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

Fernando M. Aragón† Juan Pablo Rud‡

Draft version: December 2013

Abstract

The development of modern sectors has long been linked to the displacement of traditional agriculture. The economic literature has focused on explanations associated with input reallocation, but has neglected other mechanisms, such as pollution externalities. To explore this issue, we examine the case of gold mining in Ghana. We find that mining has reduced agricultural total factor productivity, and increased rural poverty. Consistent with a pollution spillover, we document higher concentrations of air pollutants near mines. The results highlight an important channel —i.e., reduction in productivity— through which polluting industries can affect living conditions in rural areas.

Keywords: pollution, agricultural productivity, natural resources, environment and development

JEL: Q15, Q56, O13

*This project was funded by the International Growth Centre, grant RA-2010-12-005. We thank Douglas Gollin, Mushfiq Mubarak, Matt Neidell, Owen Ozier, Krishna Pendakur, Francis Teal, Chris Udry and seminar participants at BREAD/World Bank, Calgary, CSAE, EAERE, EEA-ESEM, IGC Ghana, Novafrica, Galway, Gothenburg, NEUDC, Tinbergen Institute, and Universiyy of Washington for useful comments and suggestions. 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 are associated with a displacement of traditional activities, such as agriculture, in favor of modern production[1]. In particular, there is a large literature that investigates labor reallocations across agriculture and industries (Lewis, 1954; Matsuyama, 1992; Caselli and Coleman, 2001; Hansen and Prescott, 2002; Matsuyama, 2008) and, more recently, their conflicting interests over valuable resources, such as land or water (Ghatak and Mookherjee, 2013; Keskin, 2009). However, the economic literature has put less emphasis on other negative spillover effects that are independent of input use, such as pollution and environmental degradation.

In this paper we fill this gap by providing evidence that polluting modern industries can impose a negative externality on traditional activities. Using the case of gold mining in Ghana, we show that mining has decreased crop yields and agricultural output, beyond any observable change in input use. The magnitude of the effect is economically significant and seems to be driven by cumulative pollution.

Our main contribution is to highlight the reduction on agricultural productivity as a channel through which polluting industries can affect economic activity and living conditions, specially in rural areas where agriculture is the main source of livelihood. This externality has been neglected despite the existing biological evidence linking pollution to a reduction in crops' health and yields (Emberson et al., 2001; Maggs et al., 1995; Marshall et al., 1997).

The case of gold mining in Ghana has several features that make it suitable to study the effect of modern industries on agriculture. First, most gold production is done in large scale, modern, mines. These mines are heavily mechanized and release air pollutants similar to other fuel-intensive activities, such as power plants and urban traffic.[2] Second, gold mines are located in the vicinity of fertile agricultural lands with important cash crops, such as cocoa. Finally, there is compelling evidence that the industry has a poor environmental record.[3]

We use micro-data from repeated cross-sections to estimate an agricultural production func-

[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] Gold mining also has other industry-specific pollutants, such as cyanide spills and acidic discharges. These pollutants are mostly carried by water or localized in the close vicinity of mine sites.

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

tion. We then examine the effect of mining on total factor productivity. To do so, we exploit two sources of variation: distance to a mine, and changes in mining production. The main identification assumption is that the change in productivity in areas far and close to a mine would be similar in the absence of mining. This allows us to isolate changes in agricultural output induced by input adjustments from those produced by pollution, that would affect total factor productivity. A main limitation of examining total factor productivity is, however, that we bundle all non-input channels through which pollution can affect output such as deterioriation of human and plants' health, degradation of soils, or reduction in crop growth.

To implement this approach, we use a household survey collecting agricultural data for 1997 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 total gold production by mine. As noted in the environmental literature, continuous emission of pollutants in the atmosphere by highly mechanized operations are carried over long distances by winds and can build up to levels that impoverish soils and damage vegetation.[4] Because the two rounds of surveys are distant in time, we use the cumulative production in the period to proxy for the stock of pollution generated by mining operations.

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 data limitations. Instead, we use the analytical framework of consumer-producer households (Benjamin, 1992; Bardhan and Udry, 1999) to derive a suitable empirical strategy. A standard OLS estimation would work if, when controlling for farmer's observable characteristics and district fixed effects, we fully capture productivity heterogeneity. We complement this strategy with an instrumental variables approach. We 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 check the robustness of our results using a partial identification approach proposed by Nevo and Rosen (2012). This approach allows for some correlation

[4] This can happen through a direct uptake of pollutants by trees, plants and soils or, indirectly, through acid rain.

between the (imperfect) instruments and the error term.

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 10 percent decline in productivity in areas closer to mines. Given the increase in mining activity between 1997 and 2005, this implies that the average agricultural productivity in mining areas decreased 40 percent relative to areas farther away. Similar results are obtained using partial measures of productivty, such as crop yields. The negative effects decline with distance and extend to areas within 20 km from mine sites.

The results are robust to alternative estimation methods and model specifications, and are driven by proximity to operating mines. A placebo test, for instance, shows no changes on productivity for farmers close to new mining projects that were not operating in the period of analysis. We also check that our results are not driven by (observable) changes in the composition of farmers or in agricultural practices. These may occur, for example, if there is migration of high skilled farmers, switching towards non-agricultural activities, or weakening of property rights that may affect agricultural investment (such as cocoa trees) as in Besley (1995).

We subsequently look at the effects 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 percent. The effects are present not only among agricultural producers, but extend to other residents in rural areas. There is, however, no effect on urban poverty.[5]

We interpret these results as evidence that pollution from mining activities is the most plausible channel to explain the reduction in agricultural output and productivity. The first piece of evidence supporting this interpretation comes from the finding that mining has not affected. agricultural input prices. This is contrary of what we could expect if the effects were driven by reallocation of local inputs to non-agricultural activities. The second piece of evidence is the finding of higher levels of air pollutants in mining areas. Using satellite imagery, we obtain local measures of nitrogen dioxide (NO_2), a key indicator of air pollution. We find

[5]This result contrasts with Aragon and Rud (2013) who find a positive effect of mining activities on household income. This maybe due to the scant backward linkages in the Ghanaian case. Note, however, that the increase in poverty implies that any existing positive effect has not been large enough to offset the loss of agricultural productivity.

that concentrations of NO_2 are higher in mining areas and decline with distance, in a way that parallels the reduction of agricultural productivity.

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, such as infant mortality (Chay and Greenstone, 2003; Jayachandran, 2009), school absence (Currie et al., 2009), and incidence of cancer (Ebenstein, 2012).[6] Recent papers have also started to explore other possible economic effects of health problems caused by pollution, such as reduction on labor supply and labor productivity. For example, 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 (2012) find a negative effect of air pollution on labor productivity of piece-rate farm workers in California's central valley. Our results complement their findings in two ways. First, we estimate the reductions on total factor productivity, not only on labor productivity. Thus we take into account reductions in productivity that may occur, for instance, if land becomes less productive or if crop yields decline. This distinction is relevant from a policy perspective since it provides a better overview of the total costs imposed by pollution externalities. Second, we explore how pollution ultimately affects measures of living standards, such as consumption and 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.[7] In this paper we focus on the negative spillovers due to an unexplored channel in the natural resources literature, i.e., pollution. 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

[6]See Graff Zivin and Neidell (2013) and Currie et al. (2013) for a comprehensive review of this literature.

[7]See, for example, Caselli and Michaels (2013), Brollo et al. (2010) and Vicente (2010) for (negative) political economy channels, and Aragon and Rud (2013) for more positive market channels.

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 data. Section 4 presents the main results. Section 4.3 explores possible channels, while Section 5 presents additional checks and results. Section 7 concludes.

2 Background

Our empirical analysis uses the case of gold mining in Ghana. Our dataset has information on agricultural outputs and inputs collected for the years 1997 and 2005. As shown in Figure 1, before 1997, gold production was increasing from low levels of production. This was mostly driven by the expansion of one mine, Obuasi. After 1997, gold production flattens at a higher level, and reaches a greater number of locations. Many of these mines were new or experienced a significant expansion (e.g. Tarkwa, Bibiani and Damang).[8]

Our measure of mining activity is cumulative gold production. This gives us a measure of the exposure to stock pollutants that can produce detrimental effects on soils and vegetation and affect agricultural productivity, such as heavy metals and acid rain.[9] Table 1 shows that aggregate cumulative production has almost tripled between the two relevant years (1997 and 2005) and that there is substantial variation across mines. We exploit these differences in gold production by mine in our empirical analysis.

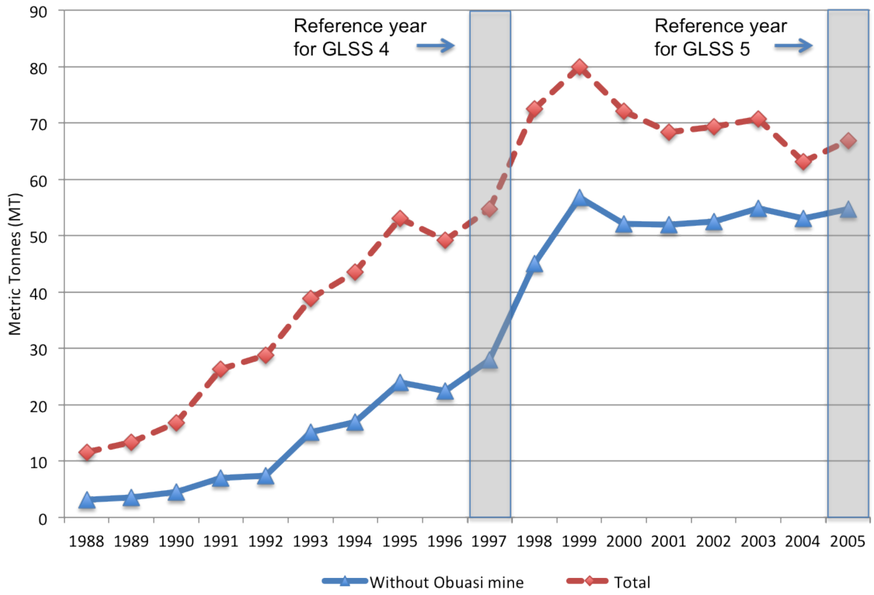
Most of the gold (around 97%) is produced by modern, large-scale mines.[10] These mines, similar to other modern mines in the world, are capital intensive, highly mechanized operations. They are located in rural areas, amidst fertile agricultural land, and have little interaction with local economies: they hire few local workers, buy few local products, their profits are not

[8] Note that the main results are robust to excluding observations in the vicinity of Obuasi mine. We report this in columns 6 and 7 in Table 7.

[9] The environmental literature distinguishes two types of pollutants: flow, or fund, pollutants and stock pollutants. Flow pollutants are dissipated or absorbed by the environment, so their effects are short-lived. In contrast, stock pollutants accumulate in the environment over time. The distinction between these types of pollutants is, however, subtle. For example, some pollutans, like NO_2, are considered flow pollutants. However, if emissions are relatively large, it can cause acid rain which has negative cumulative effects in the form of soil degradation.

[10] 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 owned by locals.

Figure 1: Total gold production (in MT), by year



Source: U.S. Geological Service, *The Mineral Industry of Ghana* 1994-2004, Infomine, and Aryeetey et al. (2007).

Table 1: Cumulative gold production by mine, in Metric Tonnes (MT)

Mine name	Type	Cumulative production 1988-1997	Cumulative production 1998-2005	Diff.
Bibiani	open pit	0.0	51.2	51.2
Bogoso/Prestea	open pit, underground and and tailings	23.9	55.9	32.0
Central Ashanti	open pit	5.4	9.7	4.3
Damang	open pit	0.0	73.6	73.6
Dunkwa placer	placer	1.2	1.2	0.0
Essase placer	placer	2.8	12.4	9.6
Iduapriem/Teberebie	open pit	19.6	61.2	41.6
Konong/Obenamasi	open pit	1.5	1.5	0.0
Obotan	open pit	2.2	19.4	17.3
Obuasi	open pit and underground	204.3	346.3	142.0
Tarkwa	open pit and underground	9.4	121.0	111.6
Wassa	open pit	0.0	10.3	10.3
TOTAL		270.3	763.7	493.4

Source: U.S. Geological Service, *The Mineral Industry of Ghana* 1991-2004, Infomine, and Aryeetey et al. (2007).

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 displace 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 (Lewis, 1954), land (Ghatak and Mookherjee, 2013), or water (Keskin, 2009).

In this paper we explore an alternative channel: the possible negative effect of environmental pollution on agricultural productivity (i.e., output conditional on quantity of inputs). 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 generated by fuel combustion, such as nitrogen oxide (NO_x) and sulfur dioxide (SO_2), on vegetation's health and yields.[11] When emitted to the atmosphere, these pollutants may remain in the air for several days and be dispersed over long distances by winds.

These airborne pollutants can affect vegetation in several ways. First, since they are poisonous they can directly affect crops' health and growth. For example, Emberson et al. (2001), Maggs et al. (1995), and Marshall et al. (1997) find drastic reductions of around 20 to 60 percent in yields of main crops -e.g. rice, wheat, and beans- due to the exposure to polluted air from urban centers.[12] Second, they can have cumulative, long-term, effects through acid rain.[13] Acid rain is caused by the combination of airborne pollutants (such as NO_x or SO_2) with rain water. Acid rain causes degradation of soils by leaching nutrients and releasing toxic substances, such as aluminum. In turns, this weakens vegetation and can cause slower growth, injury or death.[14]

[11] NO_x is a toxic gas by itself, but also contributes to the formation of tropospheric ozone. 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.

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

[13] For a summary of this evidence see, for example, the websites of the US and Canada environmental agencies (`http://www.epa.gov/acidrain/effects/forests.html` and `http://www.ec.gc.ca/air/default.asp?lang=En&n=7E5E9F00-1#ws0EF0FB73`)

[14] These negative effects could be, however, mitigated by the use of fertilizers, to replace lost nutrients, or

The above discussion suggests that air pollutants can negatively affect total factor productivity by reducing crop's health or quality of soil, a key agricultural input. These effects may occur in addition to deterioration of human health, which may also reduce worker's productivity, as documented by Graff Zivin and Neidell (2012).

It is important to note that large scale gold mines —akin to other industrial processes, power plants and motor vehicles— produce significant amounts of air pollutants, such as NO_2, SO_2 and particulate matter. The main direct sources of air emissions are petrol engines of heavy machinery, as well as fumes from smelters and refineries. This is in addition to other industry-specific pollutants, such as cyanide, heavy metals, or acid mine drainage. In modern mines, these pollutants tend to be more closely monitored and prompt mitigation actions. Importantly for our analysis, they are mostly carried by surface water. This may limit its impact on agriculture in the Ghanaian case, where most crops are rainfed.[15]

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 (Human Rights Clinic, 2010; Akabzaa, 2009; Aryeetey et al., 2007; Hilson and Yakovleva, 2007).[16] 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

crushed limestone, to reduce soil acidity.

[15]In Section 4.3 we explore the role of pollutants carried by surface waters.

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

distance to mining sites, and extend beyond mining areas, probably due to the aerial dispersion of metals from mining areas.

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 households (Benjamin, 1992; Bardhan and Udry, 1999). This framework has been used to analyze farmers' decisions when consumption (affecting utility) and production (affecting the budget constraint) are interrelated. In our case, it clarifies how mining could affect input use and agricultural output, and guides the empirical analysis.

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, sell them in local input markets (L^s, M^s) at prices w and r, or, in the case of labor, also consume it as leisure. As producers, households can buy additional labor and land (L^b, M^b).

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

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

We assume households are heterogeneous in their access to markets for inputs.[17] In par-

[17] 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. Even though, in the context of a region in Indonesia, Benjamin (1992) fails to reject separability between production and consumption, data for Ghana show that inputs markets are thin: in the area of study, around 8% of

ticular, 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.[18] 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 Ghana. Similar arguments can be made about labor markets, due to market incompleteness, farmers' preference for working on their own land, or household and market labor not being perfect substitutes.

In the case of unconstrained farmers, the maximization problem follows the separation property: the household chooses the optimal amount of inputs to maximize profits and, separately, chooses consumption and leisure 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 total factor productivity and input prices.

In the case of fully-constrained farmers, i.e., farmers unable to sell or buy inputs, the optimal input decisions are shaped by their endowments. Since the opportunity cost of land is zero, they will use all their land endowment, $M^* = E^M$. In the case of labor, however, farmers still face a trade-off between leisure and income. Solving the household's problem, the optimal level of labor, $L^*(A, E^M)$, depends now of total factor productivity and land endowment.[19]

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 among unconstrained farmers. Similar effects would occur if, for example, mines reduce supply of inputs due to land grabbings or population displacement. There would be, however, no effect on productivity, A.[20] Also note that the effect on consumption depends on the relative size of endowments. If endowments are small, so that a household

available land is rented, and only 1.4% of the total farm labor (in number of hours) is hired. As shown in Table B.4 in the Appendix, endowments are a very strong predictor of input use.

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

[19] For a fully constrained farmer, the household's problems simplifies to $\max U(c, l)$ subject to $c = F(A, L, E^M)$ and $L = E^L - l$. The first order condition is $U_c F_L = U_l$.

[20] This remark depends, however, on the assumption that input type does not change.

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).[21]

Second, mining-related pollution may affect crop's health and yields, as well as quality of inputs as discussed above. This would imply a reduction in output even if the quantity of inputs used remains unchanged. In terms of the model, this represents a drop in productivity, A. This would, unambiguously, have a negative effect on agricultural output and household's consumption. Additionally, it might reduce input use. In particular, labor use might fall either by reducing labor demand for unconstrained farmers or through a substitution of labor towards leisure for constrained farmers. In the case of land, only unconstrained farmers would reduce their land use. The empirical implication of this is that we would only observe a drop in land use in mining areas if the share of unconstrained farmers is high. Finally, contrary to what the input competition channel might deliver, input prices would decrease or remain unchanged, depending on how well markets reflect factors' marginal productivity.

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

1. 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, (iii) decrease input use among unconstrained farmers; and (iv) depending of the relative size of endowments, decrease or increase farmers' consumption.

2. If the main channel is through pollution, then mining would: (i) reduce agricultural output and productivity, A, (ii) decrease input prices, depending of the flexibility of markets, (iii) decrease input use among all farmers (except for land of constrained farmers), and (iv) unambiguously decrease farmer's consumption.

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

[21] For example, Duncan et al. (2009) suggests a reduction of around 15% in 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).

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 total factor productivity, A. We complement this approach by also studying the effect of mining on input 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:[22]

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

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. Other inputs, such as capital and materials (e.g. fertilizers, insecticides), are not widely used and thus excluded from the empirical analysis [23]. Their inclusion, however, does not change any of the results.

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 $\gamma < 0$.

As the empirical counterpart of S_{vt}, we use cumulative gold production near a farmer's locality.[24] This variable would be a reasonable proxy for exposure to pollutants under the assumption that pollutants have a cumulative effect, i.e. they are stock pollutants. As we discuss in Section 2, several pollutants released by mining operations, such as NO_2, SO_2 and heavy metals, can have negative cumulative effects on vegetation through acid rain and soil

[22]We assume a Cobb-Douglas technology for simplicity. In the empirical section, we check the robustness of the results to using a, more general, CES production function.

[23]For example, the value of tools and other capital goods is, on average, less than 1% of total output and the value of manure, seeds, fertilizers and insecticides account for less than 5%.

[24]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 the reference year of the household survey (i.e. 1997 for GLSS 4 and 2005 for GLSS 5). For areas farther than 20 km, $S_{vt} = 0$.

degradation.[25]

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 time variation in the repeated cross section to compare the evolution of productivity in mining areas relative to non-mining areas.

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 cumulative amount of gold produced in nearby mines.[26] 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.[27]

The second problem arises because 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 use farmers' observable characteristics to proxy for farmer heterogeneity, η_i. We also include district fixed effects to capture differences in average product due to local characteristics.[28] With these modifications, and taking logs, the model we estimate becomes:

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

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.

[25] In the empirical analysis, we also check the robustness of the results to measures of flow pollutants, i.e. short-lived pollutants, using annual gold production (see Table 5).

[26] We also use a simpler specification replacing S_{vt} by $(mining\_area_v) \times T_t$ where $mining\_area_v$ is an indicator of being close to a mine and T_t is a time trend. The results using this discrete treatment are, however, similar (see Table B.2 in the Appendix).

[27] In the Appendix, we explore the evolution of average agricultural output in areas closer and farther from mines for three years with available data: GLSS 2 (1988), GLSS 4 (1997) and GLSS 5 (2005). Figure A.3 shows that the evolution of output is remarkably similar in the first period (1988-1997), 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 B.1 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.

[28] Districts are larger geographical areas than localities v. We cannot use locality fixed effects due to the structure of the data.

δ_d and ψ_t represent district and time fixed effects, while $mining\_area_v$ is an indicator of being within 20 km of a mine (i.e. being in mining area). 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.

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 larger 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}.[29].

Finally, we consider the possibility that endowments are correlated to ξ_{ivt}.[30] 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 the set of parameter values.[31] The approach relies on two assumptions: (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 instrument is *less* correlated to the error than the endogenous variable. These (set) identification assumptions are weaker than the exogeneity assumption in the standard IV and OLS approaches.[32]

3.3 Data

Our main results use a repeated cross-section of household data from the rounds 4 and 5 of the Ghana Living Standards Survey (GLSS 4 and GLSS 5).[33] These surveys were collected by

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

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

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

[32]We refer the reader to Nevo and Rosen (2012) for a detailed exposition of the estimation method.

[33]We also use the GLSS 2, taken in 1988/89, 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

the Ghana Statistical Service (GSS) between April 1998 to March 1999, and September 2005 to August 2006, respectively. Note, however, that the questions on agricultural activities refer to the previous 12 months. Thus, the surveys reflect information on agricultural input and outputs mainly for years 1997 and 2005. We use these two years as the reference years to match household data with measures of mining activity.

The survey is representative at regional level and 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.[34] 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.[35]

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 first identify mines active during the period 1988 to 2005, and aggregate the annual production of each mine since 1988 to the survey's reference year for agricultural activities. Data on mining production by mine come mainly from reports prepared by the U.S. Geological Service (USGS).[36] This source covers year 1991 to 2004. We complete the remaining years with data from Infomine, and Aryeetey et al. (2007).[37]

Second, we obtain geographical coordinates of each mine site.[38] Using a geographical information system (ArcGIS), we identify the enumeration areas within different distance brackets

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.

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

[35] 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.

[36] 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`.

[37] Infomine (`http://www.infomine.com/minesite/`) provides production by mine for 2005, while Aryeetey et al. (2007) report aggregate production (measured by Ghana's Mineral Commission) for years prior to 1991. We impute production by mine for years 1988 to 1990 using mines' shares of gold production in 1991. Main results are, however, similar using only data from USGS for period 1991-2004.

[38] This information comes from proprietary industry reports prepared by Infomine.

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 A.1 displays a map of Ghana with the location of active gold mines between 1988 and 2005. Note that all mines are located in three regions: Western, Ashanti and Central. In the empirical section, we restrict the sample to these regions.[39] Figure A.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 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.

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.[40] This price index uses data from agricultural producers and varies by region and by mining and non-mining areas.[41]

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

[40] 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 B.3 in the Appendix). This consumer price has a lower geographical resolution than the one we use in this paper.

[41] 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 1997. Then, we take the median

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. The farmers are located in 42 districts in three regions of south west Ghana: Western, Ashanti and Central. Table 2 presents a simplified difference-in-difference estimation of the main variables of interest, by comparing mean values in both survey rounds for farmers located in areas close and far to any mining operations (independently of their size). A first important observation is that the log of agricultural output has shown a relative decrease near the mining areas. Consistent with the consumer-producer household framework, the poverty rate in affected areas shows a relative increase. On the contrary, there is no apparent significant difference in the use of the main inputs, land and labor. There is, however, a differential change in input prices even though the sign is not, as an increase in demand from mines would suggest, positive. A reduction in input prices might simply reflect the lower marginal productivity of inputs due to pollution.

There are also no significant differences in most farmers' characteristics, except for place of birth and land ownership. We deal with (potential) differences in farmers' characteristics in two ways. First, we include them in the main regressions. Second, we explore whether changes in farmer composition can explain our results.

value of prices and weights by region and by mining and non-mining area, i.e., six different values every survey, and construct a Laspeyres price index.

Table 2: Mean of main variables, by GLSS and location

Variable	Within 20 km of mine GLSS 4 (1)	Within 20 km of mine GLSS 5 (2)	Outside 20 km of mine GLSS 4 (3)	Outside 20 km of mine GLSS 5 (4)	Diff. columns (2-1) - (4-3) (5)
Cumul. gold prod. (MT)	41.7	84.6	-	-	-
ln(real agric. output)	6.6	6.2	6.5	6.6	-0.526***
					(0.174)
Land (acres)	7.2	17.9	8.3	9.4	9.671
					(9.505)
Labor (days)	377.3	358.8	343.1	366.3	-41.704
					(31.987)
Land owned (acres)	11.6	21.2	12.0	13.6	7.918
					(9.653)
Nr. adults equivalents	3.6	3.4	3.9	3.5	0.095
					(0.233)
ln(relative land price)	14.4	14.1	13.9	14.1	-0.519***
					(0.104)
ln(real wage)	8.6	8.8	8.4	8.8	-0.269***
					(0.042)
Age (years)	48.0	47.9	46.6	47.4	-0.944
					(1.956)
Literate (%)	53.1	46.6	54.5	45.3	0.027
					(0.063)
Born in village (%)	45.5	60.7	54.2	41.9	0.275***
					(0.062)
Owns a farm plot (%)	69.3	88.4	54.3	83.0	-0.095*
					(0.054)
Poverty headcount (%)	15.2	26.0	33.8	17.6	0.270***
					(0.050)
Nr. Observations	162	218	551	696	

Notes: Columns 1 to 4 report mean values for the sub-sample of farmers within and outside 20 km of a mine for every round of the GLSS. Means are estimated using sample weights. By definition, cumulative production in non-mining areas is equal to zero in both periods. Column 5 displays the difference in difference of columns 1 to 4. The standard errors are in parentheses. Total number of observations is 1627.

4 Main results

4.1 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[42]. As a reference, column 4 estimates the 2SLS regression using as proxy of S_{vt} the interaction between a dummy 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.[43] 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 10% in productivity.[44] 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. [45] The estimated

[42]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 B.4 in the appendix for the first stage regressions.

[43]The estimates of α and β, i.e., the participation of land and labor, also seem plausible. We cannot reject the hypothesis of constant returns to scale. Using the 2SLS estimates, the p-value of the null hypothesis $\alpha + \beta = 1$ is 0.773. We obtain a similar result of constant returns to scale when using a CES production function.

[44]The average value of the measure of mining activity (i.e., cumulative gold production within 20 km in hundreds of MT) increased from 0.417 in 1997 to 0.846 MT in 2005. The standard deviation of this variable is 0.617.

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

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 negative spillovers from modern industries in rural environments.

Columns 5 and 6 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.[46] Consistent with the results on productivity, we find a negative and significant relation between mining and crops yields.

Finally, we 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. We allow both instruments to be imperfect and run the IIV specification for different combinations of values of λ_{land} and λ_{land}, the parameters that measure the ratio of correlations of the instrument and the regressor with the error term.[47] Figure 2 shows that the effect on residual productivity is negative in the large majority of the cases (more than 95%) or, in other words, we need very specific combinations of λ_j for our main results not to hold.[48]

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

[46] 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.

[47] Note that $(\lambda_{land}, \lambda_{land}) = (0, 0)$ corresponds to the standard 2SLS estimate. For further details of the methodology see Nevo and Rosen (2012, section III.D).

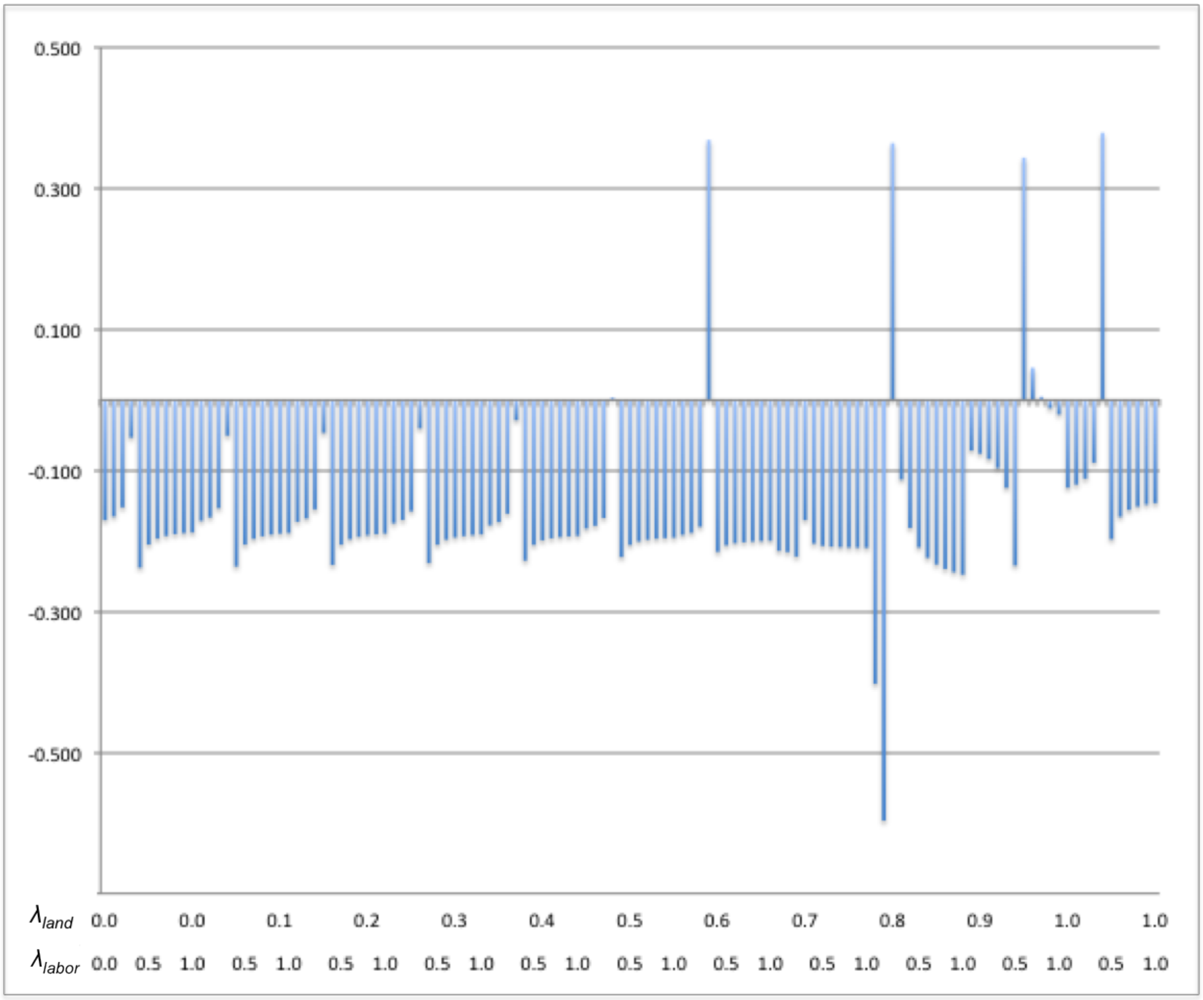
[48] For completeness, we also obtain analytical bounds proposed by Nevo and Rosen (2012) in the, more restrictive, case of only one imperfect instrument (see Table B.6).

Table 3: Mining and agricultural productivity

	ln(real agricultural output)				ln(yield cocoa)	ln(yield maize)
	(1)	(2)	(3)	(4)	(5)	(6)
Cumulative gold prod. within 20 km.	-0.149*	-0.176**	-0.170**		-0.509*	-0.420***
	(0.085)	(0.085)	(0.085)		(0.298)	(0.103)
Within 20 km of mine × GLSS 5				-0.565**		
				(0.240)		
ln(land)		0.631***	0.676***	0.678***		
		(0.038)	(0.047)	(0.046)		
ln(labor)		0.209***	0.352***	0.346***		
		(0.033)	(0.110)	(0.109)		
Estimation	OLS	OLS	2SLS	2SLS	OLS	OLS
Farmer's controls	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes		
Observations	1,627	1,627	1,627	1,627	948	605
R-squared	0.221	0.445	0.435	0.438	0.349	0.409

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 survey fixed effects, and an indicator of being within 20 km of a mine. Cumulative gold production is measured in hundreds of metric tonnes (MT).

Figure 2: Estimates of γ with multiple imperfect IVs

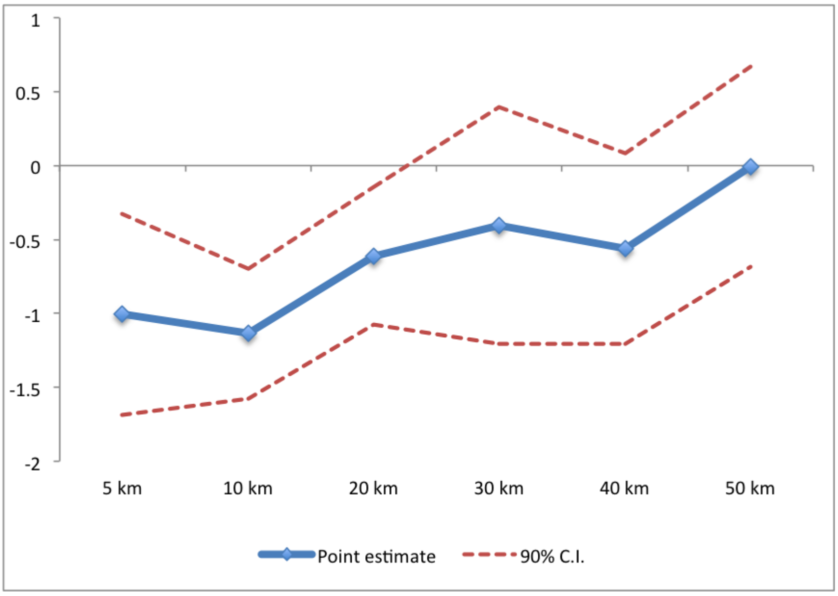


Note: Vertical axis displays estimates of γ for different values of λ_j, with $j = \{land, labor\}$. Values of λ_j in horizontal axis range from 0 to 1, with step increments of 0.1. $\lambda_j = \frac{\text{corr}(Z_j,\epsilon)}{\text{corr}(X_j,\epsilon)}$, where X is input use, Z is the instrumental variable and ϵ is the error term, measures how well the instrument satisfies the exogeneity assumption. $\lambda_j = 0$ corresponds to an exogenous, valid, instrument. The assumption that the instrument is less correlated to the error term that the endogenous variable implies that $\lambda_j < 1$.

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.

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



4.2 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.[49] 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.[50]

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.,

[49] 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).

[50] 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 farmers leaving the industry.

areas containing mineral deposits, processing units and tailings. These areas comprise, at most, few kilometers around the mine site.[51] In contrast, we document a drop 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 around a mine.[52]

A second way involves the increase in price of local inputs, i.e., the input competition channel discussed in Section 3.1. 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 relation between mining and input prices. Recall that the input competition channel has a different empirical implication for input prices than the pollution channel. If output is decreasing due to input competition, there would be a positive correlation between mining and input prices. In contrast, if results are driven by a negative shock on productivity, the relation should be negative or insignificant, depending on how competitive input markets are.

We also explore the relation between mining and input demands. Note that both channels (input competition and pollution) would predict a weakly negative relation, though for different reasons. In the first case, it would be due to an increase in input prices; while in the second, it would be due to reduction in factors' productivity. This distinction is relevant because, in the presence of lower productivity, input use may drop even if prices do not change.

Table 4 displays the results. 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.[53] To estimate input demands, we regress input use on measures of input prices, farmer's endowments

[51]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.

[52]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 5.1.

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

and proxies of total factor productivity, including mine activity.[54]

Note that the relation between mining and input prices is insignificant. This result weakens the argument that mining crowds out agriculture through increase in factor prices. Instead, it points out to a reduction in productivity as the main driver of reduction in agricultural output. The results on input demands are consistent with this interpretation. Despite no changes in input prices, demand for labor decreases with mining. This is expected in the presence of a negative productivity shock, as discussed in Section 3.1. The lack of response of input prices to this productivity shock could be due to imperfect input markets. In turn, this may explain why land demand does not change while labor demand decreases. As layed out in the analytical framework, in the absence of input markets, the opportunity cost of land is low so the whole endowment is used. In contrast, labor use is more responsive to productivity shocks since the labor endowment can always be consumed as leisure.

4.3 Pollution and productivity

We interpret the previous findings as evidence that agricultural total factor 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, several case studies show that water and soil in mining areas have higher than normal levels of pollutants (see Section 2).

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.[55]

Instead, we rely on satellite imagery to, indirectly, look for a smoking gun of the role of pollution.[56] The satellite imagery is obtained from the Ozone Monitoring Instrument (OMI)

[54]We check the robustness of these results to using annual gold production, instead of cumulative production, as proxy of mining activity, and including agricultural output as an additional control in the estimation of input demands (see Table B.7 in the Appendix).

[55]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.

[56]A similar approach of using satellite imagery to measure air pollutant is used by Foster et al. (2009) and Jayachandran (2009).

Table 4: 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.	-0.012	-0.040	-0.144**	-0.007
	(0.029)	(0.078)	(0.062)	(0.037)
ln(relative wage)			-0.093	0.019
			(0.153)	(0.117)
ln(relative land rent)			-0.085	0.009
			(0.071)	(0.038)
ln(nr. adult equivalents)			0.528***	0.022
			(0.062)	(0.021)
ln(land owned)			0.130***	0.914***
			(0.029)	(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	0.277	0.007	0.267	0.803

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 survey fixed effects and an indicator of being within 20 km of a mine. Columns 3 and 4 also include district fixed effects, and a set of farmer's controls similar to regressions in Table 3.

available at NASA.[57] 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 negative effects of NO_2 can be both short-term, by directly damaging plant's tissues, or cumulative, through acid rain and the subsequent degradation of soils. 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 large urban centers, industrial sites and heavily mechanized operations, such as large-scale mines.

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).[58] Levels of tropospheric and ground level NO_2 are, however, highly correlated.[59] 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 (2005). 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.[60] 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:[61]

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

where $NO2_v$ is the average value of tropospheric NO_2 in enumeration area v during the period

[57] 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`.

[58] 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.

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

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

[61] 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.

April-June 2005. X_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.[62] 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 5 present the empirical results using two alternative ways to measure mine activity[63]. We also replace the dummy X_v by a distance spline with breaks 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 6.

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.[64] 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.[65]

In Column 4, we run a regression that includes previous year production of the neighboring mines as well, to check whether effects on productivity result from a short run *flow* of pollution or whether the relevant variation comes from a *stock* of pollution built up due to the sustained level of production over many years. This distinction can be made because air pollutants can dissolve very quickly in the air, i.e. within a few days, even though their effects can accumulate

[62] 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.

[63] We use a semi-logarithmic specification since the relation between mining activity and NO_2 concentration is likely to be non-linear. For example, Kurtenbach et al. (2012) and Anttila et al. (2011) show that large changes in emissions (or source of emissions, such as petrol cars) are necessary to produce small changes in NO_2. We also estimate other non linear specifications, such as quadratic and third degree polynomials, with similar results.

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

[65] 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 5: Mining and pollution

	Average NO_2		ln(real agricultural output)		
			Using mining as IV	Stock vs. flow	Upstream vs downstream
	(1)	(2)	(3)	(4)	(5)
Within 20 km of mine	0.325***				
	(0.111)				
Ln (cumulative gold prod. within 20 km+1)		0.011*			
		(0.006)			
Average NO_2			-1.683*		
			(0.974)		
Cumulative gold prod. within 20 km				-0.220**	-0.162*
				(0.093)	(0.096)
Annual gold prod. within 20 km				0.016	
				(0.018)	
Cumul. gold prod. within 20 km x downstream					-0.050
					(0.081)
Estimation	OLS	OLS	2SLS	OLS	OLS
Farmer's controls	No	No	Yes	Yes	Yes
Controlling for inputs	No	No	Yes	Yes	Yes
Observations	399	399	914	1,627	1,627
R-squared	0.238	0.230		0.445	0.445

Notes: Robust standard errors in parentheses. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. Columns (1)-(3) use data for 2005 only. Column (4) uses the same data as in the baseline specification. 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. Column 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 reports 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 measures of input use as in the baseline regression (see notes of Table 3). Column 4 replicates baseline OLS regression (column 2 in Table 3) adding "annual gold production within 20 km" as a proxy for flow pollutants. This variable measures the production of gold (in hundreds of MT) from nearby mines in years 1997 and 2005. Column 5 adds to the baseline OLS regression an interaction term of the measure of mining activity and "downstream", a dummy equal to one if household is downstream of an active mine. Standard errors are clustered at the district level.

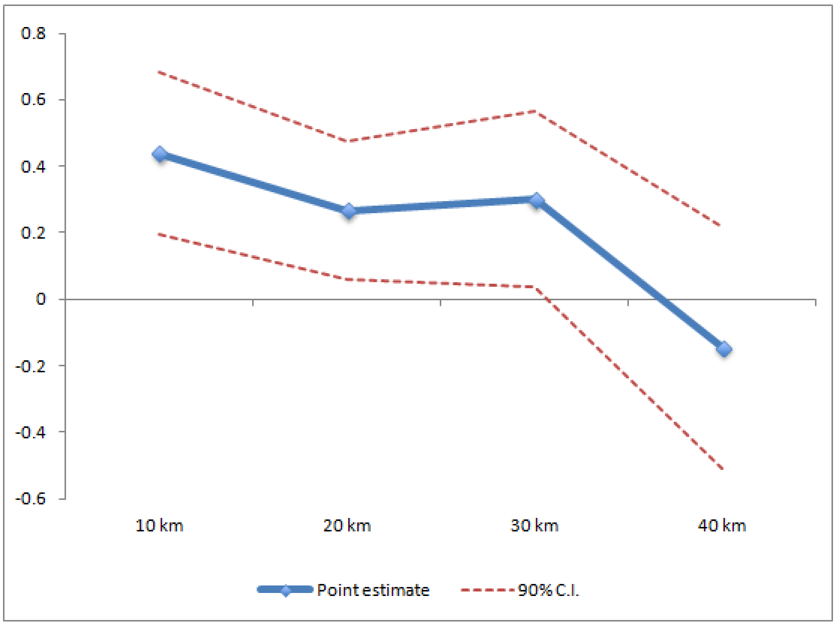
more progressively on soils, trees and plants. A non-significant coefficient for the new variable suggests that the reduction in productivity can only be explained by variation in the measure of long-term exposure to pollution.

Finally, we explore the importance of pollutants carried by surface water. To do so, we identify areas downstream of active mines and examine whether the negative effects of mining are stronger in these 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.[66] 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 5, 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.[67]

[66] 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.

[67] Additionally, there is no variation in productivity that can be explained by the direction of winds. Ghana has two main winds that come from opposite directions: the Harmattan, a dry and dusty wind, that blows from the Sahara, i.e., north east, and another wind, warm and moist, coming from the Atlantic ocean, i.e., south-west. Hence, air pollutants may be dispersed in all directions around a mine.

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



5 Additional checks

5.1 Compositional effects and property rights

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 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 probability that a worker is engaged in agriculture (either as a producer or laborer). In the presence of occupational change towards non-agricultural activies we could expect a negative correlation. Second, we look at measures of agricultural workers' demographics and mobility, such as probability of being a male in prime age (20-40 years), or being born in the same village where they reside, Third, we explore measures of human capital of agricultural workers, such as literacy and having completed secondary school.[68] 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.

An alternative story that could explain lower agricultural productivity is related to weak property rights. In the case of Ghana, two phenomena are at play: customary and weakly defined land rights and the right of the state to grant licences for the use of land where mineral wealth is located (see Botchway (1998) for a discussion). Farmers near mining sites might fear expropriation and might choose to move away from activities with long run benefits and short run costs (e.g., cocoa trees) and into crops with shorter cycles that require less attention. The link between property rights and cocoa tree planting decisions in the case of Ghana (and the Wassa region in particular) has already been discussed in Besley (1995). We follow a similar approach and check whether there is any perceptible change in cocoa shares, in crop composition, in the decision to plant new cocoa trees in the previous year, or in the decision to grow cocoa (this would capture decisions made in the last five years).[69]

Table 6 displays the results. In all cases, there is no significant relation between mining activity and observable population characteristics. Additionally, we find some adjustments to cropping decisions, but the results are the opposite of what the property rights story would suggest. If anything, there has been an increase in specialization and investment in farms near mining areas. These results weaken the argument that the reduction in productivity is driven solely by changes in perceived risk of expropriation or changes in the composition of farmers.

5.2 Alternative specifications

In Table 7, we check that our results are robust to alternative specifications. Column 1 allows for heterogeneous effects between local and non-local farmers. We define a farmer as local if she was born in the same village where she 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, unproductive, lands. Note, however, that there is not significant different in the relation between mining and agricultural output between these two types of farmers.

[68] Levels of completion of primary school are high, i.e., around 86%, while literacy levels (47.8%) and secondary school completed (36.3%) show greater variation. Results hold when using data on completed primary school.

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

Table 6: Robustness checks: compositional changes

	Works in agriculture	Male in prime age	Born in village	Literacy	Completed secondary	Share of cocoa	Crop concentration	New cocoa plants	Grows cocoa
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cumulative gold prod. within 20 km	-0.032	-0.001	-0.006	-0.004	-0.013	0.021	0.043**	0.066*	0.022
	(0.042)	(0.018)	(0.024)	(0.021)	(0.016)	(0.036)	(0.017)	(0.039)	(0.032)
Sample	All workers	Agric. workers	Agric. workers	Agric.ral workers	Agric. workers	Agric. households	Agric. households	Agric. households	Agric. households
Observations	8,932	4,978	4,929	4,971	4,978	1,627	1,627	1,627	1,627
R-squared	0.359	0.029	0.127	0.044	0.134	0.446	0.118	0.159	0.481

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, an indicator of beingin a mining area, and indicators of ecological zone and urban area. *Works in agric.* is an indicator equal to one if individuals works in agriculture as a laborer or producer. *Male in prime age* is an indicator equal to one if individual is male between 20 to 40 years old. *Born here* is an indicator equal to one if individual was born in the same village where she resides. "New cocoa plants" equals one if the farmer has planted new cocoa trees in the previous 12 months and zero, otherwise. "Grows cocoa" equals one if the farmer has grown cocoa in the previous 12 months and zero, otherwise. Columns 1 to 5, 8 and 9 are estimated using a linear probability model. Column 1 includes as additional controls: age, age^2, religion, place of birth, literacy status, and household size. Columns 3 and 4 examine the educational attainment of agricultural workers conditional on age and age^2. Columns 6 to 9 use same farmer's controls as the agricultural production function in Table 3.

Column 2 estimates a parsimonious model without farmer characteristics and district fixed effects In contrast, column 3 adds to the baseline regression indicators of use of other inputs (such as fertilizer, manure and improved seeds). Column 4 further expands this specification by adding an array of heterogeneous trends. We include the interaction of time trends with indicators of ecological zone, proximity to coast and to region capitals. This last 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 dummies of: (1) proximity to an active mine, and (2) proximity to a future mine, but not to an active one. Future mines include sites that started operations after 2005 or have not started production yet but are in the stage of advanced exploration or development.[70] 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.

Columns 6 and 7 report the baseline 2SLS results, using the cumulative measure and the dummy for households within 20km, excluding households in the vicinity of Obuasi mine. As discussed in Section 2, Obuasi mine operations were of a sizable magnitude before the period of interest. It follows that here we only exploit variation in the significant expansion of production and number of mines between 1997 and 2005. The coefficients of interest remain negative, significant and of a similar magnitude, alleviating concerns that results were driven by an outlier.[71]

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 $A_{ivt} = \exp(\gamma S_{vt} + \phi Z_i + \delta_d + \psi_t + \theta mining\_area_v)$, 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.

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

[71]This finding also holds for OLS and other specifications discussed in Table 3.

Table 7: Alternative specifications

	ln(real agricultural output)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cumulative gold prod. within 20 km	-0.199*	-0.158**	-0.163*	-0.166*		-0.162	
	(0.101)	(0.064)	(0.084)	(0.087)		(0.101)	
Cumul. gold prod. within 20 km × born here	0.049						
	(0.053)						
Within 20 km of active mine × GLSS 5					-0.800***		-0.513*
					(0.280)		(0.259)
Within 20 km of future mine × GLSS 5					0.441		
					(0.435)		
ln(land)	0.631***	0.697***	0.599***	0.603***	0.630***	0.676***	0.678***
	(0.038)	(0.041)	(0.039)	(0.039)	(0.038)	(0.047)	(0.047)
ln(labor)	0.209***	0.200***	0.207***	0.206***	0.212***	0.359***	0.358***
	(0.033)	(0.033)	(0.032)	(0.034)	(0.031)	(0.111)	(0.108)
Farmer's control	Yes	No	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	No	Yes	Yes	Yes	Yes	Yes
Other inputs	No	No	Yes	Yes	No	No	No
Heterogeneous trends	No	No	No	Yes	No	No	No
Sample	All	All	All	All	All	Excl. Obuasi	Excl. Obuasi
Observations	1,627	1,627	1,627	1,627	1,627	1,580	1,580
R-squared	0.445	0.332	0.464	0.465	0.454	0.432	0.435

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 includes an interaction term of the measure of mining activity and "born here", an indicator equal to 1 if farmer resides in the same village where she was born. Column 2 includes only region and time fixed effects, without farmer's controls nor district fixed effects. Column 3 replicates baseline regressions but includes indicators of use of other inputs, such as fertilizers, manure and improved seed. Column 4 add to the previous column the interaction of time trends with indicators of ecological zone, proximity to coast, and proximity to region capitals. In column 5, "active mines" are mines that had some production in period 1988-2005, 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. Columns 6 and 7 replicate columns 3 and 4 of Table 3) excluding observations in the vicinity of Obuasi mine. Columns 1 to 5 are estimated using OLS; while columns 6 and 7 are estimated using 2SLS.

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 cumulative gold production is associated to lower productivity.

Table 8: CES function

Parameter	Estimate	S.E.
γ	-0.165**	0.083
λ	0.911***	0.052
ρ	-0.787***	0.228
η	0.997***	0.005
Implied σ	0.560	

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

6 Effects on poverty

The standard consumer-producer household framework presented above links a household's utility function, that depends on consumption levels, to income from agricultural production. As a consequence, we expect that our previous results indicating a sizeable reduction in agricultural productivity and output imply a knock-on effect on local living standards, such as measures of poverty. There are reasons to believe that this channel can be averted. 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:

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

where *poverty* is an indicator of the household being poor, and W_i is a set of household controls.[72] The rest of the specification is similar to equation (2).[73] 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 A.4 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. Since 1997, however, poverty increased in mining areas and they become poorer than non-mining areas.[74] Note that this increase in poverty parallels the reduction in agricultural output (see Figure A.3).

Table 9 presents the estimates of equation (3) using poverty as the outcome variable.[75] 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. [76] We also check the robustness of the results to using a continuous measure of real household expenditure (see table B.8 in the Appendix).[77]

[72] 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.

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

[74] Recall that during this period, gold production reached higher levels and the number of mines increased.

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

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

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

The picture that emerges is similar to the one observed in Figure A.4. There a positive and significant relation between mining activity and poverty. The magnitude of the effect is sizable: the increase in gold production between 1997 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.[78]

The reduction in indicators of economic well-being is consistent with the decline in agricultural productivity in areas where farming activities are the main source of livelihood. Table B.9 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 and positive spillovers from mines, if any, have been insufficient to offset the negative shock to agricultural income.

Table 9: Mining and poverty

	Poverty					
			Rural			Urban
	All households		All	Farmers	Non-farmers	
	(1)	(2)	(3)	(4)	(5)	(6)
Cumul. gold prod. within 20 km.	0.059***		0.071***	0.056**	0.084**	0.054
	(0.015)		(0.019)	(0.021)	(0.032)	(0.036)
Within 20 km of mine × GLSS 5		0.186***				
		(0.055)				
Observations	5,527	5,527	3,393	2,540	853	2,134
R-squared	0.212	0.216	0.227	0.237	0.224	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 survey 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 include an indicator of being within 20 km of a mine.

[78] 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.

7 Concluding remarks

This paper examines an important externality that modern industries may impose on rural areas, namely reduction on agricultural productivity. We find robust evidence that agricultural productivity has decreased in mining areas. The reduction is economically significant: around 40% between 1997 and 2005. This effect seems to be driven by environmental pollution, not by competition for inputs, such as labor or land. We also document an increase in rural poverty associated to the decline in agricultural productivity

These findings have an important implication for environmental and industrial policies. In particular, they suggest that environmental asessments should consider the possible impact of polluting industries on agricultural productivity and farmers' income.

These potential costs are usually neglected. For instance, in the case of extractive industries, the policy debate usually focuses on the benefits they could bring in the form of jobs, taxes or foreign currency. These benefits are weighted against environmental costs such as loss of biodiversity, or human health risks. However, local living standards may be also directly affected by the reduction in agricultural productivity. In fertile rural environments, these costs may offset the benefits from extractive industries, and hinder the ability to compensate affected populations. In turn, this may have substantial re-distributive effects.

A simple back of the envelope using the Ghanaian case illustrates this argument. In 2005, mining-related revenues amounted to US$ 75 millions, which represent around 2-3% of total government revenue.[79] Most of this revenue (around 80%) was channeled to the central government.[80] In contrast, the average annual loss by farming households in mining areas, according to our main results, is in the order of US$ 97 millions.[81] 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 paper documents the negative effects of modern industries on agriculture and contrasts

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

[80] Local authorities (such as District Assemblies, Stools and Traditional Authorities) receive only 9% of mining royalties.

[81] This number is obtained by multiplying the number of producing households in mining areas, around 210,000, to the average reduction in households' annual consumption, i.e., US$ 460.

two possible channels: competition for inputs vs. pollution externalities.. However, a main limitation is that we cannot distinguish the relative importance of several plausible mechanisms through which pollution could affect productivity —such as effects on human and crops' health, quality of soil, or crops growth. While beyond the reach of this paper, examination of these issues warrant further research.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer, "Structural identification of production functions," MPRA Paper 38349, University Library of Munich December 2006.

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*, 1999.

Benjamin, Dwayne, "Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models," *Econometrica*, 03 1992, *60* (2), 287–322.

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," *American Economic Journal: Applied Economics*, 2013, *5* (1), 208–238.

___ **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* (D15), 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.

___ **, Joshua S Graff Zivin, Jamie Mullins, and Matthew J Neidell**, "What Do We Know About Short and Long Term Effects of Early Life Exposure to Pollution?," Technical Report, National Bureau of Economic Research 2013.

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," *The Review of Economics and Statistics*, 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.

Foster, Andrew, Emilio Gutierrez, and Naresh Kumar, "Voluntary Compliance, Pollution Levels, and Infant Mortality in Mexico," *The American Economic Review*, 2009, *99* (2), 191–197.

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, "Land Acquisition for Industrialization and Compensation of Displaced Farmers," *Journal of Development Economics*, 2013, (0), –.

Graff Zivin, Joshua S. and Matthew J. Neidell, "The Impact of Pollution on Worker Productivity," *American Economic Review*, December 2012, *102* (7), 3652–3673.

____ **and** ____ , "Environment, Health and Human Capital," 2013. unpublished.

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 August 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.

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.

Kurtenbach, Ralf, Jörg Kleffmann, Anita Niedojadlo, and Peter Wiesen, "Primary NO_2 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.

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," in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave McMillan, 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 December 1995.

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

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.

A Additional figures

Figure A.1: Location of active gold mines

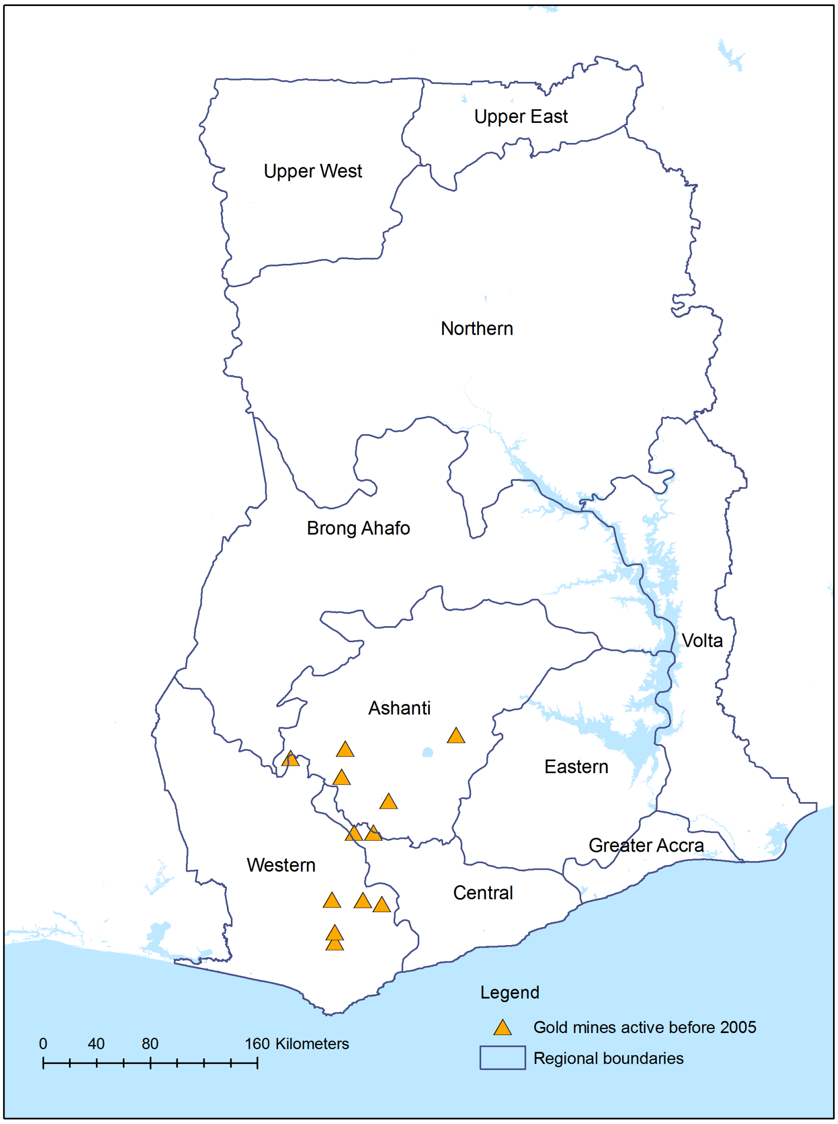


Figure A.2: Area of study and enumeration areas

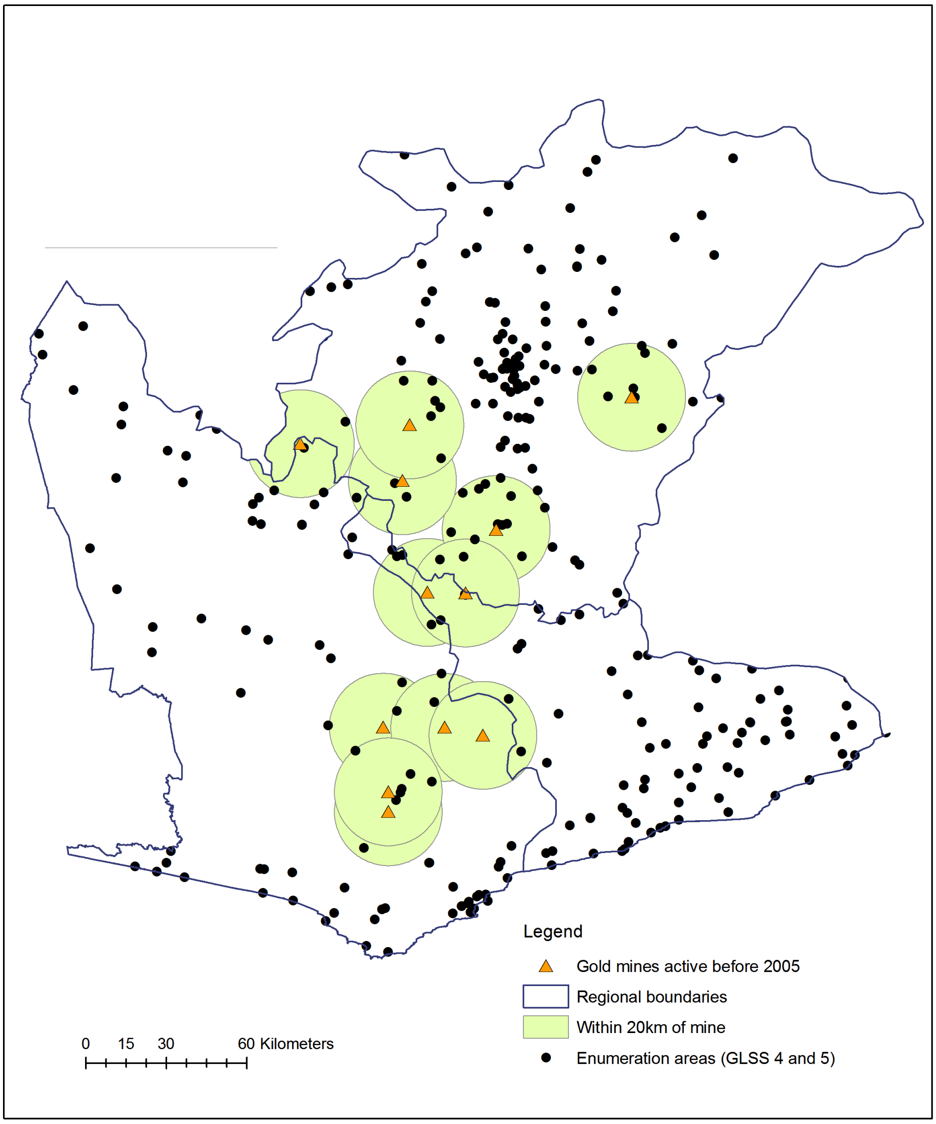


Figure A.3: Evolution of the unconditional mean of ln(real agricultural output)

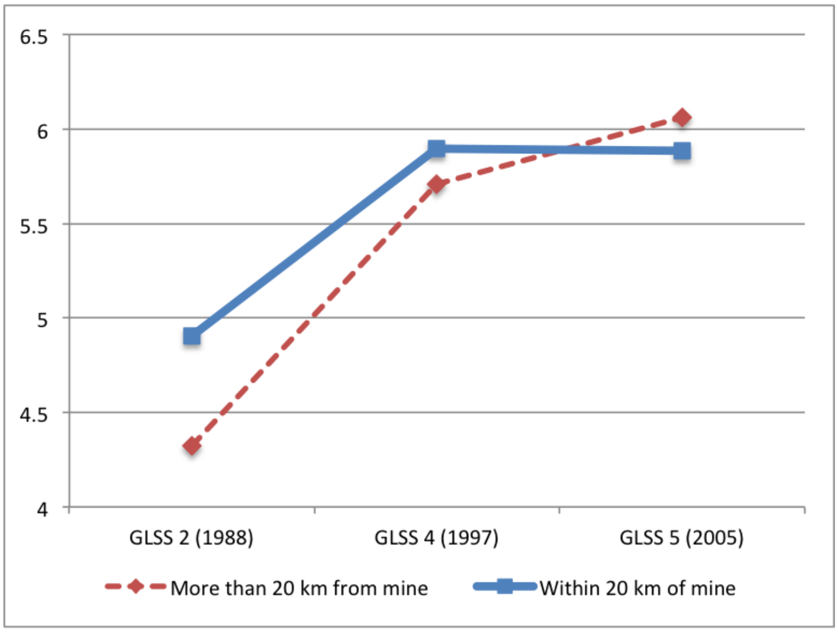
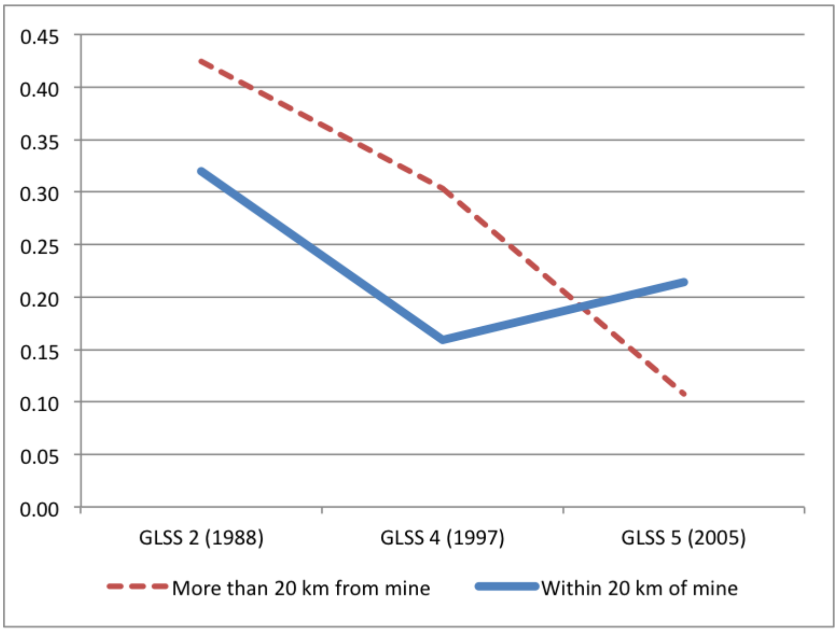


Figure A.4: Evolution of poverty headcount



B Additional results

Table B.1: Evolution of agricultural output in mining vs non-mining areas

	ln(real agricultural ouput)	
	(1)	(2)
Within 20 km of mine × GLSS 4	-0.261	
	(0.370)	
Within 20 km of mine × GLSS 5		-0.515*
		(0.256)
Sample	GLSS 2 and 4	GLSS 4 and 5
Estimation	OLS	OLS
Farmer's controls	Yes	Yes
Controlling for inputs	No	No
Observations	1,473	1,627
R-squared	0.251	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.

Table B.5: Imperfect instruments with multiple endogenous variables

$(\lambda_{land}, \lambda_{labor})$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$
(0, 0)	-0.170	0.676	0.352
(0, 0.1)	-0.165	0.657	0.422
(0, 0.2)	-0.152	0.610	0.601
(0, 0.3)	-0.053	0.249	1.967
(0, 0.4)	-0.238	0.921	-0.577
(0, 0.5)	-0.205	0.802	-0.126
(0, 0.6)	-0.197	0.771	-0.009
(0, 0.7)	-0.193	0.757	0.045
(0, 0.8)	-0.190	0.749	0.075
(0, 0.9)	-0.189	0.743	0.095

Continue on next page

Table B.5 – continued from previous page

$(\lambda_{land}, \lambda_{labor})$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$
(0, 1)	-0.188	0.740	0.109
(0.1, 0)	-0.171	0.687	0.344
(0.1, 0.1)	-0.166	0.668	0.413
(0.1, 0.2)	-0.154	0.620	0.590
(0.1, 0.3)	-0.051	0.235	1.998
(0.1, 0.4)	-0.236	0.928	-0.539
(0.1, 0.5)	-0.205	0.813	-0.115
(0.1, 0.6)	-0.197	0.782	-0.004
(0.1, 0.7)	-0.193	0.768	0.047
(0.1, 0.8)	-0.191	0.760	0.077
(0.1, 0.9)	-0.190	0.755	0.096
(0.1, 1)	-0.189	0.751	0.110
(0.2, 0)	-0.173	0.702	0.335
(0.2, 0.1)	-0.168	0.683	0.402
(0.2, 0.2)	-0.155	0.634	0.575
(0.2, 0.3)	-0.047	0.215	2.045
(0.2, 0.4)	-0.234	0.937	-0.491
(0.2, 0.5)	-0.205	0.826	-0.102
(0.2, 0.6)	-0.197	0.797	0.002
(0.2, 0.7)	-0.194	0.783	0.051
(0.2, 0.8)	-0.192	0.775	0.079
(0.2, 0.9)	-0.190	0.770	0.097
(0.2, 1)	-0.190	0.766	0.110
(0.3, 0)	-0.175	0.723	0.322
(0.3, 0.1)	-0.170	0.703	0.386
(0.3, 0.2)	-0.158	0.653	0.553
(0.3, 0.3)	-0.040	0.183	2.120

Continue on next page

Table B.5 – continued from previous page

$(\lambda_{land}, \lambda_{labor})$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$
(0.3, 0.4)	-0.231	0.949	-0.431
(0.3, 0.5)	-0.205	0.845	-0.085
(0.3, 0.6)	-0.198	0.816	0.011
(0.3, 0.7)	-0.195	0.803	0.055
(0.3, 0.8)	-0.193	0.795	0.081
(0.3, 0.9)	-0.192	0.790	0.098
(0.3, 1)	-0.191	0.786	0.110
(0.4, 0)	-0.178	0.753	0.303
(0.4, 0.1)	-0.173	0.734	0.362
(0.4, 0.2)	-0.161	0.683	0.521
(0.4, 0.3)	-0.028	0.120	2.264
(0.4, 0.4)	-0.228	0.964	-0.351
(0.4, 0.5)	-0.205	0.870	-0.060
(0.4, 0.6)	-0.199	0.843	0.023
(0.4, 0.7)	-0.196	0.831	0.062
(0.4, 0.8)	-0.194	0.823	0.085
(0.4, 0.9)	-0.193	0.818	0.100
(0.4, 1)	-0.192	0.815	0.111
(0.5, 0)	-0.183	0.802	0.272
(0.5, 0.1)	-0.178	0.783	0.324
(0.5, 0.2)	-0.167	0.732	0.466
(0.5, 0.3)	0.004	-0.048	2.651
(0.5, 0.4)	-0.223	0.985	-0.241
(0.5, 0.5)	-0.206	0.908	-0.025
(0.5, 0.6)	-0.201	0.884	0.041
(0.5, 0.7)	-0.198	0.873	0.072
(0.5, 0.8)	-0.197	0.867	0.091

Continue on next page

Table B.5 – continued from previous page

$(\lambda_{land}, \lambda_{labor})$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$
(0.5, 0.9)	-0.196	0.862	0.103
(0.5, 1)	-0.195	0.859	0.111
(0.6, 0)	-0.191	0.893	0.215
(0.6, 0.1)	-0.188	0.879	0.250
(0.6, 0.2)	-0.180	0.836	0.353
(0.6, 0.3)	0.369	-1.959	7.055
(0.6, 0.4)	-0.215	1.016	-0.079
(0.6, 0.5)	-0.206	0.969	0.034
(0.6, 0.6)	-0.203	0.953	0.071
(0.6, 0.7)	-0.202	0.946	0.089
(0.6, 0.8)	-0.201	0.941	0.100
(0.6, 0.9)	-0.200	0.938	0.107
(0.6, 1)	-0.200	0.936	0.113
(0.7, 0)	-0.214	1.129	0.067
(0.7, 0.1)	-0.216	1.139	0.047
(0.7, 0.2)	-0.222	1.177	-0.022
(0.7, 0.3)	-0.170	0.862	0.554
(0.7, 0.4)	-0.204	1.066	0.182
(0.7, 0.5)	-0.207	1.085	0.146
(0.7, 0.6)	-0.208	1.093	0.132
(0.7, 0.7)	-0.209	1.097	0.125
(0.7, 0.8)	-0.209	1.099	0.120
(0.7, 0.9)	-0.210	1.101	0.117
(0.7, 1)	-0.210	1.102	0.115
(0.8, 0)	-0.402	3.079	-1.160
(0.8, 0.1)	-0.597	4.768	-2.774
(0.8, 0.2)	0.364	-3.591	5.213

Continue on next page

Table B.5 – continued from previous page

$(\lambda_{land}, \lambda_{labor})$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$
(0.8, 0.3)	-0.113	0.562	1.245
(0.8, 0.4)	-0.182	1.160	0.674
(0.8, 0.5)	-0.209	1.399	0.446
(0.8, 0.6)	-0.224	1.528	0.322
(0.8, 0.7)	-0.233	1.609	0.245
(0.8, 0.8)	-0.240	1.664	0.193
(0.8, 0.9)	-0.244	1.704	0.154
(0.8, 1)	-0.248	1.734	0.125
(0.9, 0)	-0.072	-0.347	0.995
(0.9, 0.1)	-0.076	-0.190	1.080
(0.9, 0.2)	-0.084	0.052	1.213
(0.9, 0.3)	-0.096	0.476	1.444
(0.9, 0.4)	-0.124	1.403	1.951
(0.9, 0.5)	-0.235	5.060	3.949
(0.9, 0.6)	0.344	-14.114	-6.529
(0.9, 0.7)	0.046	-4.244	-1.135
(0.9, 0.8)	0.005	-2.881	-0.390
(0.9, 0.9)	-0.012	-2.339	-0.094
(0.9, 1)	-0.020	-2.048	0.065
(1, 0)	-0.124	0.198	0.652
(1, 0.1)	-0.120	0.226	0.757
(1, 0.2)	-0.112	0.281	0.962
(1, 0.3)	-0.089	0.435	1.539
(1, 0.4)	0.379	3.546	13.184
(1, 0.5)	-0.197	-0.289	-1.170
(1, 0.6)	-0.166	-0.078	-0.381
(1, 0.7)	-0.156	-0.013	-0.137

Continue on next page

Table B.5 – continued from previous page

$(\lambda_{land}, \lambda_{labor})$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$
(1, 0.8)	-0.151	0.019	-0.018
(1, 0.9)	-0.148	0.038	0.052
(1, 1)	-0.146	0.050	0.098

Notes: Table displays estimates used to construct Figure 2.

Table B.2: Main results using time trend as treatment variable

	ln(real agricultural ouput)			ln(yield cocoa)	ln(yield maize)
	(1)	(2)	(3)	(4)	(5)
Within 20 km of mine × GLSS 5	-0.515*	-0.566**	-0.565**	-0.913**	-1.173**
	(0.256)	(0.236)	(0.247)	(0.430)	(0.519)
ln(land)		0.631***	0.678***		
		(0.037)	(0.047)		
ln(labor)		0.210***	0.346***		
		(0.032)	(0.112)		
Estimation	OLS	OLS	2SLS	OLS	OLS
Farmer's controls	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1,627	1,627	1,627	948	605
R-squared	0.223	0.447	0.438	0.344	0.410

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. For further details on control variables and instruments see notes of Table 3.

Table B.3: Main results using official CPI as price deflator

	ln(value agricultural ouput / CPI)		
	(1)	(2)	(3)
Within 20 km of mine × GLSS 5	-0.155*	-0.183**	-0.176*
	(0.085)	(0.085)	(0.088)
ln(land)		0.631***	0.673***
		(0.038)	(0.048)
ln(labor)		0.202***	0.358***
		(0.033)	(0.114)
Estimation	OLS	OLS	2SLS
Farmer's controls	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes
Observations	1,627	1,627	1,627
R-squared	0.243	0.459	0.447

Notes: Robust standard errors in parentheses. Standard errors are clustered at district level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. For further details on control variables and instruments see notes of Table 3. CPI is the consumer price index reported by GSS. This index has a lower geographical resolution than the price index used in the paper's main results.

Table B.4: First stage regressions of Column 3 in Table 3

	ln(land)	ln(labor)
	(1)	(2)
ln(land owned)	0.917***	0.172***
	(0.027)	(0.038)
ln(nr adult equivalents)	0.024	0.475***
	(0.019)	(0.056)
F-test excl. instruments	781.6	76.8
Observations	1,627	1,627
R-squared	0.798	0.243

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 and survey fixed effects, an indicator of being within 20 km of a mine, and farmer's characteristics. See Table 3 for details on the second stage.

Table B.6: Mining and agricultural productivity - IIV approach assuming only one imperfect instrument

	ln(real agricultural ouput)	
	(1)	(2)
Cumulative gold prod. within 20 km.	[-0.180, -0.164]	[0.043, -0.097]
	(-0.189, -0.149)	(0.109, -0.125)
ln(land)	[0.740, 0.676]	[0,198, 0.676]
	(0.774, 0.616)	(-0.026, 0.773)
ln(labor)	[0.109, 0.352]	[0.652, 0.352]
	(-0.019, 0.576)	(0.793, 0.291)
Estimation	IIV	IIV
Farmer's controls	Yes	Yes
District fixed effects	Yes	Yes
Imperfect IV for:	Labor	Land
Valid IV for:	Land	Labor
Observations	1,627	1,627

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, an indicator of being within 20 km of a mine, and farmer's controls. For further details see notes of Table 3. Columns 1 and 2 identify parameter bounds using the imperfect instrumental variable approach in Nevo and Rosen (2010) assuming there is only one imperfect instrument. 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.

Table B.7: Mining, input prices and input demands - robustness checks

	ln(relative wage)	ln(relative land rent)	ln(labor)	ln(land)
	(1)	(2)	(3)	(4)
Annual gold prod. within 20 km	0.220	1.326		
	(0.627)	(1.558)		
Cumulative gold prod. within 20 km.			-0.123**	0.008
			(0.055)	(0.035)
ln(relative wage)			-0.101	0.013
			(0.157)	(0.119)
ln(relative land rent)			-0.097	0.000
			(0.070)	(0.039)
ln(nr. adult equivalents)			0.507***	0.008
			(0.061)	(0.020)
ln(land owned)			0.064**	0.867***
			(0.029)	(0.041)
ln(real agric. output)			0.095***	0.067***
			(0.025)	(0.021)
Farmer's controls	No	No	Yes	Yes
District fixed effects	No	No	Yes	Yes
Observations	194	201	1,342	1,342
R-squared	0.277	0.009	0.279	0.808

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 survey fixed effects and an indicator of being within 20 km of a mine. Columns 1 and 2 use annual instead of cumulative gold production (see notes of Table 5 for details). Columns 3 and 4 replicate results in Table 4 adding a measure of agricultural output as additional control variable.

Table B.8: 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.	-0.055		-0.048	-0.084*	0.041	-0.115
	(0.053)		(0.065)	(0.045)	(0.111)	(0.073)
Within 20 km of mine × GLSS 5		-0.214**				
		(0.102)				
Observations	5,527	5,527	3,393	2,540	853	2,134
R-squared	0.570	0.571	0.489	0.446	0.583	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 include an indicator of being within 20 km of a mine.

Table B.9: 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)	-1.144		-1.099		0.001		0.004**	
	(1.049)		(1.084)		(0.003)		(0.002)	
Within 20 km of mine x 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	0.047	0.039	0.206	0.190	0.047	0.048	0.041	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 survey fixed effects, as well as an indicator of being within 20 km of a mine.