40% relative to areas farther away [33]. The estimated effect on productivity is large. Its magnitude, however, is consistent with the biological literature that documents reductions of 30-60% in crop yields due to air pollution (see Section 2). Moreover, it highlights the importance of pollution as a source of negative spillovers from modern industries in rural environments.

Columns 5 and 6 use the imperfect instrumental variable approach developed by Nevo and Rosen (2012). This approach uses instrumental variables that *may be correlated to the error term*. Under weaker assumptions that the standard IV approach, this methodology allow us to identify parameters bounds instead of point estimates.[34] We allow one instrument at a time to be imperfect and report the estimated bounds and confidence interval of the identified set.[35] Note that the identified parameter sets of α and β remain mostly positive, though the range is quite broad. Despite this, the estimated effect of mining on agricultural productivity (γ)

[31]The estimates of α and β, i.e., the participation of land and labor, also seem plausible. We also cannot reject the hypothesis of constant returns to scale. Using the 2SLS estimates, the p-value of the null hypothesis $\alpha + \beta = 1$ is 0.651.

[32]The mean and standard deviation of the measure of mining activity (i.e., cumulative gold production within 20 km) are 22.3 and 54.7, respectively.

[33]We obtain this figure using estimates in column 4.

[34]The key identification assumptions are that (i) the instrument and the endogenous regressor have the same direction of correlation with the error term and (ii) the instrument is less correlated to the error term than the endogenous variable.

[35]In column 5 and 6, the values of the expression $(\sigma_{\tilde{x}x}\sigma_z - \sigma_x\sigma_{\tilde{x}z})\sigma_{\tilde{x}z}$ are -0.057 and -0.225, respectively. Recall that when this expression is negative there is a two-sided bound of the parameters. The 95% confidence intervals of the identified sets are obtained by adding (subtracting) 1.96 standard deviations to the upper (lower) bounds. Nevo and Rosen (2012) obtain analytical bounds only in the case when there is one endogenous regressor with imperfect instruments. In the case of multiple endogenous variables, the parameter set can be, however, obtained by simulations. The estimates of γ in this, more flexible, case are similar (see Table 14 in the Appendix).

remains negative, with values ranging between -0.301 and -0.760.

Finally, columns 7 and 8 examine the effect of mining on crop yields. Crop yields have been used as a proxy for agricultural productivity in the empirical literature and are an output of interest by themselves (see for example Duflo and Pande (2007) and Banerjee et al. (2002)). Note that crop yields use only data on physical production and land use, so they are not affected by possible errors in measuring price deflators.

We focus on the yields of cocoa and maize, the two most important crops in south west Ghana. In both cases, we estimate an OLS regression including farmer's controls and district fixed effects, but without input use.[36] Consistent with the results on productivity, we find a negative and significant relation between mining and crops yields.

The role of distance So far, we have assumed that areas within 20 km of mines experience most of the negative effect. Implicitly, this approach assumes that the effect of mining declines with distance. To explore this issue further, we estimate equation (2) replacing S_{vt} by a linear spline of distance to a mine, $\sum_c \gamma^d(distance_v^d \times T_t)$ where $distance_v^d = 1$ if enumeration area v is in distance bracket d, and T_t is a time trend. This specification treats distance more flexibly and allow us to compare the evolution of farmers' productivity at different distance brackets from the mine relative to farmers farther way (the comparison group is farmers beyond 50 km).

Figure 3 presents the estimates of γ^d. Note that the effect of mining on productivity is (weakly) decreasing in distance. Morevoer, the loss of productivity is significant (at 10% confidence) within 20 km of mines, but becomes insignificant in farther locations. This result provides the rationale for concentrating in a 20 km buffer around mines, as in the main results.

4.2 Is this driven by pollution?

We interpret the previous findings as evidence that agricultural productivity has decreased in the vicinity of mines. We argue that a plausible channel is through the presence of mining-related pollution. As we discussed before, modern mines can pollute air with exhausts from heavy machinery and processing plants, and particulate matter from blasting. This is in addition to other industry specific pollutants such as cyanide, heavy metals and acidic discharges. Indeed,

[36] We do not control for inputs since we do not have estimates of labor use by crop. However, including total input use does not change the results.

Table 3: Mining and agricultural productivity

	ln(real agricultural ouput)						ln(yield cocoa)	ln(yield maize)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cumulative gold prod. within 20 km.	-0.438*	-0.733***	-0.727**		[-0.301, -0.675]	[-0.760, -0.684]	-1.493**	-0.689**
	(0.250)	(0.268)	(0.273)		(-0131, -0.750)	(-0.800, -0.616)	(0.618)	(0.280)
Within 20 km of mine × GLSS 5				-0.567**				
				(0.246)				
ln(land)		0.630***	0.675***	0.676***	[0,195, 0.676]	[0.743, 0.676]		
		(0.038)	(0.050)	(0.048)	(-0.027, 0.773)	(0.780, 0.615)		
ln(labor)		0.219***	0.368***	0.350***	[0.667, 0.368]	[0.121, 0.368]		
		(0.036)	(0.115)	(0.113)	(0.806, 0.308)	(-0.008, 0.590)		
Estimation	OLS	OLS	2SLS	2SLS	IIV	IIV	OLS	OLS
Farmer's controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes		Yes	Yes		
Imperfect IV for:					Land	Labor		
Observations	1,627	1,627	1,627	1,627	1,627	1,627	1,076	933
R-squared	0.235	0.472	0.462	0.441			0.283	0.298

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. The set of farmer's controls includes: household head's age, literacy, and an indicator of being born in the village; as well as an indicator of the household owning a farm plot. All regressions include district and time fixed effects. All columns, except column 4, also include dummies of proximity to each mine. Column 5 and 6 identify parameter bounds using the imperfect instrumental variable approach in Nevo and Rosen (2010). Identified parameter bounds are in brackets while the 95% confidence interval is in parenthesis. Confidence intervals are calculated adding (substracting) 1.96 standard deviations to the upper (lower) bound. Cumulative gold production is measured in hundreds of MT.

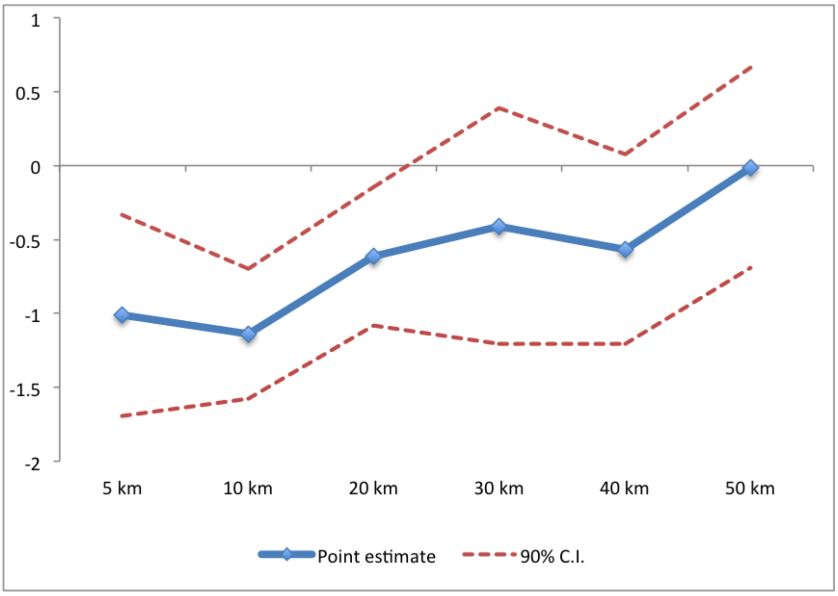

Figure 3: The effect of mining on agricultural productivity, by distance to a mine

several case studies show that water and soil in mining areas have higher than normal levels of pollutants (see Section 2).

A first indication in the previous set of results that the pollution channel might be at play is that the main specification uses a measure of cumulative production. This effectively allows for different intensities of treatment, under the presumption that the emission of pollutants is directly related to the level of production. To further explore this issue, we would need measures of environmental pollutants at local level. Then, we could examine whether mining areas are indeed more polluted. Unfortunately, this information is not available in the Ghanaian case.[37]

Instead, we rely on satellite imagery to, indirectly, assess the role of pollution. The satellite imagery is obtained from the Ozone Monitoring Instrument (OMI) available at NASA.[38] This satellite instrument provides daily measures of tropospheric air conditions since October 2004. We focus on a particular air pollutant: nitrogen dioxide (NO_2). NO_2 is a toxic gas by itself and also an important precursor of tropospheric ozone -a gas harmful to both human and crops' health. The main source of NO_2 is the combustion of hydrocarbons such as biomass burning, smelters and combustion engines.Thus, it is likely to occur near highly mechanized operations,

[37]There are, for example, air monitoring stations only in the proximity of Accra. Regarding mining areas, there are some case studies collecting measures of soil and water quality. These measures, however, are sparse, not collected systematically, and unavailable for non-mining areas. This precludes a more formal regression analysis.

[38]For additional details, see `http://aura.gsfc.nasa.gov/instruments/omi.html`. Data are available at `http://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?tree=project&project=OMI`.

such as large-scale mining.

There are three important caveats relevant for the empirical analysis. First, the satellite data reflect air conditions not only at ground level, where they can affect agriculture, but in the whole troposphere (from ground level up to 12 km).[39] Levels of tropospheric and ground level NO_2 are, however, highly correlated.[40] Thus, data from satellite imagery can still be informative of relative levels of NO_2 on the surface. Second, the data is available only at the end of the period of analysis (late 2004). For that reason we can only exploit the cross-sectional variation in air pollution. Finally, the measures of NO_2 are highly affected by atmospheric conditions such as tropical thunderstorms, cloud coverage, and rain. These disturbances are particularly important from November to March, and during the peak of the rainy season.[41] For that reason, we aggregate the daily data taking the average over the period April-June 2005. These months correspond to the beginning of the rainy season, and also to the start of the main agricultural season.

To compare the relative levels of NO_2 in mining and non-mining areas, we match the satellite data to each enumeration area and estimate the following regression:[42]

$$NO2_v = \phi_1 mine_v + \phi_2 W_v + \omega_v,$$

where $NO2_v$ is the average value of tropospheric NO_2 in enumeration area v during the period April-June 2005. $mine_v$ is a measure of mine activity such as an indicator of proximity to a mine, or the log of cumulative gold production in nearby mines; and W_v is a vector of controls variables.[43] Note that the unit of observation is the enumeration area, and that, in contrast to the baseline results, this regression exploits cross-sectional variation only.

Columns 1 and 2 in Table 4 present the empirical results using two alternative ways to measure mine activity[44]. We also replace the dummy $mine_v$ by a distance spline with breaks

[39] To obtain accurate measures at ground level, we would need to calibrate existing atmospheric models using air measures from ground-based stations. This information is, however, not available.

[40] The correlation between these two measures is typically above 0.6. OMI tropospheric measures tend, however, to underestimate ground levels of NO_2 by 15-30 % (Celarier et al., 2008).

[41] In south Ghana, the rainy season runs from early April to mid-November.

[42] The satellite data are binned to 13 km x 24 km grids. The value of NO_2 of each enumeration area corresponds to the value of NO_2 in the bin where the enumeration area lies.

[43] NO_2 is measured as 10^15 molecules per cm^2. The average NO_2 is 8.1 while its standard deviation is 1.1.

[44] We use a semi-logarithmic specification since the relation between mining activity and NO2 concentration is likely to be non-linear. (Ralf Kurtenbach and Wiesen, 2012) and (Anttila et al., 2011) for example, show that large changes in emissions (or source of emissions, such as petrol cars) are necessary to produce small changes in

at 10, 20, 30 and 40 km and plot the resulting estimates in Figure 4. Note that in this figure the comparison group is farmers beyond 40 km of a mine.

The satellite evidence suggests that mining areas have a significantly greater concentration of NO_2. Moreover, the concentration of NO_2 decreases with distance to the mine in a similar fashion as the observed decline in total factor productivity. These latest findings point out to air pollution as a plausible explanation for the decline of agricultural productivity in mining areas. This result is consistent with the biological evidence linking air pollution to reduction in crop yields and the increase in respiratory diseases that we document in Section 4.5.

Columns 3 further explores the relation between mining, air pollution and productivity. To do so, we estimate the relation between NO_2 and agricultural productivity using our measure of mine activity, i.e., cumulative gold production in nearby mines, as instruments for NO_2.[45] Since we only have measures of NO_2 for 2005, we use the sample of farmers in the GLSS 5 and thus exploit only cross sectional variation. Consistent with mining-related pollution being a possible explanation, we find a significant negative correlation between NO_2 and agricultural productivity.[46]

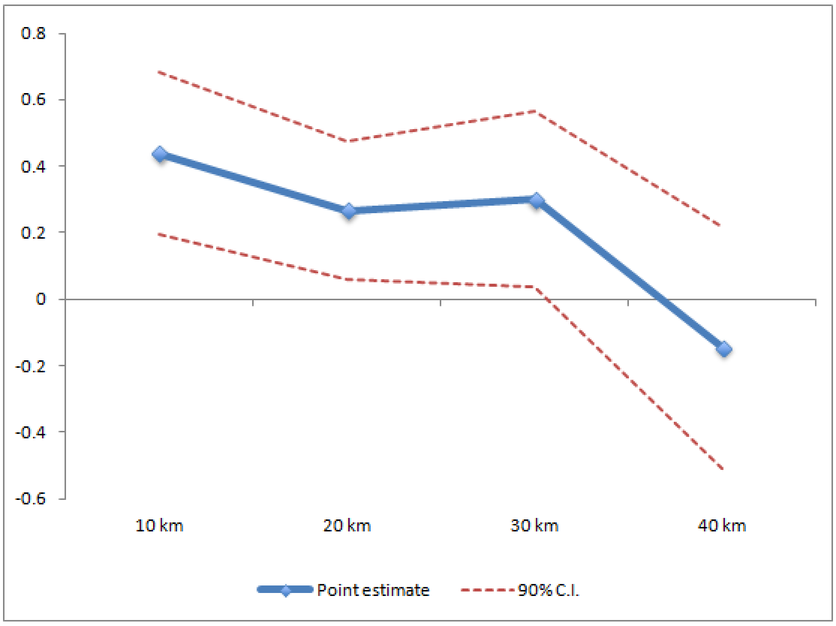

Figure 4: Increase in concentration of NO_2, by distance to a mine

NO2. We also estimate other non linear specification such as quadratic and 3rd degree polinomyal with similar results.

[45]Results are similar using an indicator of proximity to a mine, i.e., "within 20 km of mine".

[46]In the first stage the relation between NO_2 and the excluded instrument ''cumulative gold production within 20 km" is positive and significant at 5%.

Table 4: Mining and pollution

	Average NO2		ln(real agric. output)	
			Using mining as IV	Upstream vs downstream
	(1)	(2)	(3)	(4)
Within 20 km of mine	0.325***			
	(0.111)			
Ln (cumulative gold prod. within 20 km+1)		0.013**		
		(0.006)		
Average NO_2			-1.522*	
			(0.802)	
Cumulative gold prod. within 20 km				-0.794***
				(0.258)
Cumul. gold prod. within 20 km × downstream				-0.081
				(0.124)
Estimation	OLS	OLS	2SLS	OLS
Farmer's controls	No	No	Yes	Yes
Controlling for inputs	No	No	Yes	Yes
Observations	399	399	914	1,627
R-squared	0.238	0.231	0.044	0.461

Notes: Robust standard errors in parentheses. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions use data for 2005 only . Column 1 and 2 use as unit of observation the enumeration area and includes as additional controls indicators of ecological zones, urban area, and region fixed effects. Columns 3 presents 2SLS estimates of the agricultural production function using only the sample of farmers in GLSS 5. It treats "Average NO_2" as an endogenous variable and uses "ln(cumulative gold production within 20 km + 1)" as the excluded instrument. It report standard errors clustered at district level and includes the additional controls: indicators of ecological zone, urban area, region fixed effects, as well as farmer's characteristics and input use as in the baseline regression (see notes of Table 3). Column 4 replicates baseline OLS regression (column 2 in Table 3) adding an interaction term of the measure of mining activity and "*downstream*", a dummy equal to one if household is downstream of an active mine.

Finally, we explore the importance of pollutants carried by surface water. To do so, we identify areas downstream of active mine, and examine whether the negative effects of mining are stronger in this areas. Note that this is a crude way to assess exposure to pollution since some pollutants (like heavy metals and dust) can be carried by water and air, so areas upstream and downstream of mine can both be negatively affected.[47] We replicate the baseline regression including an interaction term between our measure of mining activity and a dummy ''*downstream*" equal to one if the household is located downstream of an active mine. The results, displayed in Column 4 in Table X, suggest that there is no significant difference in the effect of mining between areas downstream and upstream of a mine. Though this may be due to lack of statistical power, a conservative interpretation is that pollution of surface waters may not be driving the main results.

4.3 Competition for inputs

Mining could also affect agriculture through competition for key inputs. A first, and most obvious, way involve direct appropriation of inputs such as diversion of water sources and land grabbings.[48] A concern is that the loss in productivity simply reflects the reduction in quality of inputs associated with farmers' displacement. For example, farmers may have been relocated to less productive lands or to isolated locations.[49]

It is unlikely, however, that this factor fully accounts for the observed reduction in productivity. Population displacement, if required, is usually confined to the mine operating sites, i.e., areas containing mineral deposits, processing units and tailings. These areas comprise, at most, few kilometers around the mine site.[50] In contrast, we document drops in productivity in a much larger area, i.e., within 20 km of a mine, this represents an area of more than 1,200 km^2

[47]An alternative way to assess exposure to pollution is to use information collected by Ghana's Environmental Protection Agency (EPA). This agency collects information of environmental pollutants in some mining areas, and produces environmental assessments. This information has, however, two main limitations. First the information has been collected only since 2007, hence it may not accurately reflect the environmental conditions during the period of analysis (1998-2005). Second, there are not environmental assessments for all mines that were active before 2005, nor for non-mining areas that could be used as a control group. These issues create potentially severe measurement error, and limit the use of formal regression analysis.

[48]These phenomena are documented in the Ghanaian case and are deemed a source of conflict and increased poverty in mining areas (Duncan et al., 2009; Botchway, 1998).

[49]Note that our previous results are conditional on being a farmer, hence they underestimate the loss of agricultural output due to change of land use from agriculture to mining, or farmer's leaving the industry.

[50]For example, Bibiani mine has a license over 19 km^2; Iduapriem mine has a mining lease of 33 km^2 while Tarkwa leases cover 260 km^2. Note that not all lands in mining concessions are inhabited nor all its population is displaced.

around a mine.[51]

A second way involves the increase in price of local inputs. Mines may reduce supply of agricultural land through land grabbings, or increase demand for farming inputs such as unskilled labor. Alternatively, mines' demand for local goods and services may increase price of non-tradables (such as housing) and indirectly drive up local wages. In any case, the increase in input prices may lead to a decline in input demand, and agricultural output. This phenomena cannot be studied by equation (2) since it already controls for input use and thus it is only informative of the effect of mining on total factor productivity.

To explore this issue further, we study the effect of mining on input prices and input demand (see notes of Table 5). As measure of input prices, we use the daily agricultural wage from the GLSS community module and the price of land per acre self-reported by farmers.[52] To estimate input demands, we regress input use on measures of input prices, farmer's endowments and proxies of total factor productivity, including mine activity.

Table 5 displays the results. Columns 1 and 2 explore the effect of mining on input prices, while columns 3 and 4 estimate input demands. Note that neither input prices nor demand are affected by mining activity. These results weaken the argument that mining crowds out agriculture through increase in factor prices. Instead, they point out to direct reduction in productivity as the main driver of reduction in agricultural output.

The lack of effect of mining on input demands is surprising. Even if mining does not affect prices, with flexible inputs, we could except prices and demand to decline with productivity. The results, however, are consistent with imperfect input markets. As previously discussed, if farmers are unable to buy or sell their inputs, then their input use would simply follow their input endowment, and thus would be less responsive to prices and productivity. Indeed, the results suggest that input use is driven mostly by farmers' endowments.

4.4 Additional checks

Compositional effects We next turn our attention to changes in the composition of farmers or crops as an alternative explanation for the observed phenomena. A particular concern is that

[51] Another possibility is that the drop in productivity is driven by migrants with either lower human capital or occupying poorer lands. We discuss this alternative explanation in Section 4.4.

[52] We take the average of these variables by enumeration area, and divide them by the consumer price index to obtain relative input prices.

Table 5: Mining, input prices and input demands

	ln(relative wage) (1)	ln(relative land rent) (2)	ln(labor) (3)	ln(land) (4)
Cumulative gold prod. within 20 km.	-0.082	-0.277	-0.383	0.370
	(0.135)	(0.411)	(0.418)	(0.436)
ln(relative wage)			-0.116	0.148
			(0.175)	(0.107)
ln(relative land rent)			-0.094	-0.016
			(0.079)	(0.044)
ln(nr. adult equivalents)			0.524***	0.016
			(0.060)	(0.019)
ln(land owned)			0.142***	0.925***
			(0.030)	(0.030)
Farmer's controls	No	No	Yes	Yes
District fixed effects	No	No	Yes	Yes
Observations	194	201	1,342	1,342
R-squared	0.343	0.091	0.281	0.812

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include time fixed effects and indicators of proximity to each mine. Columns 3 and 4 also include district fixed effects, and a set of farmer's controls similar to regressions in Table 3.

the reduction in productivity is just reflecting an increase in the relative size of low productivity farmers. This is possible, for example, if high-productivity farmers are emigrating away from mining areas, or switching to non-agricultural activities. Similarly, it could reflect changes in crop composition. For example, farmers may perceive a higher risk of expropriation in the vicinity of mines and reduce the share of crops with high productivity but a long growing cycle (such as cocoa).

We examine whether mining activity is associated to several observable characteristics. As a first check, we investigate whether mining activity is associated to changes in the proportion of males in prime age (20-40 years) and people born in the same village where they reside. In the presence of migration, we could expect these indicators to change. Second, we look at measures of workers' education, such as literacy and having completed secondary school.[53] This result is informative, however, under the assumption that farming ability is positively correlated with educational attainment. This sounds a plausible assumption, given that in our baseline regression the measure of literacy is associated with an increase in agricultural product and productivity. Third, we explore the probability that a worker is engaged in agriculture (either as a producer or laborer). Finally, we examine whether farmers have changed crop composition. We focus on the share of cocoa in total agricultural value. Note that cocoa is the main cash crop, but it also has a long maturity cycle.[54] In addition, we use a Herfindhal concentration index of all main crops.

Table 6 displays the results. In all cases, there is no significant relation between mining activity and observable population characteristics. Taken together, these results weaken the argument that the reduction in productivity is driven by changes in demographics, occupational choice, farmer's ability or crop choice.

Alternative specifications In table 7, we check that our results are robust to alternative specifications. Columns 1 and 2 estimate baseline regression (2) splitting the sample between local and non-local farmers . We define farmers as local if they were born in the same village where they resides. This specification responds to concerns that the change in productivity may be driven by migrants to mining areas with lower human capital or occupying marginal,

[53] Levels of completion of primary school are high, i.e., around 88%, while literacy levels (47.7%) and secondary school completed (37.2%) show greater variation. Results hold when using data on completed primary school.

[54] Results are similar using the share of maize, the second most important crop.

Table 6: Robustness checks: compositional changes

	Male in prime age	Born in village	Literacy	Completed secondary	Works in agriculture	Share of cocoa	Crop concentration
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cumulative gold prod. within 20 km	-0.023	0.045	0.031	0.011	0.124	-0.086	0.064
	(0.017)	(0.081)	(0.029)	(0.040)	(0.089)	(0.138)	(0.060)
Sample	All individuals	People in prime age	All workers	All workers	All workers	Agricultural households	Agricultural households
Observations	22,855	5,907	9,017	9,030	8,932	1,627	1,627
R-squared	0.007	0.079	0.101	0.108	0.371	0.476	0.140

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include district and time fixed effects, indicators of proximity to each mine and indicators of ecological zone and urban area. Columns 1 to 6 are estimated using a linear probability model. Male in prime age is an indicator equal to 1 if individual is male between 20 to 40 years old. Born here is an indicator equal to 1 if individual was born in the same village where she resides. Columns 3 and 4 examine the educational attainment of workers conditional on age and age^2. Column 5 examines probability of working in agriculture conditional on worker's characteristics such as: age, age^2, religion, place of birth, literacy status, and household size. Columns 6 and 7 use same farmer's controls as the agricultural production function in Table 3.

unproductive, lands. Note that in all cases, the estimates of the effect of mining on productivity (γ) are negative and statistically significant.

Column 3 estimates a parsimonious model without farmer characteristics and district fixed effects. In contrast, column 4 saturates the baseline regression with indicators of use of other inputs (such as fertilizer, manure and improved seeds) and an array of heterogeneous trends. We include the interaction of time trends with indicators of ecological zone, region, proximity to coast and to region capitals. This specification addresses concerns that the measure of mining activity may be just picking up other confounding trends.

Column 5 performs a falsification test. To do so, we estimate the baseline regression (2) including interactions between time trends and (i) proximity to an active mine, and (ii) proximity to a future mine. This last group includes mines that started operations after 2005 or have not started production yet but are in the stage of advanced exploration or development.[55] The results show that the negative relation between mining and agricultural productivity occurs only in the proximity of mines active during the period of analysis, but not in future mining areas.

Finally, we relax the assumption of a Cobb-Douglas production. Instead, we estimate the following CES production function using non-linear least squares:

$$y_{ivt} = A_{ivt}[\eta M_{it}^{-\rho} + (1-\eta) L_{it}^{-\rho}]^{-\frac{\lambda}{\rho}},$$

where $A_{ivt} = \exp(\gamma S_{vt} + \phi Z_i + \delta_d + \psi_t + \theta mine_v)$, M and L represent land and labor use, while S_{vt} is the measure of mining activity, i.e., cumulative gold production within 20 km. The parameter of interest is γ, the effect of mining activity on total factor productivity.

Table 8 displays the results. The implicit elasticity of substitution, $\sigma = \frac{1}{1-\rho}$, is less than one, and we cannot rule out constant returns to scale ($\lambda = 1$). Similar to the baseline results, the estimate of γ is negative, suggesting that the increase in mining activity is associated to lower productivity.

[55] Note that we cannot use cumulative gold production (our preferred measure of mine activity) in this case because there is not production for future mines.

Table 7: Alternative specifications

	ln(real agricultural output)				
	Locals	Non-locals			
	(1)	(2)	(3)	(4)	(5)
Cumulative gold prod. within 20 km	-1.400*	-1.025***	-1.101***	-0.516**	
	(0.762)	(0.274)	(0.292)	(0.216)	
Within 20 km of active mine × GLSS 5					-0.777***
					(0.276)
Within 20 km of future mine × GLSS 5					0.941***
					(0.335)
ln(land)	0.614***	0.683***	0.643***	0.603***	0.629***
	(0.048)	(0.051)	(0.046)	(0.039)	(0.037)
ln(labor)	0.290***	0.176***	0.210***	0.229***	0.227***
	(0.053)	(0.044)	(0.044)	(0.035)	(0.033)
Farmer's control	Yes	Yes	No	Yes	Yes
District fixed effects	Yes	Yes	No	Yes	Yes
Other inputs	No	No	No	Yes	No
Heterogeneous trends	No	No	No	Yes	No
Observations	780	847	1,633	1,627	1,627
R-squared	0.470	0.513	0.308	0.485	0.453

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. Columns 1 and 2 split the sample between locals and non-locals where local is someone who resides in the same vilalge where she was born. Column 3and 4 examine the robustness of the baseline regression to different set of controls. Column 3 does not include any control. Column 4 includes farmer's controls, indicators of use of other inputs such as fertilizers, manure and improved seed as well as the interaction of time trends with indicators of ecological zone, region, proximity to coast, and proximity to region capitals. In column 5, "active mines" are mines that had some production in period 1988-2004, while "future mines" are mines that started operations after 2005 or have not started production yet but are in the stage of advanced exploration or development.

Table 8: CES function

Parameter	Estimate	S.E.
γ	-0.728***	0.003
λ	0.931***	0.050
ρ	-0.791***	0.222
η	0.997***	0.005
Implied σ	0.558	

Note: * denotes significant at 10%, ** significant at 5% and *** significant at 1%. Regression includes district and time fixed effects, indicators of proximity to each mine, and farmer's characteristics as in Table 3. Regression estimates $y_{ivt} = A_{ivt} = \exp(\gamma S_{vt} + \phi Z_i + \delta_d + \psi_t + \theta mine_v)[\eta M_{it}^{-\rho} + (1-\eta)L_{it}^{-\rho}]^{-\frac{\lambda}{\rho}}$ using non-linear least squares.

4.5 Effects on poverty

The previous results indicate a sizeable reduction in agricultural productivity and output. These effects, in turn, could worsen local living standards.

We focus on household poverty. Since agriculture is the main source of livelihood in rural Ghana, the loss of productivity could reduce income, and increase poverty.[56]. The net effect, however, is unclear. Mining companies or the government could, for example, promote local development projects, employ local workers, compensate local residents, or transfer part of the mining surplus. These policies are often implemented by the industry to mitigate potential negative side-effects of mining, and may offset the decline in productivity.

To examine this issue, we use data from the GLSS on poverty to estimate the following regression:

$$poverty_{idvt} = \phi_1 S_{vt} + \phi_2 W_i + \delta_d + \omega_{it} \tag{3}$$

where *poverty* is an indicator of the household being poor, and W_i is a set of household con-

[56]The standard framework presented above links a household's utility function that depends on consumption levels which in turn are linked to income form agricultural production.

trols.[57] The rest of the specification is similar to equation (2).[58] The parameter of interest is ϕ_1 which captures the difference in the evolution of poverty in mining areas, relative to non-mining areas. Note that the identification strategy is a difference in difference, similar to the one used in the estimation of the production function.

Figure 5 depicts the evolution over time of poverty headcount in areas close and far from mines. There are two relevant observations. First, poverty declined steadily between 1988 and 2005 in areas far from mines. This trend is similar to the dramatic poverty reduction experienced in the rest of Ghana since the early 1990s (Coulombe and Wodon, 2007). Second, during the 1990s mining areas were less poor than non-mining areas, and poverty evolved similarly in both areas. During the expansion of mining, however, there is a significant trend change: poverty increases in mining areas between 1998-99 and 2005. As a result, mining areas become actually poorer than non-mining areas. Note that this increase in poverty parallels the reduction in agricultural output (see Figure 6).

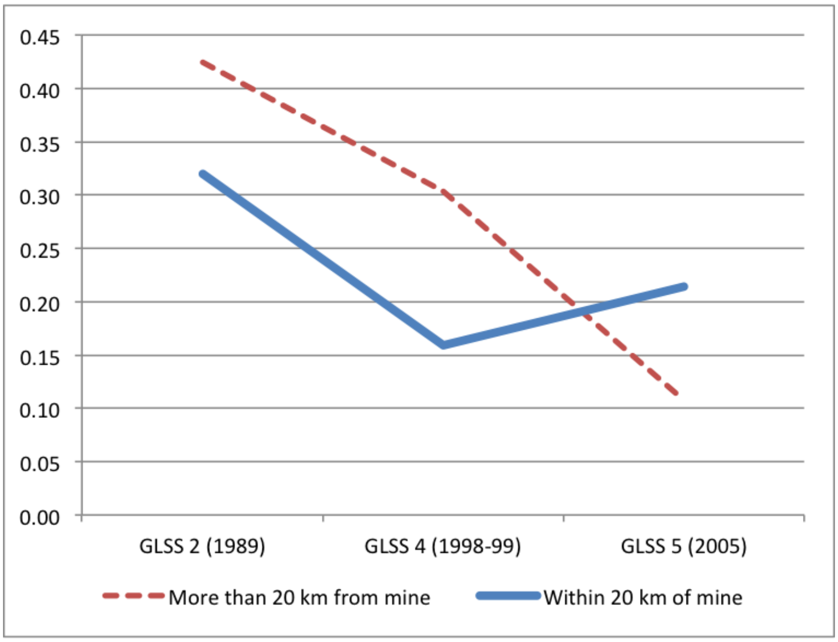

Figure 5: Evolution of poverty headcount

Table 9 presents the estimates of equation (3) using poverty as the outcome variable.[59]

[57]We use the poverty line used by the Ghana Statistical Service, i.e., 900,000 cedis per adult per year in 1999 Accra prices. The poverty line includes both essential food and non-food consumption (Ghana Statistical Service, 2000). We check the robustness of the results to alternative poverty lines such as USD 1.25 PPP a day.

[58]We also estimate this model by OLS using sample weights and clustering the errors at district level.

[59]We estimate equation (3) using only data from the last two rounds of the GLSS. We do not use data from

Column 1 shows results for all households using our preferred specification. As a reference, column 2 uses as proxy of S_{vt} the interaction between a dummy of proximity to a mine and a time trend to obtain the average effect of mining on poverty. Columns 3 and 6 split the rural sample between urban and rural households, respectively. Column 4 looks at rural households that are engaged in household production (and thus were included in the estimation of the agricultural production function,) while column 5 looks at rural households that did not report any agricultural production. [60] We also check the robustness of the results to using a continuous measure of real household expenditure (see table 15 in the Appendix).[61]

The picture that emerges is similar to the one observed in Figure 5. There a positive and significant relation between mining activity and poverty. The magnitude of the effect is sizeable: the increase in gold production between 1998 and 2005 is associated to an increase of almost 16 percentage points in poverty headcount. The effect is concentrated among rural inhabitants, regardless of whether the households are producers or not. Non-producers could be affected either directly, by the reduction in agricultural wages associated to lower total factor productivity, or indirectly, if they sell good or services locally.[62] The reduction in indicators of economic well-being is consistent with the reductions in agricultural productivity found above, in an area where farming activities are the main source of livelihood. Table 16 in the Appendix shows two additional results among children that are also consistent with levels of poverty induced by pollution: malnutrition and acute respiratory diseases have both increased in mining areas. Taken together, these findings suggest that compensating policies, if any, may have been insufficient to offset the negative shock to agricultural income.

GLSS 2, which are available, in order to keep the estimates comparable to the results on agricultural productivity. The results including this survey round are, however, similar.

[60] Note that households whose members are engaged in farming as wage laborers are around 65% of the sample.

[61] To construct the measure of real expenditure, we deflate nominal expenditure per capita with the index oflocal agricultural prices used to obtained measures of real agricultural output. The results using the official consumer price index are, however, similar.

[62] Aragon and Rud (2013) discuss the conditions under which these effects would be present and show evidence for the households in the area of influence of a gold mine in Peru.

Table 9: Mining and poverty

			Poverty			
			Rural			Urban
	All households		All	Farmers	Non-farmers	
	(1)	(2)	(3)	(4)	(5)	(6)
Cumulative gold prod. within 20 km.	0.162***		0.275***	0.296***	0.235***	0.062
	(0.045)		(0.075)	(0.079)	(0.061)	(0.041)
Within 20 km of mine × GLSS 5		0.159***				
		(0.058)				
Observations	5,527	5,527	3,393	2,540	853	2,134
R-squared	0.222	0.222	0.245	0.257	0.251	0.199

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions are estimated using ordinary least squares, and include district and year fixed effects as well as household controls, such as: age, age^2, religion, place of birth and literacy status of household head, household size, and an indicator of urban areas. All columns, except column 2, also include dummies of proximity to each mine.

5 Concluding remarks

The results of this paper suggest that in cases where highly mechanized activities occur in the proximity of agricultural areas, environmental policy should consider the possible impact of production-related pollution on crop yields and local income. In particular, the loss of agricultural productivity, and farmers' income should be an important part of the policy debate on the costs and benefits of activities such as mining. As such, the presence of this externalities would overestimate the net contribution of modern sectors, such as mining, to an economy.

We find that total factor productivity, and crop yields have decreased in mining areas. Our estimates suggest a reduction of around 40% in agricultural productivity between 1998 and 2005. The negative effect is associated to polluting mines, and decreases with distance. The reduction in agricultural productivity is associated with an increase in rural poverty. During the analyzed period, measures of living standards have improved all across Ghana. However, households engaged in agricultural activities (whether as producers or workers) in areas closer to mining sites have been excluded from this process. As a consequence, their household poverty indicators have deteriorated.

In the case of mining, the policy debate usually focuses on the benefits mining could bring in the form of jobs, taxes or foreign currency. These benefits are weighted against environmental costs such as loss of biodiversity or health risks. However, local living standards may be also directly affected by the reduction in agricultural productivity. In fertile rural environments, such as Ghana, these costs may offset the country's benefits from mining. It also means that the scope of mitigation and compensation policies should be much broader.

Usually mitigation and compensation policies focus on populations directly displaced by mining. The negative effects of air and water pollution, however, can extend to a broader population, beyond the boundaries of mining licenses. These groups should also be considered in the area of influence of a mine. As a consequence, such activities introduce substantial redistributive effects on the economic activity and wealth of a developing country.

A simple back-of-the envelope calculation shows that, in this case, the scope for compensating transfers is low. In 2005, mining-related revenues amounted to US$ 75 millions, which represent around 2-3% of total government revenue.[63] Most of these revenue (i.e. 80%) is

[63]The low contribution of mining to fiscal revenue has been attributed to relatively low royalties (Akabzaa, 2009). For example, in the period of analysis, royalties were fixed at 3% of profits, even thought the regulatory

channeled to the central government. Local authorities (such as District Assemblies, Stools and Traditional Authorities) receive only 9% of mining royalties. Between 1999 and 2005, this represented in total around US$ 8 million (World Bank, 2006, p. 91). On the other hand, the average loss by farming households in mining areas, according to our main results, is in the order of US$ 97 millions[64]

These rough numbers show that the amount of tax receipts might not be enough to compensate losing farmers and that this situation is even worsened by the fact that a only small proportion of the tax receipts go back to affected localities. This strong redistribution can help to better understand local opposition to mining projects and demands for better compensation. More generally, and in the context of highly rural developing countries, this paper points out to the risks of not addressing the external effects of pollution generated by modern industries on agricultural production.

---

framework set by the Minerals Royalties regulations allowed for rates of up to 12%.

[64]This number is obtained by multiplying the number of producing household in mining areas, around 210,000, to their average loss,i.e. US$ 460.

References

Ackerberg, Caves, and Frazer, "Structural Identification of Production Functions," 2006. unpublished.

Akabzaa, Thomas, "Mining in Ghana: Implications for National Economic Development and Poverty Reduction," in Bonnie Campbell, ed., *Mining in Africa: Regulation and Development*, International Development Research Group, 2009, chapter 1.

—— **and Abdulai Darimani**, "Impact of mining sector investment in Ghana: A study of the Tarkwa mining region," `http://www.saprin.org/ghana/research/gha_mining.pdf` 2001.

Anttila, Pia, Juha-Pekka Tuovinen, and Jarkko V. Niemi, "Primary NO2 emissions and their role in the development of NO2 concentrations in a traffic environment," *Atmospheric Environment*, 2011, *45* (4), 986 – 992.

Aragon, Fernando M. and Juan Pablo Rud, "Natural resources and local communities: evidence from a Peruvian gold mine," *American Economic Journal: Economic Policy*, May 2013, *5* (2).

Armah, F. A., S. Obiri, D. O. Yawson, A. N. M. Pappoe, and A. Bismark, "Mining and Heavy Metal Pollution: Assessment of Aquatic Environments in Tarkwa (Ghana) using Multivariate Statistical Analysis," *Journal of Environmental Statistics*, 2010, *1* (4).

Aryeetey, Ernest, Osei Bafour, and Daniel K. Twerefou, "Impact of Mining Sector Reforms on Output, Employment and Incomes in Ghana, 1980-2002," Technical Publication 75, Institute of Statistical, Social and Economic Research (ISSER), November 2007.

Banerjee, Abhijit V., Paul J. Gertler, and Maitreesh Ghatak, "Empowerment and Efficiency: Tenancy Reform in West Bengal," *Journal of Political Economy*, April 2002, *110* (2), 239–280.

Bardhan, Pranab and Christopher Udry, "Development Microeconomics," *Oxford University Press*, 2012.

Besley, Timothy, "Property Rights and Investment Incentives: Theory and Evidence from Ghana," *The Journal of Political Economy*, October 1995, *103* (5), 903–937.

Blundell, Richard and Stephen Bond, "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 2000, *19* (3), 321–340.

Botchway, F. Nii, "Land Ownership and Responsibility for the Mining Environment in Ghana," *Natural Resources Journal*, Fall 1998, *38*, 509–536.

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini, "The Political Resource Curse," NBER Working Paper No. 15705 January 2010.

Caselli, Francesco and Guy Michaels, "Do Oil Windfalls Improve Living Standards? Evidence from Brazil," `http://personal.lse.ac.uk/casellif/papers/brazil.pdf`, June 2009.

___ **and Wilbur John Coleman**, "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 2001, *109* (3), 584–616.

Celarier, E. A., E. J. Brinksma, J. F. Gleason, J. P. Veefkind, A. Cede, J. R. Herman, D. Ionov, F. Goutail, J.-P. Pommereau, J.-C. Lambert, M. van Roozendael, G. Pinardi, F. Wittrock, A. Schönhardt, A. Richter, O. W. Ibrahim, T. Wagner, B. Bojkov, G. Mount, E. Spinei, C. M. Chen, T. J. Pongetti, S. P. Sander, E. J. Bucsela, M. O. Wenig, D. P. J. Swart, H. Volten, M. Kroon, and P. F. Levelt, "Validation of Ozone Monitoring Instrument nitrogen dioxide columns," *Journal of Geophysical Research (Atmospheres)*, May 2008, *113*, 15.

Chay, Kenneth Y. and Michael Greenstone, "The Impact of Air Pollution on Infant Mortality: Evidence From Geographic Variation In Pollution Shocks Induced By A Recession," *The Quarterly Journal of Economics*, August 2003, *118* (3), 1121–1167.

Coulombe, Harold and Quentin Wodon, "Poverty, livelihoods, and access to basic services in Ghana," `http://siteresources.worldbank.org/INTGHANA/Resources/CEM_poverty.pdf` 2007.

Currie, Janet, Eric A. Hanushek, Matthew Kahn E. Megana nd Neidell, and Steven G. Rivkin, "Does Pollution Increase School Absences," *The Review of Economics and Statistics*, November 2009, *91* (4), 682–694.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken, "Climate Change and Economic Growth: Evidence from the Last Half Century," NBER Working Papers 14132, National Bureau of Economic Research, Inc 2008.

Duflo, Esther and Rohini Pande, "Dams," *The Quarterly Journal of Economics*, 2007, *122* (2), 601–646.

Duncan, Edward E., Jerry S. Kuma, and Seth Frimpong, "Open Pit Mining and Land Use Changes: An Example from Bogosu-Prestea Area, South West Ghana," *The Electronic Journal of Information Systems in Developing Countries*, 2009, *36*, 1–10.

Ebenstein, Avraham, "The Consequences of Industrialization: Evidence from Water Pollution and Digestive Cancers in China," *Review of Economics and Statistics*, 2013/02/01 2012, *94* (1), 186–201.

Emberson, L.D., M.R. Ashmore, F. Murray, J.C.I. Kuylenstierna, K.E. Percy, T. Izuta, Y. Zheng, H. Shimizu, B.H. Sheu, C.P. Liu, M. Agrawal, A. Wahid, N.M. Abdel-Latif, M. van Tienhoven, L.I. de Bauer, and M. Domingos, "Impacts of Air Pollutants on Vegetation in Developing Countries," *Water, Air, & Soil Pollution*, 2001, *130*, 107–118.

Environment Canada, "Enviromental code of practice for metal mines," `http://www.ec.gc.ca/Publications/0B566B84-FBCC-4C75-81EA-ABCF62942A19/Environmental-Code-of-Practice-for-Metal-Mines-2009.pdf` 2009.

Ghana Statistical Service, "Poverty Trends in Ghana in the 1990s," `http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1181743055198/3877319-1190217341170/PovProf.pdf` 2000.

Ghatak, Maitreesh and Dilip Mookherjee, 2013.

Graff Zivin, Joshua S. and Matthew J. Neidell, "The Impact of Pollution on Worker Productivity," NBER Working Paper No. 17004 2011.

Greenstone, Michael and Rema Hanna, "Environmental regulations, air and water pollution, and infant mortality in India," NBER Working Paper No. 17210 2011.

Guiteras, Raymond, "The Impact of Climate Change on Indian Agriculture," `http://www.econ.umd.edu/research/papers/34` 2009.

Hanna, Rema and Paulina Oliva, "The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City," NBER Working Paper No. 17302 2011.

Hansen, Gary D. and Edward C. Prescott, "Malthus to Solow," *American Economic Review*, September 2002, *94* (4), 1205–1217.

Hilson, Gavin and Natalia Yakovleva, "Strained relations: A critical analysis of the mining conflict in Prestea, Ghana," *Political Geography*, 2007, *26*, 98–119.

Hornbeck, Richard, "The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe," *American Economic Review*, June 2012, *102* (4).

Human Rights Clinic, "The Cost of Gold: Communities Affected by Mining in the Tarkwa Region of Ghana," 2010.

Jayachandran, Seema, "Air Quality and Early-Life Mortality: Evidence from Indonesias Wildfires," *Journal of Human Resources*, 2009, *44* (4), 916–954.

Keskin, Pinar, "Thirsty Factories, Hungry Farmers: Intersectoral Impacts of Industrial Water Demand," 2009.

Kleffmann, Anita Niedojadlo Ralf Kurtenbach Jrg and Peter Wiesen, "Primary NO2 emissions and their impact on air quality in traffic environments in Germany," *Environmental Sciences Europe*, 2012, *24* (21).

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70*, 317–342.

Lewis, W. Arthur, "Economic Development with Unlimited Supplies of Labour," *The Manchester School*, 1954, *22* (2), 139–191.

Maggs, R., A. Wahid, S. R. A. Shamsi, and M. R. Ashmore, "Effects of ambient air pollution on wheat and rice yield in Pakistan," *Water, Air, Soil Pollution*, 1995, *85*, 1311–1316. 10.1007/BF00477163.

Marshall, Fiona, Mike ashmore, and Fiona Hinchcliffe, "A hidden threat to food production: Air pollution and agriculture in the developing world," Gatekeeper Series SA73, International Institute for Environment and Development (IIED) 1997.

Matsuyama, Kiminori, "A Simple Model of Sectoral Adjustment," *The Review of Economic Studies*, 02 1992, *59* (2), 375–387.

——, *structural change*, Basingstoke: Palgrave Macmillan, 2008.

Mehlum, Halvor, Karl Moene, and Ragnar Torvik, "Institutions and the Resource Curse," *The Economic Journal*, 2006, *116* (508), 1–20.

Nevo, Aviv and Adam Rosen, "Inference with Imperfect Instrumental Variables," *Review of Economics and Statistics*, 2012.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Sachs, Jeffrey D. and Andrew M. Warner, "Natural Resource Abundance and Economic Growth," NBER Working Papers 5398, National Bureau of Economic Research, December 1995.

—— **and** ——, "The Curse of Natural Resources," *European Economic Review*, May 2001, *45* (4-6), 827–838.

Serfor-Armah, Y., B. Nyarko, S. Dampare, and D. Adomako, "Levels of Arsenic and Antimony in Water and Sediment from Prestea, A Gold Mining Town in Ghana and its Environs," *Water, Air, & Soil Pollution*, 2006, *175*, 181–192. 10.1007/s11270-006-9127-9.

Tetteh, Samuel, Adem A. Golow, David K. Essumang, and Ruphino Zugle, "Levels of Mercury, Cadmium, and Zinc in the Topsoil of Some Selected Towns in the Wassa West District of the Western Region of Ghana," *Soil and Sediment Contamination: An International Journal*, 2010, *19* (6), 635–643.

Vicente, Pedro C., "Does oil corrupt? Evidence from a natural experiment in West Africa," *Journal of Development Economics*, May 2010, *92* (1), 28–38.

WACAM, "Water report: impact of gold mining on water quality in Western Ghana," `http://www.wacamghana.com/app/webroot/img/documents/4accc8ce1913a.pdf` 2010.

World Bank, "Ghana Country Environmental Analysis," Report 36985-GH, The World Bank 2006.

A Additional results

Table 10: Evolution of agricultural output in mining vs non-mining areas

	ln(real agricultural ouput) (1)	(2)
Within 20 km of mine × GLSS 4	0.403 (0.282)	
Within 20 km of mine × GLSS 5		-0.515* (0.256)
Sample	GLSS 4 (1998/99) and GLSS 2 (1989)	GLSS 5 (2005) and GLSS 4 (1998/99)
Estimation	OLS	OLS
Farmer's controls	Yes	Yes
Observations	1,523	1,627
R-squared	0.483	0.223

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include district and survey fixed effects, as well as a set of farmer characteristics as in Table 3. *GLSS 4* and *GLSS 5* are indicators equal to 1 if survey is GLSS 4 or 5, respectively. *Within 20 km of mine* is a dummy equal to 1 if household is in a mining area.

Figure 6: Evolution of the unconditional mean of ln(real agricultural output)

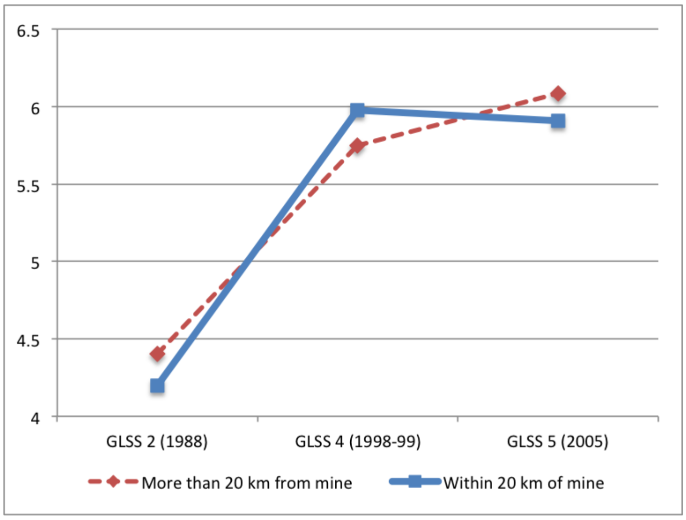

Table 11: Main results using time trend as treatment variable

	ln(real agricultural ouput)				
	(1)	(2)	(3)	(4)	(5)
Within 20 km of mine × GLSS 5	-0.515*	-0.567**	-0.567**	[-0.377, -0.511]	[-0.563, -0.554]
	(0.256)	(0.237)	(0.246)	(-0.316, -0.538)	(-0.569, -0.549)
ln(land)		0.628***	0.676***	[0.195, 0.676]	[0.735, 0.676]
		(0.037)	(0.048)	(-0.027, 0.773)	(0.769, 0.615)
ln(labor)		0.225***	0.350***	[0.656, 0.350]	[0.131, 0.350]
		(0.033)	(0.113)	(0.799, 0.288)	(0.001, 0.577)
Estimation	OLS	OLS	2SLS	IIV	IIV
Farmer's controls	Yes	Yes	Yes	Yes	Yes
Imperfect IV for:				Land	Labor
Observations	1,627	1,627	1,627	1,627	1,627
R-squared	0.223	0.449	0.441		

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. The set of farmer's controls includes: household head's age, literacy, and an indicator of being born in the village; as well as an indicator of the household owning a farm plot. All regressions include district, survey and mine fixed effects. Column 4 and 5 identify parameter bounds using the imperfect instrumental variable approach in Nevo and Rosen (2010). Identified parameter bounds are in brackets while the 95% confidence interval is in parenthesis. Confidence intervals are calculated adding (substracting) 1.96 standard deviations to the upper (lower) bound. GLSS 5 is an indicator equal to 1 if survey is GLSS 5 (year 2005).

Table 12: Main results using official CPI as price deflator

	ln(value agricultural ouput / CPI)				
	(1)	(2)	(3)	(4)	(5)
Cumulative gold	-0.442	-0.733***	-0.727**	[-0.297, -0.677]	[-0767, -0.687]
prod. within 20 km.	(0.276)	(0.268)	(0.273)	(-0.125, -0.752)	(-0.807, -0.619)
ln(land)		0.630***	0.675***	[0.185, 0.675]	[0.746, 0.675]
		(0.038)	(0.050)	(-0.038, 0.772)	(0.782, 0.614)
ln(labor)		0.219***	0.368***	[0.673, 0.368]	[0.109, 0.368]
		(0.036)	(0.115)	(0.813, 0.308)	(-0.021, 0.592)
Estimation	OLS	OLS	2SLS	IIV	IIV
Farmer's controls	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes
Imperfect IV for:				Land	Labor
Observations	1,627	1,627	1,627	1,627	1,627
R-squared	0.256	0.472	0.462		

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. The set of farmer's controls includes: household head's age, literacy, and an indicator of being born in the village; as well as an indicator of the household owning a farm plot. All regressions include district, survey and mine fixed effects. Column 4 and 5 identify parameter bounds using the imperfect instrumental variable approach in Nevo and Rosen (2010). Identified parameter bounds are in brackets while the 95% confidence interval is in parenthesis. Confidence intervals are calculated adding (substracting) 1.96 standard deviations to the upper (lower) bound.

Table 13: First stage regressions

	ln(land)	ln(labor)
	(1)	(2)
ln(land owned)	0.926***	0.183***
	(0.026)	(0.038)
ln(nr adult equivalents)	0.016	0.470***
	(0.017)	(0.051)
F-test excl. instruments	723.7	88.2
Observations	1,627	1,627
R-squared	0.804	0.254

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All columns include district, survey and mine fixed effects, as well as farmer's characteristics. See Table 3 for details on the second stage.

Table 14: Imperfect instruments with multiple endogenous variables

$(\lambda_{land}, \lambda_{labor})$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$
(0,0)	-0.723	0.676	0.368
(0,0.2)	-0.640	0.610	0.608
(0,0.4)	-1.076	0.957	-0.654
(0,0.6)	-0.851	0.778	-0.004
(0,0.8)	-0.821	0.754	0.086
(0,1)	-0.808	0.744	0.121
(0.2,0)	-0.741	0.703	0.351
(0.2,0.2)	-0.659	0.635	0.582
(0.2,0.4)	-1.064	0.972	-0.557
(0.2,0.6)	-0.864	0.805	0.007
(0.2,0.8)	-0.835	0.781	0.088
(0.2,1)	-0.823	0.772	0.121
(0.4,0)	-0.776	0.756	0.318
(0.4,0.2)	-0.696	0.685	0.529
(0.4,0.4)	-1.046	0.997	-0.398
(0.4,0.6)	-0.886	0.854	0.027
(0.4,0.8)	-0.860	0.832	0.093
(0.4,1)	-0.850	0.823	0.120
(0.6,0)	-0.873	0.904	0.226
(0.6,0.2)	-0.814	0.845	0.362
(0.6,0.4)	-1.011	1.044	-0.096
(0.6,0.6)	-0.938	0.970	0.074
(0.6,0.8)	-0.924	0.957	0.106
(0.6,1)	-0.918	0.951	0.119
(0.8,0)	-2.619	3.573	-1.437
(0.8,0.2)	2.196	-3.227	4.637
(0.8,0.4)	-0.918	1.170	0.709
(0.8,0.6)	-1.219	1.596	0.328
(0.8,0.8)	-1.333	1.757	0.184

Note: $\lambda_X = \frac{\rho_{Z_X,\epsilon}}{\rho_{X,\epsilon}}$ where $X = land, labor$ and Z_X is the instrument for X, i.e., the endowment of input X. λ_X measures how well the instrument satisfies the exogeneity assumption. $\lambda_X = 0$ corresponds to an exogenous, valid, instrument. Note that the assumption that the instrument is less correlated to the error term that the endogenous variable implies that $\lambda_X < 1$. Table displays estimates of main parameters for values of $\lambda_X \in (0.0, 0.8)$

Table 15: Mining and household expenditure

	ln(real expenditure per capita)					
			Rural			Urban
	All households		All	Farmers	Non-farmers	
	(1)	(2)	(3)	(4)	(5)	(6)
Cumulative gold prod. within 20 km.	-0.134		-0.239	-0.331***	-0.034	-0.128
	(0.095)		(0.168)	(0.084)	(0.203)	(0.086)
Within 20 km of mine $\times$ GLSS 5		-0.214**				
		(0.102)				
Observations	5,527	5,527	3,393	2,540	853	2,134
R-squared	0.576	0.571	0.507	0.465	0.595	0.585

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions are estimated using ordinary least squares, and include district and year fixed effects as well as household controls, such as: age, age^2, religion, place of birth and literacy status of household head, household size, and an indicator of urban areas. All columns, except column 2, also include dummies of proximity to each mine.

Table 16: Mining, child nutrition and health

	Under 5 weight-for-age		Under 5 height-for-age		Diarrhea		Acute respiratory disease	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln (cumulative gold prod. within 20 km + 1)	-2.353**		-0.584		0.002		0.005**	
	(1.095)		(1.316)		(0.003)		(0.002)	
Within 20 km of mine × post 2003		-26.407**		2.852		0.020		0.054*
		(12.570)		(14.785)		(0.032)		(0.031)
Mother and child controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,554	3,304	2,486	3,236	2,711	3,522	2,712	3,520
R-squared	0.064	0.039	0.212	0.190	0.054	0.048	0.047	0.033

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. Mother and child controls include: mother education, child age and its square, child gender, access to piped water, and an indicator of being in a rural area. All regressions are estimated using OLS and include district and year fixed effects, as well as dummies of proximity to each mine.