Local Effects of
Artisanal and Industrial Mining

Evidence from Ghana

Martin Guenther

A thesis presented for the degree of
Doctor of Philosophy

Department of Economics
Royal Holloway, University of London
7 March 2020

Declaration

I, Martin Guenther, hereby declare that this thesis and the work presented in it is entirely my own, other than where I have clearly indicated that it is the work of others.

The artisanal mining data, described in chapter 2, has been created in a joint project with 11 data scientists from Microsoft. I conceived the project idea, managed the process and guided image labelling, while all technical work on designing and implementing the machine learning detection of artisanal mining sites has been carried out by Microsoft staff.

Signed .... (Martin Guenther)

Date: 7 March 2020

Acknowledgements

I would like to thank my PhD supervisors Andrew Seltzer and Juan Pablo Rud for their guidance and support throughout my time at Royal Holloway. They were always available for any questions, helped me with applications for funding, conferences and work and most of all provided valuable input to my thesis at all stages.

I am also grateful to Ija Trapeznikova and Arnaud Chevalier, who as PhD coordinator and head of department did a great job in supporting my work. Thank you also to the Economics department academic staff for helpful comments and to the admin and IT staff for all their help over the past 3 years.

I thank all participants of the Royal Holloway and Microsoft Hackathon 2017, which resulted in the main data source for my thesis. Support from the Microsoft Azure for Research award is gratefully acknowledged. I am further grateful for the support from the Leverhulme Magna Carta Doctoral Studentship.

Last but not least, I want to thank my family and friends for bearing with me over the past 30 years. A special thank you to my brother Marian for his amazing support at all stages of my studies. Thank you to my parents and to my sister Constanze for always being there for me. Thank you to all my London friends and Danke and Alaaf to my friends from Cologne, Freiburg and Crock. Finally, thank you Lea, for your love and encouragement.

Abstract

I estimate local economic, environmental and health effects of artisanal and small-scale gold (ASM) and industrial large-scale mining (LSM) in Ghana. For that purpose I use a novel dataset of artisanal gold mines created with machine learning techniques and satellite imagery. ASM is an informal, low-tech, but highly labour-intensive form of resource extraction that is typically associated with environmental and health damages, social problems and poverty. In contrast to positive expectations from field research, I find that local incomes and expenditures are not affected by nearby ASM sites, while LSM does contribute to both. Environmental degradation by both ASM and LSM is documented in the form of forest cover loss, which has accelerated in more recent years. Finally, I find child health burdens of ASM in the form of increased malaria infection rates and child mortality and different disease symptoms for LSM. These results shine a new light on the trade-offs of mineral extraction between local economic gains and health and environmental costs.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Many developing countries with high poverty rates are also blessed with abundant mineral resources. This mineral wealth is typically extracted through industrial large-scale mines, which are highly capital intensive and have been shown to provide very few direct employment opportunities for local populations (Chuhan-Pole et al. (2017), Cust and Poelhekke (2015)). In contrast, over the last 15 years smaller mining operations have become increasingly common, particularly in the gold mining sectors of many African, Asian and Latin-American countries. While these artisanal or small-scale mining operations employ millions of workers, they are also linked to environmental damages, health hazards, social costs, poverty and even political tensions (Hilson and McQuilken (2014), Financial Times (2017) and National Geographic (2017)). Because many artisanal mines operate informally, data and thereby empirical evidence on this industry are extremely scarce (Cust and Poelhekke (2015) and World Bank (2015)).

I estimate local economic, environmental and health effects of artisanal and small-scale mining (ASM) in Ghana by using a novel data set covering the geolocation of these mining operations. Ghana is one of the main gold producers in Africa and its ASM sector has seen strong growth over recent years, now accounting for one third of total gold output. The ASM data is generated in collaboration with Microsoft by training a machine learning model to detect artisanal mining sites from satellite images. As a result, I am able to map a large share of artisanal mines in Ghana for a cross-section in 2014. These more than 3,600 detected mines are then matched with data on household, health and environmental outcomes.

By exploiting the aggregate increase in ASM production volumes between 1998 and 2013, I can identify outcome changes in surrounding areas after ASM operations become active. Using household survey data, I find that per-adult income is surprisingly unaffected by nearby ASM sites. This result is robust to controlling for large-scale gold production, a number of individual and household controls and unobserved heterogeneity at the district level. This zero result seems to be partly driven by increases in immigration into ASM areas, which creates a downward pressure on real incomes. In contrast to these results, large-scale mining overall contributes positively to economic outcomes, as measured by income, expenditure and nighttime lights.

Whereas the economic effects of ASM and LSM differ, I find that both contribute to environmental and health costs. Using a grid-level panel approach spanning most of southern Ghana, cells with artisanal mining show an increase of 0.2 percentage points in forest cover loss relative to non-mining cells in years after 2010. This is significant in magnitude as it represents a roughly 30 percent increase on the sample mean and over 3,000 hectares of forest loss per year due to ASM.[1] The effect per cell is even higher for industrial mining, but less precisely estimated due to fewer such mines. Forest conservation is an important goal by itself as forests are a source of local livelihoods, help regulate world climate, filtrate water and protect soils. Loss in forest cover has also been shown to deteriorate child health (Berazneva and Byker, 2017).

From a health perspective a common claim is that (artisanal) gold mining creates breeding spots for disease-carrying mosquitoes, with poor sanitary conditions and excessive chemical use exacerbating the problem (World Health Organization, 2016). Following the grid-level approach I find that malaria prevalence rates are in fact higher in the presence of artisanal mining. Infection rates in Ghana have strongly decreased from 65 to 34 percent from 2000 to 2015, but this reduction has been weaker by roughly one percentage point in areas close to ASM sites. On the other hand, using demographic and health survey data for over 6,600 children between 1992 and 2014 there is no significant effect of mining activity on fever, cough or diarrhoea rates. I also show that large-scale mining, while not related to malaria prevalence, is associated with increased fever and cough symptoms along with overall higher child mortality rates.

The relationship between natural resources and economic outcomes has been studied extensively at the national and sub-national level. A common observation is that many countries remain poor despite being rich in minerals or fossil fuels. The theoretical key mechanisms behind this resource curse include Dutch disease, rising commodity prices and rent-seeking activities.[2] However, empirical studies at the national level lead to divergent results: Gylfason et al. (1999), Leite and Weidmann (1999), Sachs and Warner (1995) and Sachs and Warner (2001) find that national gdp is negatively affected by resource exports, but Brunnschweiler (2008), Lederman and Maloney (2006) and Lederman and Maloney (2008) show that this effect is often not robust or even reverse. These contradictory findings have been linked to the observation that resource exports are endogenous to countries' institutions and political stability (Sala-i Martin and Subramanian (2013) and Bulte et al. (2005)) and that a large mineral sector may actually represent natural resource dependence instead of abundance (Brunnschweiler and Bulte, 2008).

[1]Compared to approximately 38,000 hectares of total forest loss per year for the sample area, ASM constitutes around 10 percent of all FCL.

[2]See Chuhan-Pole et al. (2017) for a review of the resource curse literature.

Going one step further, a more recent literature has started to analyse local impacts of resource extraction and finds that the opening of mines is associated with reduced agricultural productivity (Aragón and Rud, 2016), increased local corruption (Knutsen et al., 2017) and conflict incidence (Berman et al., 2017). These studies focus on large-scale industrial mining (LSM). This is in part because LSM produces the biggest share in mineral output, but also because it is relatively easy to identify the location, opening and closing dates, as well as employment and production figures for these operations.

The increasing importance of artisanal and small-scale mining both in terms of employment and output in many developing countries means that these operations must be taken into account in order to accurately describe the local impacts of resource extraction. To the best of my knowledge to date only two studies investigate this relationship empirically. Bazillier and Girard (2017) make use of ASM register and geological data in Burkina Faso and find a positive effect of ASM on local consumption expenditure in years when the gold price is high. Using machine learning techniques and satellite images Saavedra and Romero (2017) identify illegal artisanal mines in Colombia. While they mainly focus on the political determinants of ASM they also find that illegal mining leads to deteriorated health outcomes.

This study adds to the literature above by creating a dataset on the precise geolocation of artisanal gold mines in Ghana that, in combination with data on industrial mines, can be used to measure the local economic effects of resource extraction. Going beyond economic effects, I also estimate environmental damages in the form of forest cover loss and health costs. With these empirical results I show that the inclusion of artisanal mining is essential to understand the impact of natural resources on development.

This thesis is structured as follows. Chapter 2 provides institutional background on artisanal and industrial mining in Ghana and introduces the machine learning approach to detect ASM sites. This chapter also describes the different data sources used for the empirical analysis and how they are linked to the ASM and LSM data. Chapter 3 analyses how ASM and LSM affect economic outcomes and chapter 4 estimates health and environmental effects. Finally, chapter 5 concludes.

Chapter 2

Detecting artisanal mining in Ghana

2.1 Introduction

Artisanal and small-scale mining (ASM) is reported to employ over one million people in Ghana, but is also blamed with causing environmental and health damages. Attempts to measure these effects to date have relied predominantly on case study data, which are naturally limited in geographic and time scope (Hilson (2010), Rajaee et al. (2015) and Cobbina (2012)). On the other hand, empirical work on the effects of mineral extraction usually abstracts from ASM and focusses solely on industrial mining (Aragón and Rud (2016), Knutsen et al. (2017) and Berman et al. (2017)). Both strands of literature are held back by the unavailability of data on ASM.

This chapter aims to overcome this obstacle, by providing the first accurate mapping of artisanal mining activity in the African context. This mapping is based on the idea that ASM sites can be visually identified by the human eye. The approach makes use of Machine Learning (ML) techniques in combination with over four million high-resolution satellite images. After using secondary data on the location of large-scale mines to correct some predictions by hand, I obtain a map of roughly 3,600 artisanal mining sites in Ghana.[1]

While this novel approach creates a unique dataset on the precise location of a mostly informal industry, some caveats remain. The potentially biggest drawback of the data is that the ML detection images are only available for a cross-section in 2014. To overcome this challenge, I provide two approaches. First, as will be described in the subsequent chapters 3 and 4, I develop an identification strategy that uses secondary data to compare detected ASM areas before and after mines become active. As a secondary approach, I use colour band values from public Landsat satellites to directly identify land use changes within detected ASM sites. This generates years of activity for all ASM sites.

This chapter is structured as follows. Section 2.2 introduces the Ghanaian artisanal

[1] At this point, I want to reiterate that the technical work on designing and implementing the ML detection has been provided by a team of data scientists from Microsoft.

mining context and describes the ML approach to detect ASM sites. This also includes a discussion of how to identity the timing of these sites. Section 2.3 summarises data used as outcome and control variables in subsequent chapters and section 2.4 concludes.

2.2 Detecting artisanal mining

2.2.1 Background: Artisanal mining in Ghana

Mining is a major industry in Ghana, directly contributing 8 percent to GDP and 16 percent to fiscal revenue. Gold accounts for 97 percent of this revenue and makes up 23 percent of all exports (US Geological Service (2014), GHEITI (2015), Simoes and Hidalgo (2011)). These macro figures relate to both large-scale mining (LSM) and artisanal and small-scale mining (ASM). It is important to distinguish between these two industries, as they differ greatly in organisation, technology and labour used. Local economic outcomes are therefore likely to be different. Large-scale mining is mostly operated by large international mining companies, with the Ghanaian government required to hold at least 10 percent of shares (GHEITI, 2015). It is highly capital-intensive and therefore generates only relatively few local employment opportunities (Aryeetey et al., 2007). As of 2014, approximately 12,300 workers, out of a labour force of 11 million people in Ghana, are employed in the large-scale mining sector (US Geological Service, 2014).

Local effects of large-scale mining, using the GPS location and production of large industrial mines, have been found in terms of decreased agricultural productivity (Aragón and Rud, 2016), higher local corruption (Knutsen et al., 2017) and more conflicts (Berman et al., 2017).[2] Some studies also use remote sensing to observe environmental damages in the form of forest loss around large mines (Schueler et al., 2011). The seemingly negligible local economic impact of large-scale mining is in sharp contrast to the artisanal and small-scale mining sector, which in Ghana alone is estimated to employ 1.1 million workers directly (Hilson and McQuilken, 2014).

Artisanal gold mining in Ghana can be dated back to the 6th century, long before European colonisation started in the 15th century (Hilson (2001) and Ofosu-Mensah (2011)).[3] Since the end of the 19th century Ghana's gold sector experienced an increased industrialisation, which led to a decline of traditional forms of gold extraction (Ofosu-Mensah, 2011). Looking at more recent data in figure 2.1 it can be seen that by the beginning of the 1990's the gold production of ASM was only 15,000 ounces (425 Kg).[4] Compared to

[2]For a recent survey of local effects of (large-scale) mineral extraction see Cust and Poelhekke (2015) and Chuhan-Pole et al. (2017).

[3]During colonial years Ghana was known as the Gold Coast under Portuguese and later British rule (Crawford et al., 2015).

[4]ASM here includes both legal and illegal mining. This data is originally based on annual reports from Ghana's Minerals Commission, but made available by Crawford et al. (2015) and Hilson and Potter (2005). It is not clear how data for ASM, especially the illegal type, is collected. However, because both legally

the industrial gold output of 851,000 ounces (24,000 Kg) this represented an ASM sector share of under two percent. In the following years small-scale gold production increased substantially, reaching 145,000 ounces in 2000. The formalisation of the artisanal mining sector from 1986 onwards, which among other things allowed for the legal sale of gold from ASM sources, is seen as one of the main drivers of this rise.[5] The formalisation provided an opportunity to escape unemployment and poverty for many young Ghanaians (Hilson and Potter (2003), Hilson (2004)). Still, in terms of national gold production, the small-scale sector remained negligible with a share of under ten percent until 2004.

The early 2000's marked the beginning of the worldwide commodities super cycle. From 2001 to 2012 the international gold price rose from 280 to over 1,600 USD per ounce. This increase in value in combination with increasing mechanisation and an influx of foreign capital and labour made ASM in Ghana much more profitable and productive (Mantey et al. (2016), Crawford et al. (2015)). Consequently, small- and large-scale gold production increased strongly (see figure 2.1). In 2014 artisanal gold mining had reached a national share of 36 percent at 1.5 million ounces production. This second, stronger increase in artisanal and small-scale mining is the main focus of this thesis.

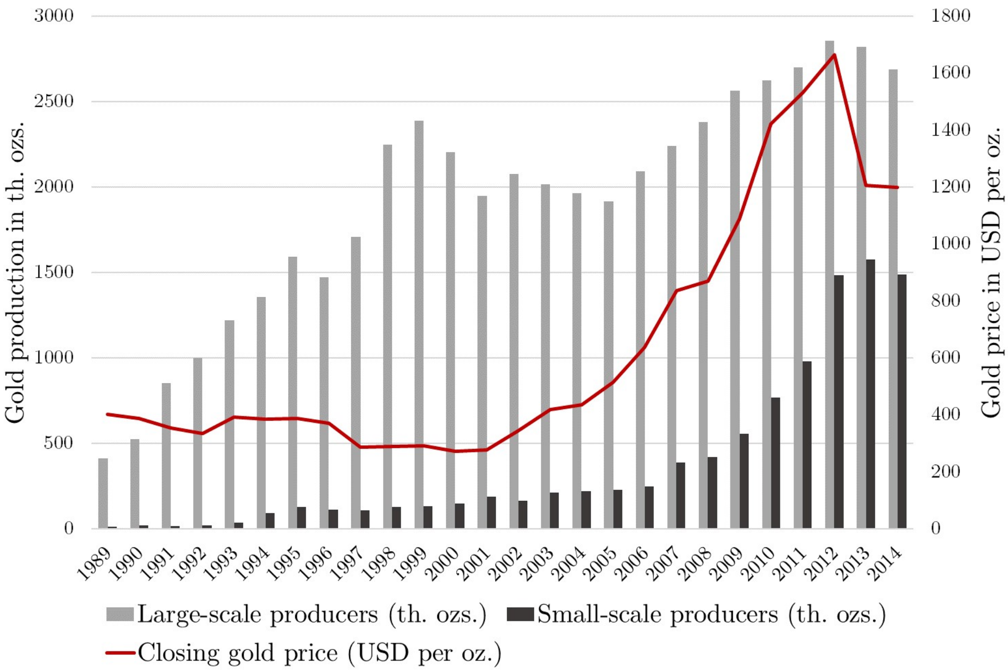


Figure 2.1: Gold production in Ghana and international gold price

Source: Author's visualisation based on data from Crawford et al. (2015), Hilson and Potter (2005).

and illegally produced ASM gold can be sold to the Precious Minerals Marketing Company (PMMC), the production data is likely based in their estimates. So while the ASM data here includes both legal and illegal mining it may still underestimate the true amount if gold is illegally exported instead of sold to the PMMC.

[5]The legalisation of SSM was part of the Minerals and Mining Law 1986 (PNDCL 153) and further implemented in Mercury Law (PNDCL 217), Small-Scale Gold Mining Law (PNDCL 218), and Precious Minerals and Marketing Law (PNDCL 219) in 1989, see Hilson and Potter (2005).

Artisanal mining in Ghana mainly takes place in the form of unlicensed operations, called "galamsey" (Hilson (2002) and Teschner (2012)). It is estimated that at least 70 to 85 percent of ASM sites operate without a license (Crawford et al. (2015) and Teschner (2012)). This is despite the fact that ASM has been legalised since the late 1980s. Official licenses can be acquired for areas of up to 25 acres (100,000 square metres), but the process is described as prohibitively lengthy, bureaucratic and expensive (Hilson et al., 2014). It is therefore not surprising that most ASM companies choose to operate without a license.

There is a sharp contrast between how ASM is perceived both in public and in the academic literature. Public perception is largely driven by reports on environmental damages, such as mercury contamination of soil and water (Mantey et al. (2016), Van Straaten (2000)) and deforestation (Hirons (2011), Rahm et al. (2015) and Schueler et al. (2011)), as well as health hazards and numerous social problems, like child labour, prostitution, drug abuse, corruption and violence (World Health Organization (2016) and Basu et al. (2015)). As a result, ASM operations are often blamed to cause extreme poverty (Mantey et al. (2016), Hilson and McQuilken (2014)). Others argue that many of the detrimental outcomes result from the illegality of most ASM operations but are not inherent to ASM itself. Thereby, formalisation of ASM and harmonisation with large-scale activities would mitigate many negative effects. Proponents of ASM claim that instead of causing poverty ASM has developed into a major industry in many developing countries, accounting for more than one million jobs in Ghana directly, and over four million indirectly through downstream industries (Hilson, 2016).

This thesis adds to this discussion by providing evidence on ASM's local economic, health and environmental effects. The major obstacle for this analysis is that due to its illegality, data on ASM is extremely scarce. For the relatively few legal ASM operations licensing data is available (see section 2.3.3). However, this data cannot be exactly geolocated to measure local effects and is further substantially underrepresenting the size of illegal ASM activities, which again are assumed to make up most of Ghana's small-scale mining (Hilson and McQuilken, 2014).

Many studies have looked at interview and case-study evidence to measure the effects of ASM, but lack consistent measures and the scope to draw inferences on a larger scale[6]. Only recently more credible attempts that employ low-level ASM data have been made. Asner et al. (2013) identify forest loss through satellite images that is caused by both industrial and artisanal mining in the Amazon region. Bazillier and Girard (2017) combine official ASM registers and geological data to infer the location of legal small-scale mines

[6]For ASM in Ghana, see for example Rajaee et al. (2015) and Long et al. (2015) on health effects, Cobbina (2012) on environmental effects and Hilson (2010) and Owusu and Dwomoh (2012) on social outcomes. See further (Fisher et al. (2009) and Kitula (2006) on ASM and environmental and socio-economic effects in Tanzania.

and estimate their effect.[7] They find that higher gold prices increase wealth of households near artisanal mines. The most similar attempt in terms of data generation to this article is conducted by Saavedra and Romero (2017). They create a panel of ASM locations in Colombia using satellite images and machine learning techniques to show how a reform on tax revenue distribution from local to central governments increases illegal mining activity. While they report some negative health effects of illegal mining, they do not cover household income or expenditure. Therefore it remains still unclear (1) how these results translate to the context of mostly illegal ASM in Ghana and (2) how effects on individual households and the environment compare.

2.2.2 Generating artisanal mining data

The main artisanal and small-scale mining data for this thesis has been generated in a joint project with Microsoft. The basic idea is that a machine learning model, based on a convolutional neural network, is trained to identify mines on satellite images from Ghana.[8] These identified mines can then be mapped against other georeferenced data to draw inferences about the effects of ASM. The latest release of Bing Maps for Ghana from 2014 is used for this project. Satellite images for previous years are not available from Bing Maps, so it is unfortunately impossible to construct a panel that way.

An alternative ASM data source is official mining register data, which is used in the robustness section and described below. There are however several advantages of using satellite-based over register data. First, all visible mining activities regardless of legal status are included. This is especially important given the large number of informal ASM operations in Ghana (see background section 2.2.1). That way a more comprehensive mapping of ASM activity is achieved. Second, this approach circumvents the selection issue of which firms or individuals obtain licenses. Licensed and unlicensed operators likely differ in a number of characteristics, such as financial assets and political connectedness. By including both legal and illegal mining an average effect can be measured. Finally, the high resolution of the satellite images of 153 by 153 metres allows for a very accurate localisation of the individual mining sites (see below). By comparison, the best available geographic matching of license data is at the district level. A notable caveat of the satellite-based approach is that underground mining is excluded, which is one reason why license data is used as a robustness test.

Both industrial and artisanal gold mining are confined to areas that fulfil certain geological prerequisites (Hilson, 2001). Birimian and Tarkwaian are the two major rock groups in Ghana that can potentially contain gold belts. Access to the hardrock deposits

[7]Different to Ghana, artisanal and small-scale mining in Burkina Faso is mostly carried out with legal extraction licenses, see also Jaques et al. (2006).

[8]For an intuitive introduction on convolutional neural networks see https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets

of these belts is typically limited to highly capital-intensive industrial mines. However, erosion has caused displacement of parts of these rocks to as so-called placer or alluvial gold deposits in streams, flood plains and terraces of Ghana's major rivers Offin, Pra, Ankobra, Birim and Tano. Figure A.4 in the appendix shows the major rivers in Ghana, which is where alluvial gold mining is possible. The possible locations of ASM are therefore to a large degree dictated by the geological setting.

The most common forms of ASM, alluvial and surface mining, can be spotted by the human eye on high-resolution satellite images (see figure 2.2).[9] The distinctive visual features of surface ASM sites are described in Mantey et al. (2016), Rahm et al. (2015) and Unitar (2016). These are pits of unnatural shape, often found next to rivers, which have changed to a light-brown colour. When abandoned, but also to some degree during production, these pits fill with rain or river water and remain highly visible over a long time. The detected sites can therefore be either active or past ASM operations. While this reduces the accuracy of the ASM measure, it also allows to detect mining sites that have already been exhausted with a single cross section of satellite images. Starting from known ASM centres such as Dunkwa-on-Offin and Obuasi, satellite images were then labelled as either "mine", "not mine", or "cloud" by trained users at the Hackfest.

The following describes the software used for the ML detection of ASM sites. The image recognition model is designed in Python using Keras[10] with a Tensorflow model.[11] Microsoft Azure is used as the main platform for storing of data and running of the ML detection.[12] [13]

The validity of this labelling process is verified afterwards by comparing the results to a smaller sample of georeferenced ASM data by Mantey et al. (2016). Notably, the image detection process in this expert-knowledge based approach does not depend on defining which features are important. Instead, based on the true labelled training dataset, through repeated filtering of layers and backpropagation the best fit is achieved.

Most of the Southern half of Ghana is included in this approach. Other regions are excluded because of high computational costs and because we could not visually identify any small-scale surface mining sites in the Northern half of the country during the training phase. This is not a large restriction to the analysis, as the majority of artisanal surface mining activity takes place in this region, see also figure A.4. This results in roughly four million satellite image tiles at a resolution of 0.6 metres per pixel, with each tile

[9] Other, less frequent, forms of artisanal mining not covered in this analysis include underground mining, mill-house operation and pilfering mining. Note that there are many different forms of surface mining, which cannot be distinguished by the approach in this article.

[10] https://keras.io/

[11] https://www.tensorflow.org/

[12] https://azure.microsoft.com/en-gb/

[13] All Python codes and background information can be found at https://github.com/noodlefrenzy/ImageRecognitionInKeras and https://github.com/noodlefrenzy/StrataDataLondon2018.

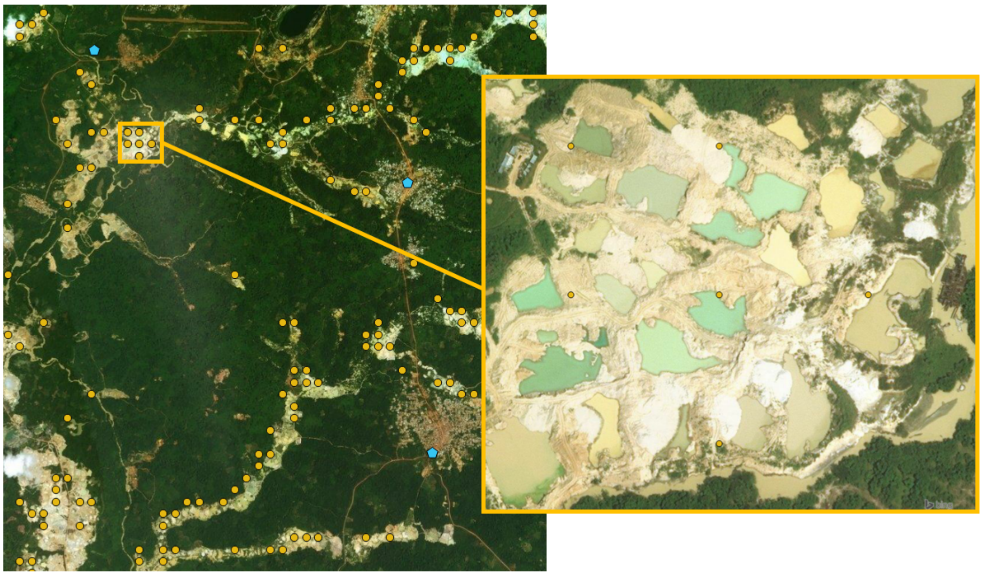


Figure 2.2: Detected ASM sites and household locations in Ghana

Notes: Detected artisanal mines in yellow, GLSS household locations in blue.
Source: Microsoft, Bing Maps, GLSS, author's visualisation.

spanning 152.9 by 152.9 metres (equalling 0.02 square kilometres of surface). The area covered by the detection model is depicted in figure 2.3. The training data set consists of 14,887 images with 907 labelled as mine, 12,671 as not mine and another 1,309 as cloud. Figure 2.4 shows the confusion matrix for the validation set, based on part of the training data.[14] The chance of false positives for "mine" can be derived by dividing the sum of the off-diagonal elements of column 2 (0+9) by the total sum of column 2 (0+25+9). The probability of false positives is therefore 26.5 percent. The probability of false negatives can be calculated similarly by the second row and is 7.4 percent. By construction the model is designed to produce more false positives in favour of fewer false negatives. This is because it is much easier to detect and correct false positives.[15]

After the machine-learning based image detection, I increase the accuracy by removing obvious false positives by hand. These falsely classified mines can be towns, forests or sometime large-scale mines, all of which are easy to distinguish from artisanal mines. Examples of such incidences are shown in the appendix figures A.2 and A.3. Large-scale mines that are wrongly predicted to be artisanal mining sites are identified through the

[14] A confusion matrix is a metric for the accuracy of the model. Elements on the diagonal from top left to bottom right indicate images that are correctly detected by the model. Elements off the diagonal are mispredictions. For more details see https://machinelearningmastery.com/confusion-matrix-machine-learning/.

[15] Figure A.1 in the appendix shows the amount and location of detected ASM sites by model prediction probability threshold. Note that the prediction probabilities of the three classes "mine", "not mine", "cloud" by design sum to 1. From figure A.1 it can be seen that using a 80 or 90 percent probability is clearly not sufficient, as this generates too much noise. I therefore choose 99 percent prediction probability as the threshold for the "mine" class.

collected LSM location data. In this manual process 327 labels are changed from "mine" to "not mine". Note that this step removes false positives, while false negatives, with a probability of occurrence of 7.4 percent (see previous paragraph), remain. These false negatives constitute measurement error in the independent variable, thus leading to attenuation bias.

Thereby 3,639 images are detected to contain a small-scale mine as depicted in figure 2.3. As anticipated, the small-scale mines form different clusters either in bulks or following rivers. Most detected mines lie in the Ashanti region, with some clusters also being in the Central, Eastern and Western regions. These are also the main large-scale mining areas.

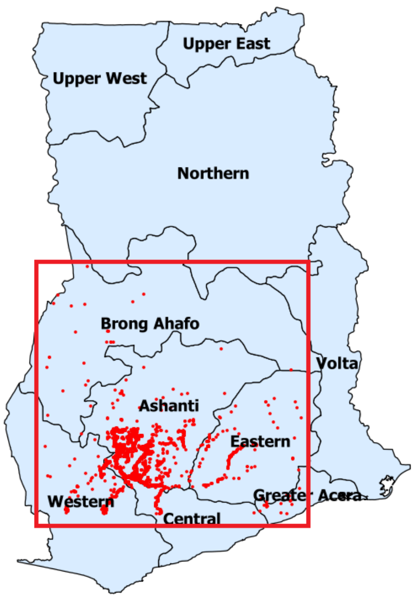


Figure 2.3: Overview of detected mines in Ghana

Notes: Detected artisanal mines in red. The red box shows the covered area by the machine learning prediction. Source: Microsoft, author's visualisation.

To verify whether the detected images in fact contain artisanal mining sites, I conducted field research in Ghana. Although it was not possible to visit many active ASM sites, I was able to confirm some of the detected ASM locations. Figure 2.5 shows images taken at an ASM site near Dunkwa-on-Offin. Some of the mining pits had seemingly been abandoned for some time, while there was also active (illegal) small-scale excavation during my visit.

Another piece of evidence in verifying the ASM data comes from comparing it to the license boundaries of industrial mining, shown in figure 2.6.[16] What is clear from this is that most of the detected ASM sites fall either outside or just on the edge of industrial licenses. This suggests that these are in fact small-scale and not industrial mining sites.

[16]The boundary data is from March 2019, so it may not perfectly coincide with the ASM detection from 2014.

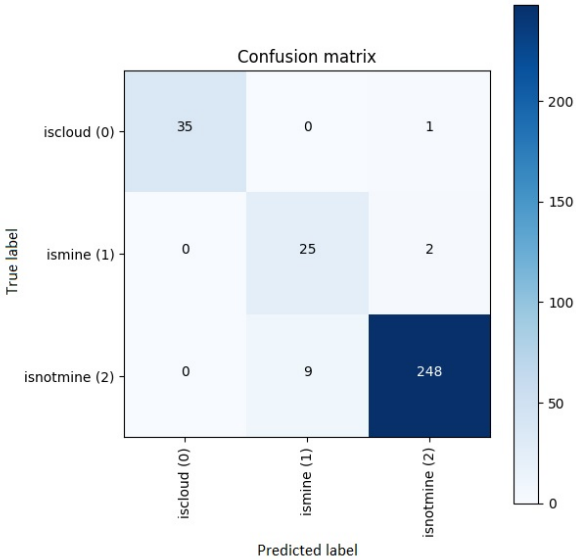


Figure 2.4: Confusion matrix for validation set

Source: Microsoft

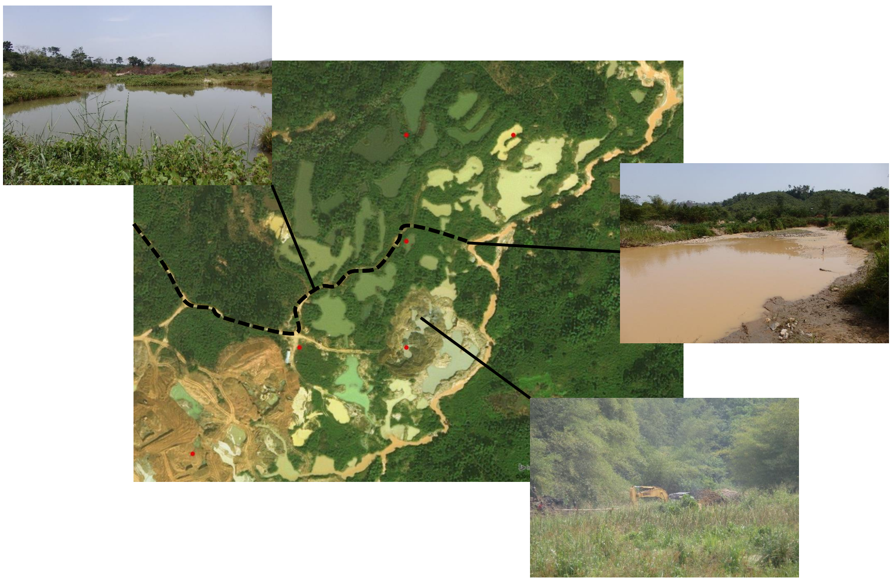
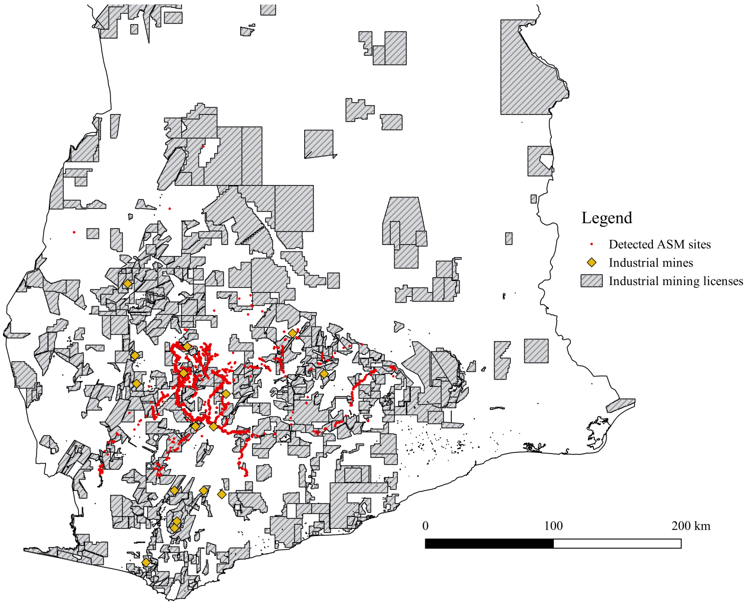
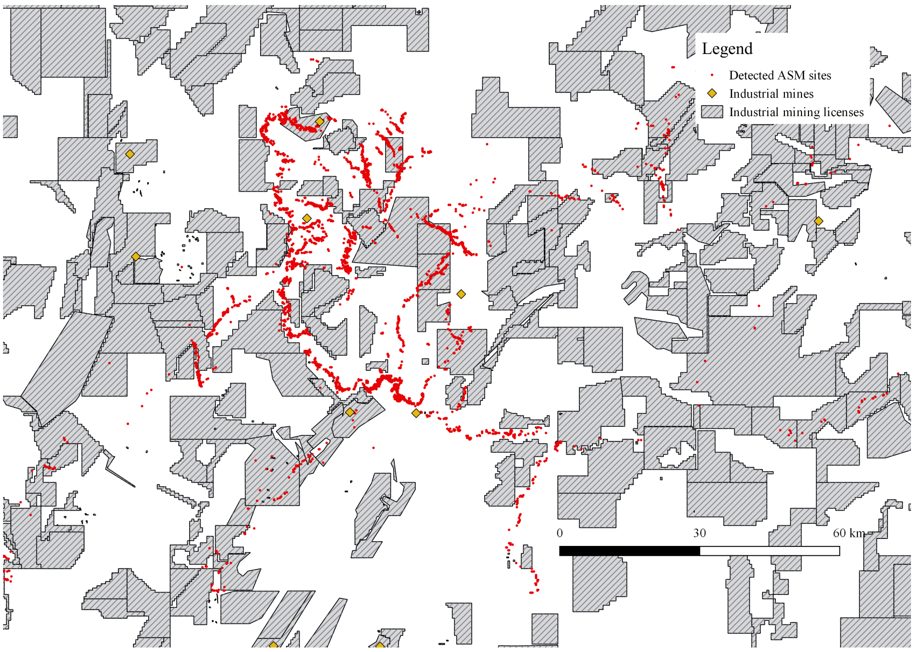


Figure 2.5: ASM site visit in March 2019

Notes: Red dots show detected ASM sites. Photographs come from field research in Ghana during March 2019.



(a) Overview



(b) Zoom

Figure 2.6: Location of ASM sites and LSM boundaries

Notes: Author's visualisation, based on license data from Ghana's Minerals commission and Microsoft.

2.2.3 Time variation of ASM sites

An important limitation of the machine-learning ASM data is that it is based on only a cross section of satellite images. Empirical analysis that aims to go beyond cross-sectional comparison then requires the additional assumption that most of the detected sites only become active in years after 2010 (see chapters 3 and 4). While the development of gold prices and aggregate ASM production as well as the literature support this assumption, there remains uncertainty.

This section introduces a method to infer the timing of the detected ASM sites. The goal is to aid the subsequent analysis in two ways: First to validate the timing assumption stated above by at the amount of active artisanal mines over time; Second to use the activity status of individual ASM sites in a separate robustness analysis.

I use a time series of publicly available satellite images to infer when a given ASM site first becomes visible, i.e. active. For this I choose data from the Landsat 7 programme, administered by U.S. Geological Survey (USGS, 2019) and obtained through Google Earth Engine (2019). This data series runs from April 1999 to today, where each location on Earth is covered every 16 days.

Landsat 7 images contain seven bands, three of those visible, at a 30 by 30 metre resolution, i.e. pixel size. By comparison the pixel size of the Bing Maps images used for the machine-learning detection in the previous section is 60 by 60 centimetres. I use USGS Landsat 7 Surface Reflectance Tier 1 images from 1999 to 2018. The images are already pre-processed by USGS for atmospheric correction (Google Earth Engine, 2019).[17]

A common problem in this region is persistent cloud cover (Coulter et al., 2016). I therefore apply standard cloud-masking provided by Google Earth Engine (2019). This step excludes all images with high cloud confidence, as determined by the Landsat 7 pixel quality layer. As a result, many pixel-year observations are empty. I therefore pool images from a two-year image series. The result is a panel that contains values of bands one to seven for each 30m by 30m pixel for year pairs 1999-2000, 2001-2002 etc. until 2017-2018.[18]

The Landsat pixel values are then combined with the ASM data from the previous section. Recall that the ASM data is based on Microsoft Bing grid cells of roughly 153 by 153 metres (at the equator). I calculate the average values of each Landsat 7 image band within each of the ASM cells. Results for this are illustrated in figures 2.7 and 2.8.

Figure 2.7 shows an example area near the ASM centre Dunkwa-on-Offin. The Landsat 7 images, which are band 1, 2 and 3 composites, are overlaid with the shape of one of the detected ASM sites (blue square). The average amount of reflected surface light

[17]For more details on pre-processing of Landsat data, see Young et al. (2017).

[18]I collect the minimum band values of each pixel as a conservative measure. This is because clouds have high values for all bands.

from band 3 (red) is also reported. In panels (a) and (b) there is no visible evidence of any economic activity. This coincides with low values of reflected red light. In panel (c) however, which shows years 2013-14, there is a clear change in land cover. The average red light value within the detected ASM cell is accordingly much higher. This suggests that high values of red light can be used to identify land use changes. Because this land use change takes place within the previously detected ASM sites, which are narrow in size, it is very likely that it reflects the start of mining activity.

Figure 2.8 shows the average for band 3 (red light) for all ASM and non-ASM cells from 1999-2017. As suggested by the previous figure 2.7, the band 3 values in 1999 to 2009 are relatively low compared to 2011 and beyond. Moreover, ASM and non-ASM cells follow a very similar time trend until 2009, but from 2011 ASM cells exhibit significantly higher values. The non-ASM cells meanwhile remain essentially flat. This suggests a land use change within the ASM cells, which again is most likely related to mining activity.

The findings thereby support the ASM timing assumption. It now seems clear that most detected ASM sites only become active after 2010. To further make use of these insights, I use a uniform decision rule to determine individual mining site activity. Such a rule should allow for annual changes in vegetation conditions, which would affect both ASM and non-ASM areas. To also be consistent between years, I use as a rule that a mine becomes active if it is one standard deviation above the non-ASM cell long-term average.[19] Figure 2.9 shows the corresponding number of active ASM sites.

2.3 Outcome and control variables

2.3.1 Socio-economic and health data

I use two different surveys to estimate effects of ASM and LSM on economic and health outcomes. The Ghana Living Standards Survey (GLSS) provides data on economic development. To analyse health effects I use data on child diseases and child mortality from the Demographics and Health Survey (DHS). These data sources are described in more detail below.

The regionally representative Ghana Living Standards Survey (GLSS), produced by the Ghana Statistical Service (GSS), provides three waves that are suitable for this analysis: GLSS 4 (1998/99), GLSS 5 (2005/06) and GLSS 6 (2012/13). Together they sum to a total of 29,459 households and 124,170 individuals. The clusters in GLSS are drawn randomly within the ten regions of Ghana for each survey year. This means that the clusters do not necessarily overlap between survey years but only provide a repeated cross-section.

[19]The results are robust to different thresholds of the decision rule, for example using half of a standard deviation.

(a) 2001-2002, average red light = 446

(b) 2007-2008, average red light = 177

(c) 2013-2014, average red light = 1602

Figure 2.7: Illustration of colour change

Notes: Author's visualisation, taken from Google Earth Engine (2019). Blue square is one of 3639 detected ASM sites. Average red light values refer to the ASM site average.

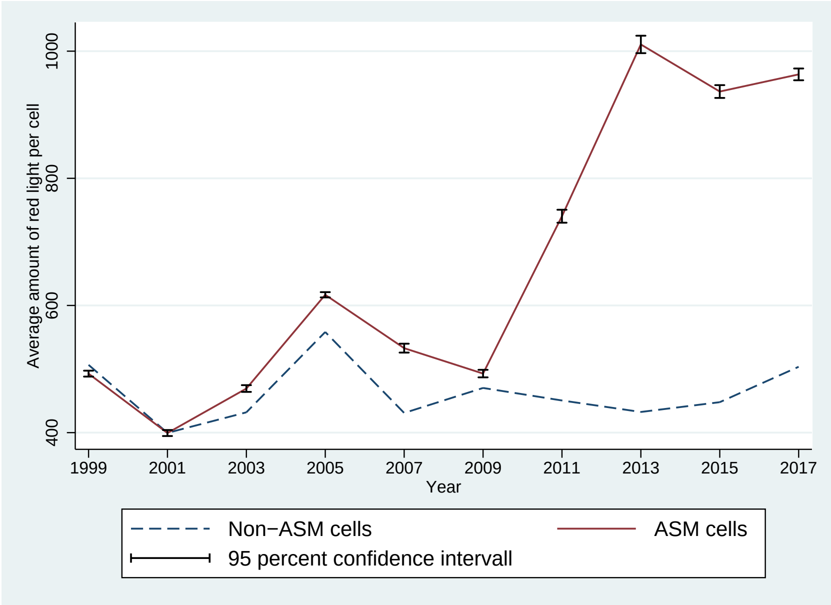


Figure 2.8: Amount of red light in ASM and non-ASM cells

Notes: Based on annual Landsat 7 minimum image bands for year pairs from 1999 to 2017.

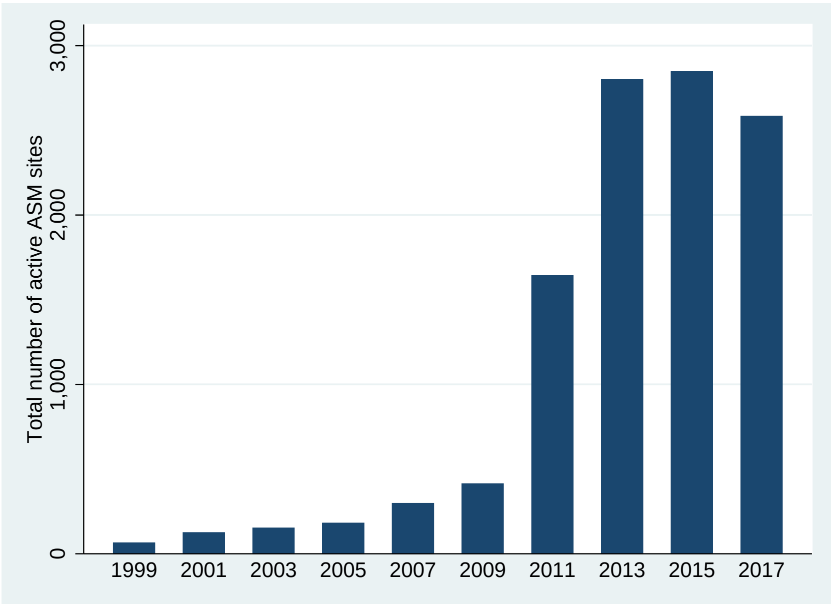


Figure 2.9: Active ASM sites over time

Notes: Based on annual Landsat 7 minimum image bands for 1999-2017.

The GLSS provides a rich set of questions at the individual, household and community level (one cluster can have multiple communities). Amongst these are detailed questions on types of income and expenditure, assets, employment, education and health. Price levels at the cluster level allow for a precise comparison of living standards between households in different regions. In addition, the community questionnaire of GLSS provides information on each community's major economic activities, infrastructure (roads, hospitals, electricity, piped water etc.), migration and average wages.

This thesis improves on previous studies in the use of GLSS wave 6 data, for which the public download has incomplete GPS coordinates for many clusters. Following up with representatives from the World Bank and Ghana's Statistical Service, I obtained the complete geo-locations for all survey clusters. This increases the amount of survey clusters from 839 to the complete 1,200. While this increases the accuracy of GLSS household locations compared to previous studies, there still remains some measurement error. This is because cluster centres instead of actual household locations are recorded in the data. However, given that clusters, i.e. villages or neighbourhoods, are small in size, the resulting measurement error in calculating the distance to ASM and LSM sites is most likely limited. Further, the ASM and LSM measures are affected in the same way by the potentially wrong household locations, which means that the comparison between the two forms of mining is not hindered.

Table 2.1 shows the number of household-level observations for each GLSS wave and region covered in the image detection introduced in the previous section. Note that the total number of observations in this table is still lower than outlined above for two main reasons. First, the area covered by the satellite image predictions only spans eight out of Ghana's ten regions partially (see figure 2.3). Second, in the main sample only observations with non-missing values for the main variables are included, giving a total sample size of 12,361 households in 892 different clusters across the three survey waves.[20] Figure 2.10 shows the combined data from GLSS and the machine learning predictions.

In addition to GLSS, I use Demographics and Health Survey (DHS) data for Ghana from 1993 to 2014 to explore health effects of artisanal mining. Similar to the GLSS surveys this data consists in repeated cross-sections that are geo-coded at the cluster level and nationally representative (GLSS is regionally representative). The unit of observation in this data is mothers when estimating child mortality and children under five years of age when analysing child diseases. Along with health outcomes, DHS collects socio-economic variables such as level of education and age, which are used as controls. Note that DHS geo-locations are, unlike GLSS, randomly displaced between 0-5 km, with 1 percent of the sample being displaced by up to 10 km. Similar to the GLSS analysis, and as is best practice, I will use distance buffers of five km and above to account for this displacement. Still, the resulting measurement error in the independent variable is going

[20] These households represent 49,549 individuals.

Table 2.1: Number of sample household observations per region and GLSS wave

	4	5	6	Total
Ashanti	963	1526	1819	4308
Brong Ahafo	457	613	1334	2404
Central	133	171	394	698
Eastern	702	753	1446	2901
Greater Accra	106	325	587	1018
Northern	0	15	47	62
Volta	25	43	38	106
Western	209	231	424	864
Total	2595	3677	6089	12361

Source: Ghana Statistical Service (1998), Ghana Statistical Service (2008), Ghana Statistical Service (2012).

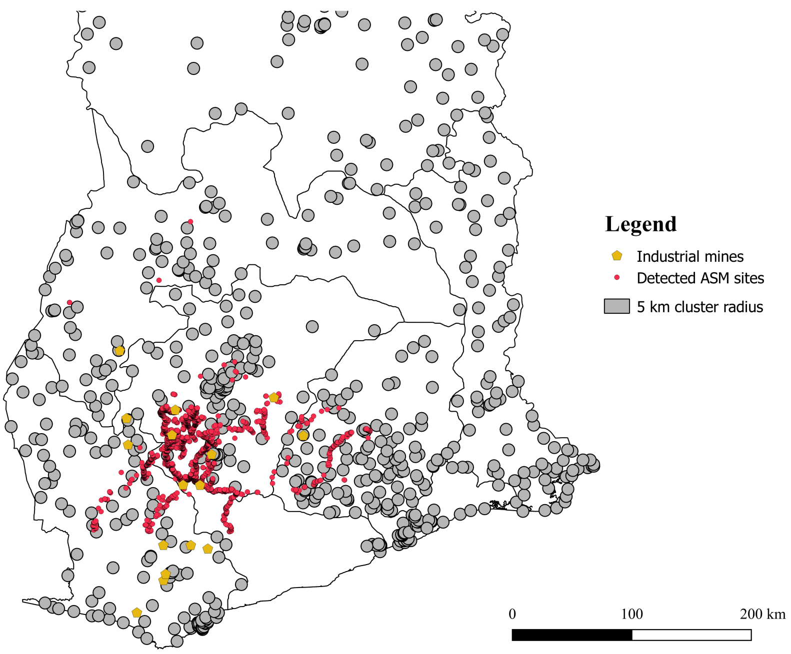


Figure 2.10: Artisanal mines, industrial mines and household locations

Source: author's visualisation of Ghana Statistical Service (2012), Microsoft.

to be larger than in the GLSS survey. As with GLSS, the attenuation bias introduced by this will affect the results for ASM and LSM in the same way.

2.3.2 Grid level data

The main focus of this article is based on household-level survey data described above. As a complementary strategy a grid cell approach is used in chapters 3 and 4. Figure 2.11 shows the basic set-up for this exercise. The main estimations are based on a grid sample that contains the entire area covered by the machine-learning detection of ASM sites described in section 2.2.2. This full grid sample contains 56 by 56 square cells for a total of 3,136 cells, where each cell has a side length of 0.05 decimal degrees or roughly 5.6 kilometres at the equator.[21] This cell size and location are based on the Normalised Difference Vegetation Index (NDVI) data, which has the lowest spatial resolution of all data sources used in this analysis. For all layers with a higher resolution the 0.05 degree cell average is taken. The different datasets used in this grid cell approach are described below.

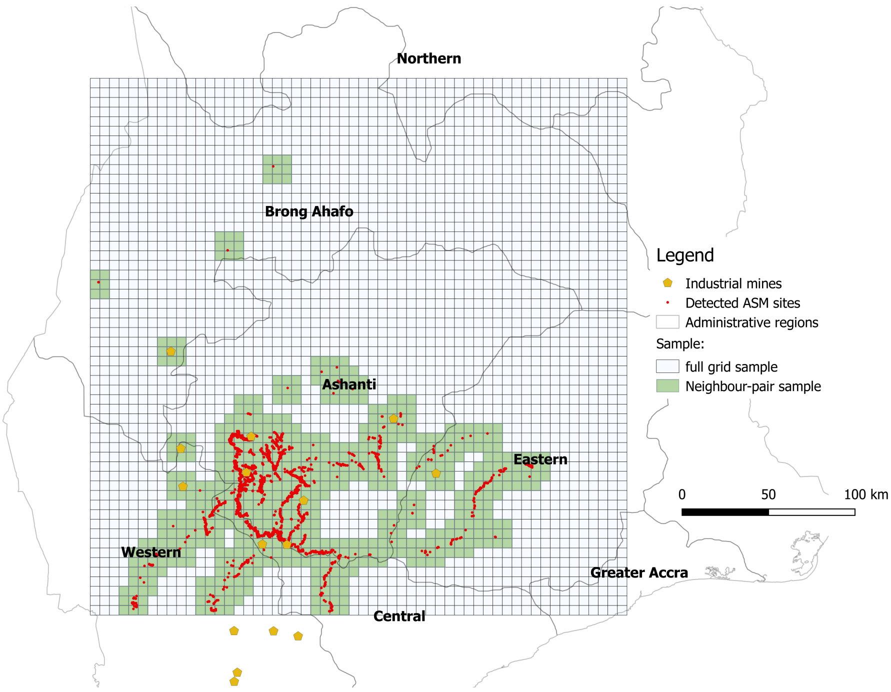


Figure 2.11: Sample used for grid cell analysis

I use the DMSP nighttime lights series from the National Oceanic and Atmospheric Administration (NOAA) to measure aggregate economic activity at the cell level over time (NOAA, 2019). The NOAA nighttime lights data contains cloud-free lights from

[21] In an alternative specification instead of the full grid sample a neighbour-pair sample is used (see figure 2.11). In this neighbour pair sample only cells that contain at least one ASM or LSM site and their first-degree neighbours are included. This strategy is similar to Acemoglu et al. (2012), Buonanno et al. (2015) and Berman et al. (2017).

cities, towns, and other sites with persistent lighting, including gas flares.[22] Data values are expressed in Digital Number (DN) and range from 0 (darkest) to 63 (brightest). They cover the years 1992 to 2013. Nighttime lights have been found to be a good proxy for economic development (Henderson et al. (2012), Stelios Michalopoulos and Papaioannou (2014) and Bruederle and Hodler (2017)). They will be used here to provide an overall estimate on the correlation of small-scale mining and local economic activity.

There are a number of caveats associated with using DMSP NOAA nightlights to measure economic activity. First, the satellites lack onboard image calibration and are subject to trajectory shifts and degradation. Atmospheric conditions also differ by year. which means the resulting DN of nightlights cannot directly be compared across time (Li and Zhou (2017) and Gruendler and Link (2019)). To reduce these distortions, I follow Li and Zhou (2017), who develop a step-wise calibration approach (see figure A.5). This calibration corrects for systematic over- or underestimation of nightlights by using years with overlapping satellite observations. They estimate a second order regression model based on these overlapping years to infer correction coefficients. I apply the coefficients provided by Li and Zhou (2017) to calculate the calibrated nighttime lights for each satellite for each year.[23]

The second major caveat of the DMSP nightlights is that they are top-coded at 63 DN and this maximum luminosity is reached frequently in both developed and developing countries (Bluhm and Krause (2018) and Gruendler and Link (2019)). This effectively makes urban areas appear dimmer than they actually are. I use the approach developed by Bluhm and Krause (2018) to correct for top coding. The intuition is to use the average luminosity of pixels below the "top-coding" threshold and the share of these pixels to correct the pixels above this threshold upwards.[24] Finally, to obtain the corrected average luminosity for each grid cell in each year, I take the average DN for all years with two satellite observations. The cell-level average nightlight luminosity in DN from 1992 to 2013 are shown in figure 2.12.

To measure environmental effects of ASM, I use data on forest cover loss (FCL) from Hansen et al. (2013) and Normalised Difference Vegetation index (NDVI) data from Goodman et al. (2017). The FCL data records areas where canopy cover above 5 metres height was lost between 2001 and 2016 at a resolution of 1 arc-second per pixel, or 30 by 30 metres at the equator. FCL has been used in similar applications to show the connec-

[22] Gas flaring is a common method to dispose excess gas in petroleum production and will upward bias nighttime lights. It is best practice to exclude areas with gas flaring, which are captured by NOAA, from the analysis (Elvidge et al., 2009). However in Ghana, all gas flaring takes place off the Atlantic coast and so falls outside the sample area.

[23] The formula for this is $DN_{ref} = a_0 + a_1 DN + a_2 DN^2$, where DN_{ref} is the brightness in DN of the year to be corrected, and DN is the DN of the year used for correction.

[24] The according formula for the mean luminosity L_i in grid cell i is $L_i = \omega_{iN}\lambda_{iN} + (1-\omega_{iN})\frac{\alpha}{(\alpha-1)}\phi_i$, where ω_{iN} is the share of pixels below the "top-coding" threshold, λ_{iN} is their average luminosity, α is a shape parameter and ϕ_i is the value of the top-coding threshold. Following Bluhm and Krause (2018) and Gruendler and Link (2019) I use α equal to 1.5 and ϕ_i equal to 55.

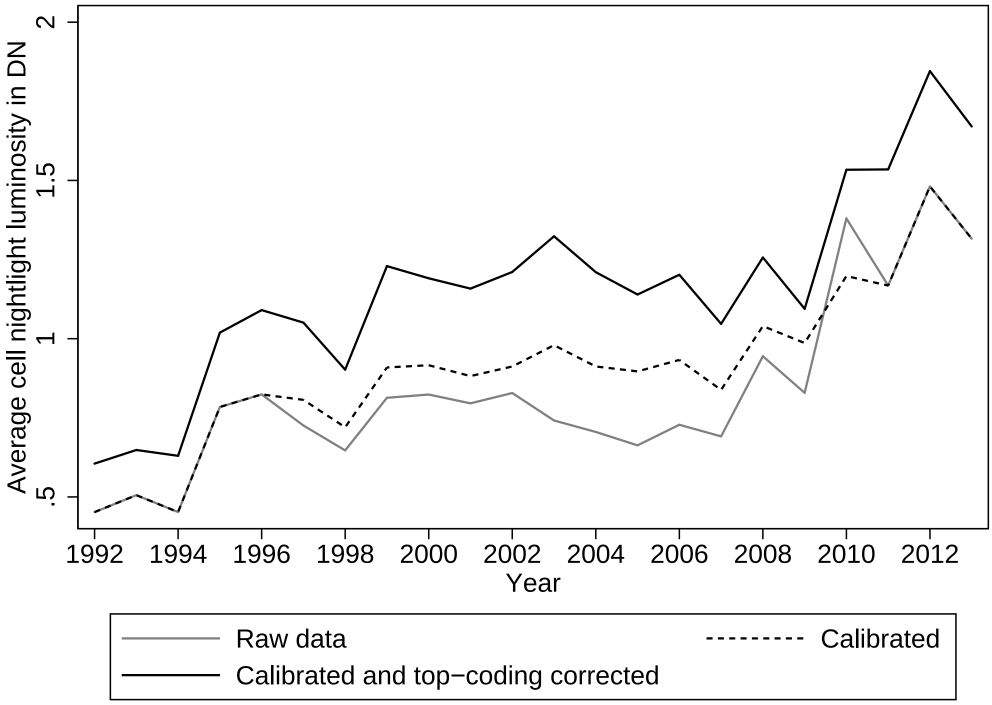


Figure 2.12: Average nightlight luminosity in sample area 1992 to 2013

Source: Author's calculation based on NOAA (2019). Image and Data processing by NOAA's National Geophysical Data Center. DMSP data collected by the US Air Force Weather Agency. Notes: Based on 3636 grid cells in sample area in Ghana.

tion between political incentives and deforestation (Burgess et al., 2012) and to identify small-scale mining in Colombia (Saavedra and Romero, 2017).

The NDVI data is provided by AidData and stems from the Land Long Term Data Record (LTDR) created by Pedelty et al. (2007). It has a spatial resolution of 0.05 degrees or roughly 5.6 kilometres at the equator and covers years from 1981 to 2014. NDVI measures uses the difference between near infrared and visible lights to infer the health of vegetation. The annual ndvi index ranges is aggregated up from daily and monthly observations and ranges from 0 (unhealthy) to 10,000 (healthy).

The malaria dataset is developed by Bhatt et al. (2015). They combine geo-referenced malaria tests from survey and study data with spatial data on e.g. climate, land cover, elevation and population to infer a cell-level annual panel (see figure A.6 in the appendix for the covariates used). This panel contains the plasmodium falciparum infection prevalence and clinical incidence rate among 2-10 year olds, which is the most common type of malaria in Sub-Sahara Africa. The resolution is 2.5 arc-minutes (5 x 5 kilometres at the equator). Detailed background on this dataset is provided in the online appendix to Bhatt et al. (2015).

2.3.3 Other sources

In addition to the machine learning based data introduced in section 2.2.2, other sources on small- and large-scale mining are used in this analysis. For small-scale mining, data on official licenses issued between 1992 and 2017 from Ghana's Minerals Commission Small Scale Mining database is applied.[25] For each registered ASM operation the database contains name, location (given by the nearest village/town plus district and post address of the company), issue and expiry date as well as size in acres. Note that the issue and expiry date is only provided for licenses that expired before April 25th 2017, the date the data was released to me.

Tables 2.2 and 2.3 show the distribution of ASM licenses over the regions in Ghana during the different GLSS survey years in number of licenses and their size. Most licenses are located in the Ashanti, Eastern, Western and Central regions, which is also where the bulk of detected ASM sites are located. Note that there are a number of caveats in comparing the two types of ASM measures. Official licenses can also be underground, detected mines can be inactive and it is not clear how many of the valid licenses are actually used for mining.

Licenses are granted for five years. Therefore, the licenses currently valid, issued between the years 2012 to 2017, cannot be dated precisely. To get some idea about the

[25]Official data on SSM licenses is available from 1992 onwards, following the legalisation of this industry. The legalisation of SSM was part of the Minerals and Mining Law 1986 (PNDCL 153) and further implemented in Mercury Law (PNDCL 217), Small-Scale Gold Mining Law (PNDCL 218), and Precious Minerals and Marketing Law (PNDCL 219) in 1989, see Hilson and Potter (2005).

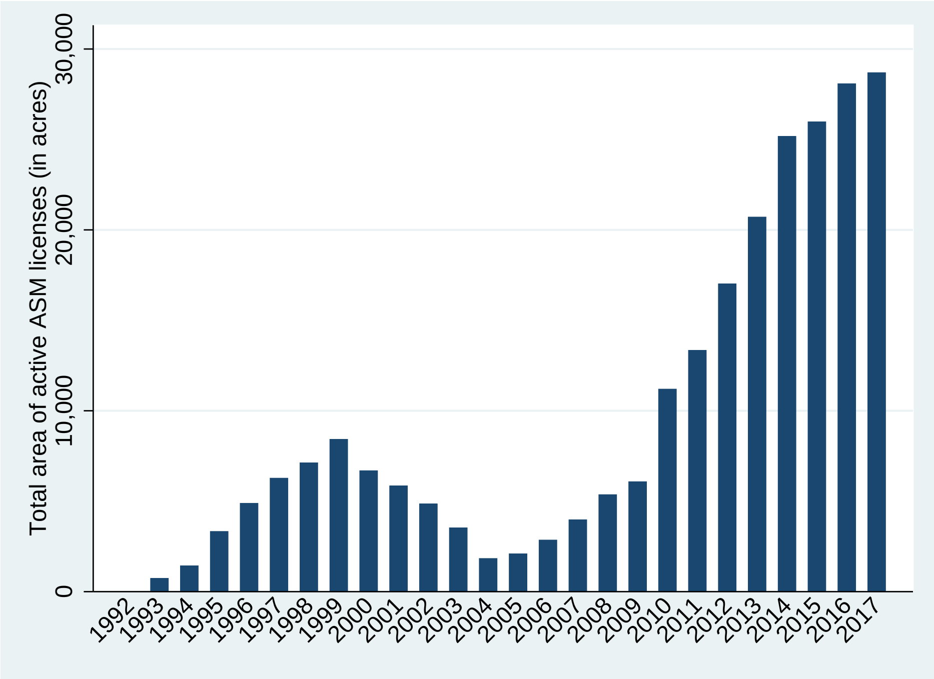


Figure 2.13: Area of active ASM licenses by year in acres

Source: Author's calculation based on data from the Ghanaian Minerals Commission Small Scale Mining Database, April 25th, 2017. This graph combines the data sets "Expired licenses" (incl. year issued, year expired) and "Valid licenses" (year issued, expired not available). For the valid licenses the year issued is assigned uniformly between 2012 and 2017.

Table 2.2: Number of licensed and detected ASM sites by region, only covered area

	ASM Licenses 98/99	ASM Licenses 05/06	ASM Licenses 12/13	ASM Licenses 12-17	Detected ASM sites 2014
Ashanti	62	17	198	486	2582
Brong Ahafo	28	0	9	61	4
Central	111	23	70	125	592
Eastern	64	42	136	284	195
Greater Accra	0	0	0	0	0
Western	64	6	69	264	266
All regions	329	88	482	1224	3639

Source: ASM licenses: Author's calculation based on data from the Ghanaian Minerals Commission Small Scale Mining Database, April 25th, 2017. This data consists of two sets "Expired licenses" (including year issued, year expired) and "Valid licenses" (year issued, year expired not available). Detected ASM sites: based on machine-learning detection of Bing satellite images.

Table 2.3: Size of licensed and detected ASM sites by region, only covered area (in acres)

	Size ASM Licenses 98/99	Size ASM Licenses 05/06	Size ASM Licenses 12/13	Size ASM Licenses 12-17	Size Detected ASM sites 2014
Ashanti	1429	420	4505	10297	14936
Brong Ahafo	699	0	213	1285	23
Central	2634	575	1579	2544	3424
Eastern	1230	903	3113	6009	1128
Greater Accra	0	0	0	0	0
Western	1370	135	1573	5838	1539
All regions	7362	2033	10982	26070	21050

Source: See table 2.2.
Notes: For license data the actual size of each license is used. The average is 23 acres per license. For the detected ASM sites I assume that the whole image is covered by the ASM site, which is equal to 5.8 acres.

magnitude of how many licenses were issued from 2012 onwards and how much area they covered, I assume that the year of issuance is distributed uniformly between 2012 and 2017. The resulting development of area covered by official ASM operations is shown in figure 2.13. This data can be geo-located at the district level and can thus be used to distinguish effects between legal (this data) and illegal (satellite image based data) ASM. As of April 2017 there are 1,500 active small-scale mining licenses in Ghana. During 2012/13 there were 978 licensed ASM companies.

To account for large-scale gold mining, combined data on the GPS location and annual production of these mines is obtained for years between 1992 and 2017 from Aragón and Rud (2016), GHEITI (2017) and Ghana Chamber of Mines.[26] The validity of the GPS location provided in these sources is double-checked by satellite-imaged based observation and information from the respective mining company websites.

2.4 Conclusion

Non-experimental studies in developing countries often have to deal with data scarcity and this thesis is certainly no exception. This chapter has illustrated how in these contexts machine learning techniques in combination with remote sensing data can overcome that challenge. The result is the first detailed map of artisanal and small-scale mining activity in the African context. The second contribution of the chapter is to gather outcome data from numerous sources, which will allow for the first comprehensive impact evaluation of ASM in the remainder of this thesis.

[26] http://ghanachamberofmines.org/wp-content/uploads/2016/11/Performance-of-the-Industry-2017.pdf

While some caveats of this novel approach, such as identifying the timing of ASM sites, have been mitigated in this chapter, some restrictions remain. Notably, the ASM detection is crude in the sense that it does not distinguish between the many different types of artisanal mining, which for example vary in their degree of mechanisation and use of manpower. Further, only mining pits are identified, where only part of the production chain takes place. Related operations, such as processing of materials and sales activities, also generate income and don't necessarily take place on the detected sites. Note that this caveat is a bigger restriction when analysing economic effects, but not so much when looking at health and environmental outcomes. The approach provided here is therefore an important first step in understanding impacts of artisanal mining, but further research is needed to obtain a more complete picture.

Chapter 3

Economic impacts of artisanal and industrial mining

3.1 Introduction

Mineral extraction is a booming industry in many Sub-Sahara African countries. Between 1989 and 2014 gold production in Ghana increased nearly by a factor of 10 to over four million ounces (118 tonnes) and it now accounts for 8 percent of total GDP (GHEITI, 2015). Whether this mineral wealth also contributes to raising incomes and reducing poverty has been studied both at the national (Sachs and Warner (2001) and Brunnschweiler (2008)) and the local level (Chuhan-Pole et al. (2017), Aragón and Rud (2016) and Cust and Poelhekke (2015)) with mixed results (see discussion in chapter 1). Importantly, all these studies have in common that they only consider part of the mining sector, namely large-scale industrial extraction. This limitation is important, because, along with the general increase in mining output, artisanal and small-scale mining (ASM) operations have become increasingly common in many developing countries. In Ghana, the ASM sector now accounts for over one-third of the total gold production and employs over one million people directly (Hilson and McQuilken, 2014). In settings with such a sizeable ASM concentration it is therefore crucial to account for both large- and small-scale production.

The main question in this chapter is therefore whether artisanal and small-scale mining improves economic livelihoods and how its effects compare to industrial mining. To do so, I make use of the novel ASM location data introduced in the previous chapter 2. The effect of ASM on local incomes and expenditures is a priori ambiguous. ASM Workers and their families likely benefit from the employment opportunity and higher wages. On top of that neighbouring industries, such as hospitality, transport and manufacturing may benefit from increased demand. On the other hand, ASM activity will attract immigration and can potentially lead to higher prices, which may both create downward pressures on real wages. Finally, environmental and health burdens linked to ASM (World Health Organization, 2016) can reduce individual productivity and create additional costs. The net effect is therefore an empirical question, which will be the main focus of this chapter.

To find out whether ASM or LSM in fact provide economic benefits, I make use of two data sources and respective identification strategies. To understand how household level outcomes are affected by mineral extraction I use survey data from the Ghana Living Standards Survey (GLSS). This has the advantage of accurate measures of economic and other household characteristics. A downside is that GLSS consists in a repeated cross-section, where different clusters or villages are surveyed in each wave, with waves being on average five years apart. The unit of observation is therefore not constant over time. The alternative approach of using nighttime lights overcomes this challenge by comparing constant grid cells every year from 1992 to 2013.

This chapter is structured as follows. Section 3.2 provides descriptive results on how ASM and non-ASM areas compare in socio-economic outcomes. Section 3.3 introduces the identification strategy and reports results on how ASM and LSM affect economic outcomes. This includes different robustness checks. Next, section 3.4 provides a complementary economic analysis using nighttime lights as the outcome variable. Section 3.5 discusses potential channels and 3.6 concludes.

3.2 Descriptive results: Cross section 2012/13

I start the analysis with a cross-sectional comparison of household outcomes in ASM and non-ASM locations. I define an ASM area simply by having at least one detected ASM site within five kilometres distance. Because the ASM locations are based on satellite images from 2014, I combine them with the nearest available household data, which is wave six of the Ghana Living Standard Survey (GLSS) from 2012/13. For the sample area, GLSS wave 6 consists of 6,117 households, of which roughly 800 are within five kilometres of an ASM site.

The geographic distribution of small-scale mining, large-scale mining and district-level per-adult income is shown in figure 3.1. Darker colours indicate higher incomes. A large share of ASM sites are situated in high-income districts, within the Ashanti and Western regions. This is also true for most large-scale industrial mines. However, many neighbouring districts also exhibit high incomes, so it is not clear whether any differences can be attributed to the presence of either type of mining.

How do other socio-economic characteristics differ by proximity to artisanal mining? Table 3.1 shows summary statistics for the main variables. The first two columns contain averages or shares for households that live more (1) or less than five kilometres (2) away from the nearest detected ASM site. As shown on the previous map, large-scale mining is also more prevalent next to ASM. Average income and expenditure however are almost

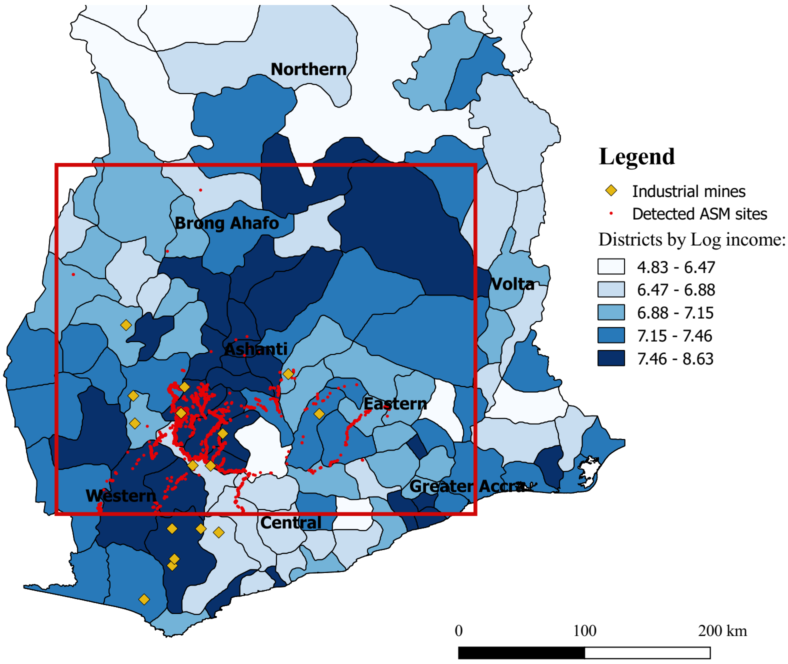


Figure 3.1: Small- and large-scale mining and income in Ghana

Source: Author's visualisation based on data from Ghana Statistical Service (2012), Microsoft, Aragón and Rud (2016), GHEITI (2015). Notes: The red box indicates the sample area covered by the machine learning detection. Mean income in natural logarithm of per-adult net income in 2010 GHC at the district level.

identical in both areas.

Eight percent of household in ASM areas have at least one member active in the extractive sector, which is four times more than in areas without ASM. But agriculture shares are also higher, which is consistent with the lower share of *urban* households. The sector shares can be interpreted in two ways. First, even in *ASM areas*, extractive resources play a negligible role compared to agriculture. Then it would not be surprising that average incomes and expenditures are not highly affected by it. Alternatively, knowing that most ASM sites operate illegally, the low extractive share could show that individuals are hesitant to disclose engagement in these activities.[1] In the second case, (reported) incomes might then also be *under-reported*, but I would still expect to find differences in expenditure, which is harder to link to potentially illegal sources. Because both income and expenditure are similar between ASM- and non-ASM areas, it seems plausible that the first explanation is true and that simply not many people are affected by mining.

Apart from that, gender composition and education levels seem comparable, as are immigration and prices. Perhaps surprisingly, more households have electricity in ASM areas than outside. Full summary statistics for the main sample are presented in figure B.1 in the appendix.

3.3 Main results

3.3.1 Identification

The descriptive results from the previous section suggest that while there are differences in socio-economic characteristics between ASM and non-ASM areas, economic outcomes for households are identical. However, to understand the effect of artisanal mining on these economic outcomes it is important to consider the development of mining and non-mining areas over time. It is for example possible that ASM activity has allowed initially poorer areas to catch up. ASM- and non-ASM areas may also differ in other characteristics, which are related to income and expenditure levels, such as geographic conditions (climate, soil fertility, connection to water bodies), infrastructure, human capital, institutional settings and so on.

To address these endogeneity concerns, I make use of two earlier GLSS surveys from 1998/99 and 2005/06. As there are no satellite images available for ASM detection for these years, I have to make an assumption about the respective ASM incidence. I assume that the majority of ASM sites detected in 2014 only become active mines after 2010. In this sense, observations from 1998/99 and 2005/06 indicate pre-treatment outcomes, and observations from 2012/13 represent post-treatment. 2010 is chosen as the cut-off point following the discussion in the previous chapter and illustrated in figure 2.9. By 2010,

[1] On top of that many people may engage in some mix of agricultural and mining activity.

Table 3.1: Summary statistics in 2012/13 by proximity to nearest ASM site

	(1) Non-ASM area	(2) ASM area	(3) Difference	s.e.
Number of ASM sites (5 km)	0.00	28.40	28.40***	(1.43)
Large-scale gold prod. (5 km) in tonnes	0.15	0.54	0.39***	(0.10)
Log net income p.a. in GHC (real)	7.34	7.21	-0.12*	(0.05)
Log expenditure p.a. in GHC (real)	6.51	6.47	-0.03	(0.04)
Share any HH member in agriculture	0.50	0.57	0.07***	(0.02)
Share any HH member in extractives	0.02	0.08	0.07***	(0.01)
Share HH urban	0.51	0.37	-0.14***	(0.02)
Share HH with electricity	0.67	0.73	0.06***	(0.02)
Share male	0.50	0.50	0.00	(0.01)
Average age	28.79	30.43	1.64**	(0.61)
Education of HH head on scale 0-4	2.36	2.40	0.04	(0.05)
Share HH head literate	0.57	0.56	-0.01	(0.02)
Share HH head moved here	0.31	0.32	0.02	(0.02)
Share HH head always here	0.69	0.68	-0.02	(0.02)
Cluster price index	1.00	0.99	-0.01***	(0.00)
Observations	5308	809	6117	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Based on GLSS data from 2012/13. Columns (2)-(1) might not equal (3) due to rounding error.

the gold price had also increased significantly, presumably making ASM operations more financially viable, sparking the increase in aggregate ASM production (see figure 2.1).

To implement the timing assumption above in a regression framework, I interact the ASM variable with separate survey year dummy variables for GLSS 4 (1998/99), GLSS 5 (2005/06) and GLSS 6 (2012/13). The remaining non-interacted coefficient for ASM then indicates the difference between ASM and non-ASM areas before this form of mining started to take place. The same is done with LSM to check whether its effect changes over time. With the interaction terms the specification takes the form of a differences-in-differences approach. The first difference is spatial variation in the number of ASM sites within different distance bands from the household. The second difference is time variation from the three survey waves GLSS 4, 5 and 6. The estimating equation then is:

$$
\begin{aligned}
Outcome_{it} = \beta_0 &+ \sum \beta_1^d ASM_i^d + \sum \beta_2^d (ASM_i^d * GLSS_t) \\
&+ \sum \beta_3^d LSM_{it}^d + \sum \beta_4^d (LSM_{it}^d * GLSS_t) \\
&+ \beta_5 \mathbf{X}_{it} + \gamma_k + \delta_t + \varepsilon_{it},
\end{aligned} \tag{3.1}
$$

where $Outcome_{it}$ is the logarithm of per-adult equivalent income or expenditure of household i in year t. ASM_i^d indicates the number of detected small-scale mines within different distance bands d from household i. Using the number of mines instead of the simple distance or incidence has the advantage that ASM operations of different sizes are identified separately. The distance bands are 0-5, 5-10, 10-20, 20-30, 30-40 and 40-50 kilometres. The omitted category is therefore mines that are further than 50 kilometres away from the household location. This distance band approach is similar to Chuhan-Pole et al. (2015) and Aragón and Rud (2016) and allows for non-linear effects of proximity to a mine. $ASM_i^d * GLSS_t$ is the interaction between the artisanal mining variable and the GLSS survey waves 4, 5 and 6. The interactions with wave 6, when most artisanal mines are assumed to be active, are the main treatment variables of interest. To distinguish the effect between artisanal and large-scale mining the same structure is used for the latter. LSM_{it}^d is the gold production in tonnes during survey years t in the respective distance bands d from household i. $LSM_{it}^d * GLSS_t$ is then the interaction with the different survey years.

$\mathbf{X}_{it}$ is a set of household (head) controls including sex, age, age squared, educational attainment, religion and migration status of the household head. $\mathbf{X}_{it}$ also contains whether the household location is urban or rural and the number of adult equivalents living in the household.[2] γ_k and δ_t are district and survey fixed effects. With district fixed effects, initial differences in geographic conditions, infrastructure and human capital are controlled

[2]This equivalence scale is based on calorie intakes of individuals of different sex and age. A 25 to 50 year old male has an equivalence scale of one. Younger, older and female household members have a lower scale.

for. While ideally this would be done at the cluster or even individual level, this is the lowest level the data permits. Time fixed effects account for changing macroeconomic conditions. The error term ε_{it} is clustered at the GLSS cluster level.

The validity of this difference-in-difference approach relies on parallel trends in ASM- and non-ASM areas in absence of the treatment, i.e. before GLSS 6. The parallel pre trends from GLSS 2 to 6 are shown in figures B.1 and B.2 in the appendix.

An important limitation of this analysis is that data on artisanal mines is only available for 2014, while the closest household data is from 2012/13. However there are two features of ASM operations that mitigate this caveat. First, artisanal alluvial and surface mines typically run for approximately ten years, which makes an overlap between 2012/13 and 2014 very likely (Mantey et al., 2016). Second, even if mines became inactive before 2014 they would still be captured by the machine learning detection, because the mining pits remain visible for many years after operation has ceased. Considering the development of aggregate ASM production and the gold price, many of these abandoned ASM sites will have been active in the period between 2010 and 2013 (GLSS wave 6). The main explanatory variable ASM_i and its interaction with the GLSS survey waves therefore allow to measure an effect of ASM activity even when the timing of the explanatory and outcome variables do not exactly overlap. Lastly, the number of ASM sites created after the date of the GLSS survey 2012/13 and before the date of the satellite images 2014 is likely to be small, as the time window for this is narrow. Even if the data included some false positives in the explanatory variable, this would only lead to attenuation bias of the effect size. A measured positive income effect of ASM can in this sense be interpreted as a lower bound.

3.3.2 Results

This section presents the main findings on the relation between artisanal mining, large-scale mining and economic outcomes. First, cross-sectional and repeated cross-sectional results for income and expenditure are shown in table 3.2. All columns include the full set of household (head) controls, district fixed effects and distance brackets of ASM and LSM up to 40-50 kilometres. Coefficients are only reported for the closest distance band 0-5 kilometres to show immediate effects of ASM and LSM. Columns (1) and (3) report cross-sectional correlations only from the last GLSS survey wave 6 from 2012/13. They are based on equation (3.1), excluding the interaction terms $ASM_i^d * GLSS_t$ and $LSM_{it}^d * GLSS_t$. Here neither ASM nor LSM significantly correlate with economic outcomes.

Columns (2) and (4) show results for the full model including the survey interaction terms and using all three GLSS surveys. The results for artisanal mining, as indicated by the interaction term $ASM_{5km} * GLSS6$, are zero throughout. Neither before nor after 2010 is there a significant difference in household income or expenditure. From this evidence, local households do not seem to benefit from artisanal mining activity.

Large-scale mining seems to at least partially benefit local inhabitants, with similar incomes in 1998/99 and gains of 10 percent per tonne (35 percent per s.d.) in 2005/06. Note that large-scale gold production volumes are captured annually, so that the un-interacted coefficient on LSM_{5km} is not a pre-treatment difference as in the case of ASM_{5km}. Overall the effect of LSM on incomes is therefore zero in 1998/99, 9.4 percent in 2005/06 (which is statistically significant) and zero in 2012/13. The income increase due to LSM also translates to expenditure gains of 5.7 percent during GLSS 5 (see column 4).

Table 3.2: Effects of mining on income and expenditure in (repeated) cross-section

	Income (1)	Income (2)	Expenditure (3)	Expenditure (4)
ASM_{5km}	0.002 (0.002)	–0.002 (0.004)	–0.001 (0.002)	–0.002 (0.002)
$ASM_{5km} * GLSS5$		0.004 (0.003)		0.000 (0.002)
$ASM_{5km} * GLSS6$		0.002 (0.004)		0.002 (0.002)
LSM_{5km}	0.022 (0.014)	–0.008 (0.007)	0.014* (0.008)	0.001 (0.004)
$LSM_{5km} * GLSS5$		0.102*** (0.019)		0.057*** (0.011)
$LSM_{5km} * GLSS6$		0.009 (0.015)		0.002 (0.008)
Mean (dep. var.)	7.412	6.818	6.600	6.186
SD (dep. var.)	1.400	1.486	1.224	1.295
Observations	6117	12389	6117	12389
R^2	0.205	0.289	0.703	0.743
Household controls	Yes	Yes	Yes	Yes
Distance brackets <= 50km	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Time FE	No	Yes	No	Yes

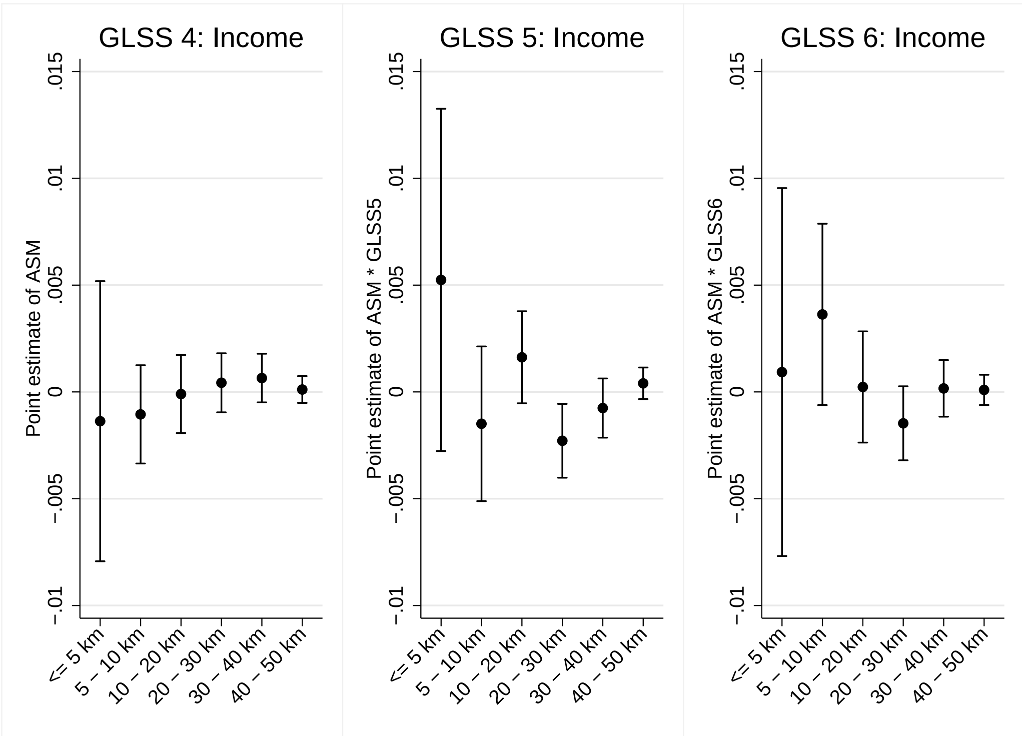
* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at GLSS cluster level in parentheses. All regressions include sample weights. Net income and expenditure in 2010 log per-adult equivalent household values. Additional distance brackets of the number of artisanal mines and large-scale gold production include 5-10, 10-20, 20-30, 30-40, 40-50km.

Table 3.2 shows that households nearby ASM operations do not benefit economically. But perhaps mining in the immediate neighbourhood is detrimental, while having it within reach, or commuting distance, actually helps? In other words, are there more far reaching effects of ASM or LSM? Figures 3.2 and 3.3 show the coefficients of ASM and LSM from equation 3.1 for all distance brackets for the three different survey waves.

In short, there are no far-reaching effects of either ASM or LSM. Figure 3.2 shows flat and non-significant coefficients for ASM throughout. The same is true for LSM in figure 3.3, except for the previously reported large and positive coefficient within five kilometres in GLSS 5.

Figure 3.2: Effect of ASM on income by distance bracket



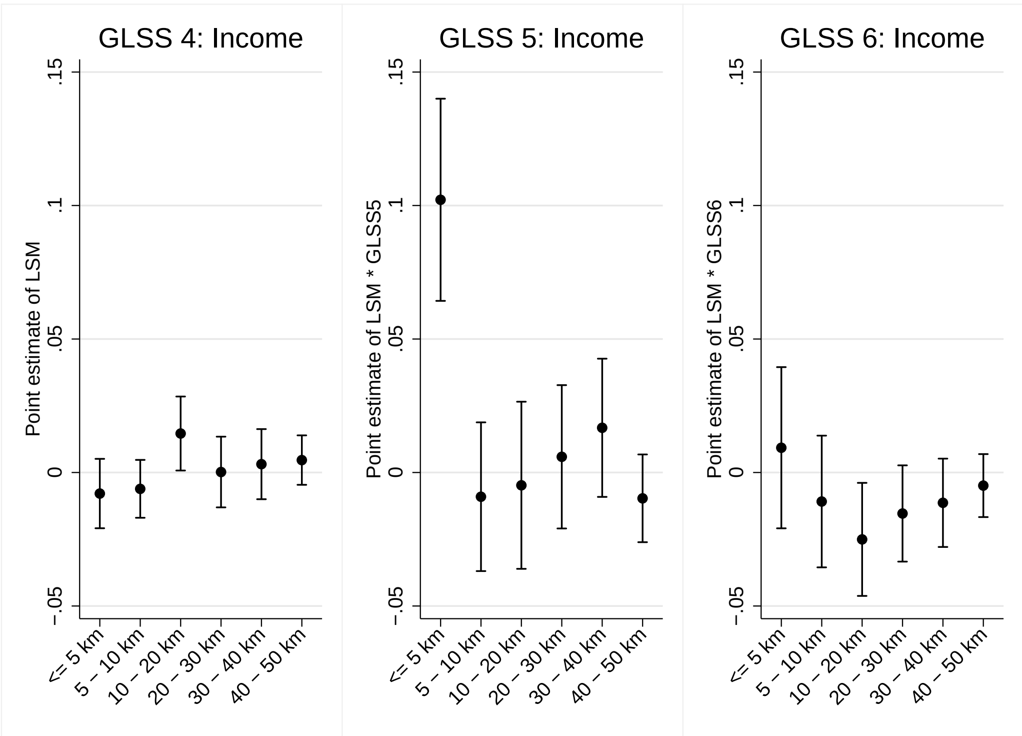
Notes: 95 percent confidence intervalls are used. Point estimates are the coefficients of (1) *ASM*, (2) $ASM * GLSS5$ and (3) $ASM * GLSS6$ on log net income for different distance bands from the specification described in equation (3.1).

3.3.3 Robustness checks

The evidence so far suggests artisanal mining fails to raise local living standards. In this section I test whether this result holds under different model specifications, for a binary instead of a continuous treatment of ASM and whether the zero effect is related to legal or illegal small-scale mining.

First, I test if the results on incomes are sensitive to different model choices. Table 3.3 reports coefficients for the main specification from equation 3.1 (column 1) and three additional specifications. In column (2) region-survey fixed effects are used instead of district fixed effects. The reasoning behind this is to control not just for initial differences,

Figure 3.3: Effect of LSM on income by distance bracket



95 percent confidence intervalls are used. Point estimates are the coefficients of (1) *LSM*, (2) $LSM * GLSS5$ and (3) $LSM * GLSS6$ on log net income for different distance bands from the specification described in equation (3.1).

but also for linear time trends that are different between areas.[3] The coefficients on ASM are slightly larger in magnitude, but still statistically insignificant.

Another concern is that the artisanal mining variable measures also an effect of living close to water, as many artisanal mines operate in or near rivers. In fact households that have any artisanal mines within five kilometres distance have an average distance of 2.6 kilometres to the next river, while households farther than five kilometres from ASM sites have an average distance of 4.4 kilometres. Therefore in column (3) distance to the nearest river in kilometres is used as an additional control variable.[4] The results from the main specification are not affected by controlling for distance to water.

Lastly in column (4) industry of employment of the household head and industry-survey fixed effects are included as controls. The goal of this is to account for both differences between sectors as well as income changes in these industries over time that are not due to artisanal mining. The zero effect of artisanal mining remains again unchanged, while the coefficients on large-scale mining are now somewhat smaller. While this gives confidence to the general findings this specification is not preferred as it does not allow for individuals choosing different occupations due to ASM, which is one potential channel of how ASM might affect economic outcomes.[5]

Next I want to test if the extensive instead of the intensive margin matters in terms of mining effects on incomes. It may be that having any artisanal mining nearby affects income differently than the number of sites. For this, I define the explanatory variable as a dummy, that is one if there are any detected ASM sites within a given distance band and zero otherwise. Similarly, I define an LSM dummy as one if there is any large-scale gold production within a distance band and zero otherwise. Figure 3.4 shows the resulting point estimates from using these dummy variables within equation 3.1. Again, the interaction with GLSS 6 is of interest, but yields no effects. The coefficients on ASM are flat for every distance band. As in the case of the continuous measure, LSM remains strongly positive during GLSS 5.

Figure 3.2 in the main results section suggests that the point estimate of ASM on incomes is zero. The confidence interval around it however is relatively large with lower and upper bounds of -0.75 and +0.75 percent per ASM site. This might suggest that there is sizeable variation in the effect of artisanal mining on incomes for households within five kilometres distance. Here I test whether part of this heterogeneity between different treated areas or households can be explained by legal versus illegal artisanal mining. Note that the measure of ASM used so far picks up both licensed and unlicensed

[3]Ideally this would be done by including district-survey fixed effects, but due to the limited number of observations many district-survey categories would be omitted in the estimation.

[4]The results remain unchanged if distance to water is interacted with the different GLSS survey waves.

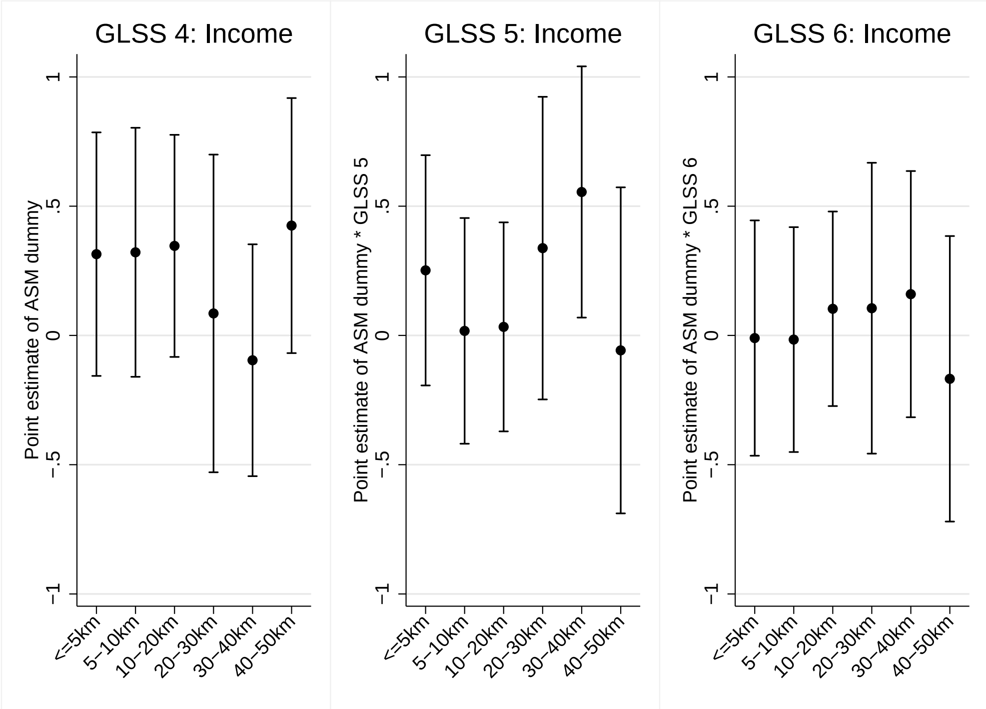
[5]As an additional robustness check, I use the inverse hyperbolic sine transformation of net income to deal with zero values. With this larger sample the results remain unchanged. Other robustness checks include alternative income definitions, using per-capita instead of per-adult equivalents and clustering standard errors at the district instead of cluster level. The results remain unchanged in these cases.

Table 3.3: Effects of ASM and LSM on income in different specifications

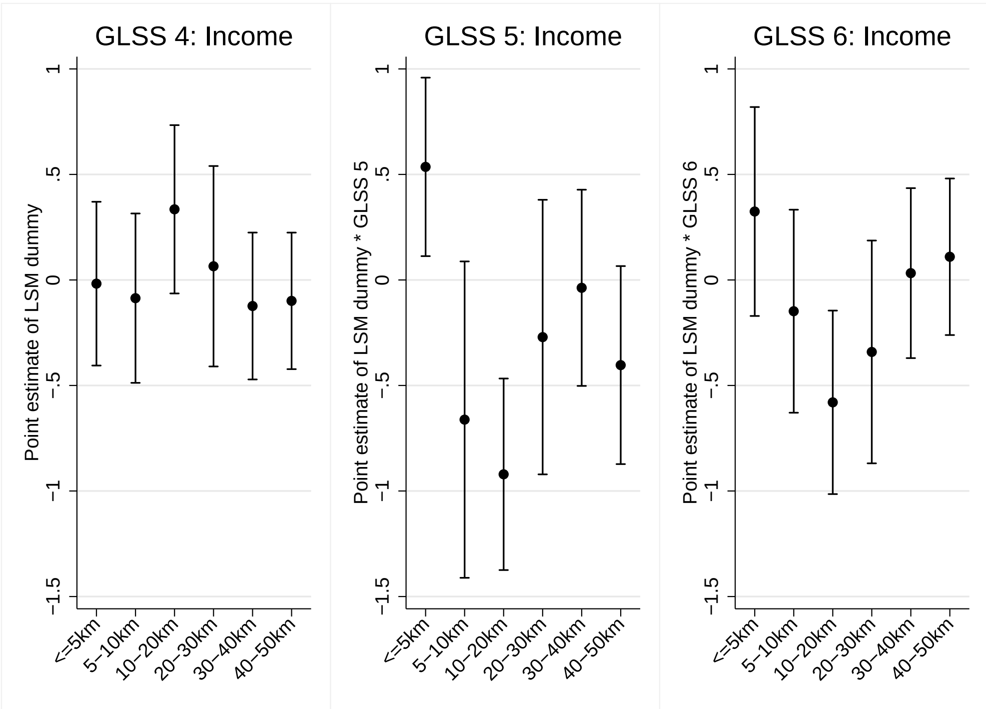
	Log net income			
	(1)	(2)	(3)	(4)
ASM_{5km}	–0.002	–0.006	–0.001	–0.003
	(0.004)	(0.004)	(0.004)	(0.003)
$ASM_{5km} * GLSS5$	0.004	0.008*	0.003	0.004
	(0.003)	(0.005)	(0.004)	(0.004)
$ASM_{5km} * GLSS6$	0.002	0.007	0.001	0.003
	(0.004)	(0.005)	(0.004)	(0.004)
LSM_{5km}	–0.008	–0.025***	–0.009	–0.004
	(0.007)	(0.007)	(0.007)	(0.007)
$LSM_{5km} * GLSS5$	0.102***	0.101***	0.099***	0.065***
	(0.019)	(0.019)	(0.019)	(0.020)
$LSM_{5km} * GLSS6$	0.009	0.029**	0.008	0.003
	(0.015)	(0.012)	(0.015)	(0.018)
Mean (dep. var.)	6.818	6.818	6.818	6.862
SD (dep. var.)	1.486	1.486	1.486	1.483
Observations	12389	12389	12389	11257
R^2	0.289	0.273	0.289	0.333
Household controls	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes
Distance brackets <= 50km	Yes	Yes	Yes	Yes
District FE	Yes	No	Yes	Yes
Region-Survey FE	No	Yes	No	No
Distance to water	No	No	Yes	Yes
Industry-Survey FE	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at GLSS cluster level in parentheses. All regressions include sample weights. Net income in 2010 log per-adult equivalent household values. Additional distance brackets of the number of artisanal mines and large-scale gold production include 5-10, 10-20, 20-30, 30-40, 40-50km.



(a) ASM dummy



(b) LSM dummy

Figure 3.4: Effect of ASM and LSM dummy on income by distance bracket

95 percent confidence intervalls are used. Point estimates are the coefficients of (1) ASM, (2) $ASM * GLSS5$ and (3) $ASM * GLSS6$ on log net income for different distance bands from the specification described in equation (3.1) in panel a and (1) LSM, (2) $LSM * GLSS5$ and (3) $LSM * GLSS6$ in panel b.

small-scale operations. Therefore, I now use official ASM license data from Ghana's Minerals Commission Small Scale Mining database, which includes years 1992-2017. Because the exact GPS location for these licensed operations is not available, I match them with the household data by district. This is not ideal, since districts in Ghana have different sizes, meaning that some households can be far away from licensed ASM sites within a district. Further, the analysis so far has shown that the effects of artisanal mining are not very far reaching and will therefore likely not show up at the district level. These limitations in mind, I re-estimate the previous specification with additional coefficients for legal ASM in table 3.4. Overall there are no income-increasing effects of legal or detected ASM sites.[6]

The main outcome variables income and expenditure aim to capture the possible benefit of ASM on overall economic well-being. In this check I re-run the estimation using the poverty headcount variable based on Ghana's national poverty line of 90 GHC (70 GHC for extreme poverty) in 1999 prices for GLSS 4 and 5 and 1314 GHC (792.05 GHC for extreme poverty) in 2013 prices for GLSS 6. The values are based on per adult equivalent per year consumption. The corresponding household headcount rates for absolute (extreme) poverty are 25.9 (14.4) percent for GLSS 4, 13.5 (6.3) percent for GLSS 5 and 14.5 (3.8) percent for GLSS 6.[7]

Table 3.5 shows results from linear probability models using the absolute (extreme) poverty headcount as an outcome variable and the same controls as in the main specification. The results here are overall consistent with the main ones, with the ASM coefficients being zero overall.

3.4 Grid-level analysis of nighttime lights

The analysis so far has been guided by the availability of survey data to measure economic outcomes. The advantage of using survey data is the rich set of control and outcome variables available. Household income for example is largely determined by household or household member characteristics such as education, age, industry of employment, household size etc. External factors, such as proximity to ASM, are then likely to have different effects according to these household and individual characteristics. The big downside of survey data however is the limited spatial and temporal coverage due to the high costs involved.

Nighttime lights on the other hand do not depend on the availability of expensive survey data, but are instead based on satellite images. This means they are consistently

[6]The same results hold if instead of the number of detected mines and licenses the size of both is used. The results are also unchanged if I keep using the previous measure of detected ASM sites within different distance bands from the household location and only to control for ASM licenses at the district level.

[7]Converted to individual headcount rates these figures are consistent with estimates from the Ghana Statistical Service (GSS) and a recent UNICEF report (Cooke et al., 2016).

Table 3.4: Effect of detected and licensed ASM on income at district level

	Income (1)	Income (2)	Income (3)
ASM_d	–0.0011	–0.0011	
	(0.0010)	(0.0011)	
$ASM_d * GLSS5$	–0.0006**	–0.0006*	
	(0.0003)	(0.0003)	
$ASM_d * GLSS6$	0.0001	0.0006	
	(0.0003)	(0.0004)	
$LegalASM_{d,t}$		0.0065	0.0086*
		(0.0049)	(0.0049)
$LegalASM_{d,t} * GLSS5$		0.0220*	0.0251**
		(0.0128)	(0.0126)
$LegalASM_{d,t} * GLSS6$		–0.0139**	–0.0066
		(0.0058)	(0.0052)
$LSM_{d,t}$	–0.0051	–0.0036	–0.0052
	(0.0057)	(0.0057)	(0.0056)
$LSM_{d,t} * GLSS5$	0.0028	0.0056	0.0046
	(0.0133)	(0.0134)	(0.0133)
$LSM_{d,t} * GLSS6$	–0.0204***	–0.0207***	–0.0218***
	(0.0074)	(0.0074)	(0.0074)
Observations	12389	12389	12389
R^2	0.279	0.280	0.278

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at GLSS cluster level in parentheses. Regressions include sample weights, district and time fixed effects and individual controls. Net income in log per-adult equivalent household values. $LegalASM_{d,t}$ are active ASM licenses during each GLSS wave.

Table 3.5: Effects of ASM and LSM on poverty

	Poor	Poor	Poor (extreme)	Poor (extreme)
	(1)	(2)	(3)	(4)
ASM_{5km}	–0.002*	0.002	–0.001***	0.001
	(0.001)	(0.002)	(0.000)	(0.001)
$ASM_{5km} * GLSS5$		–0.001		–0.000
		(0.002)		(0.001)
$ASM_{5km} * GLSS6$		–0.003*		–0.001
		(0.002)		(0.001)
LSM_{5km}	–0.002	–0.003	0.004***	–0.000
	(0.003)	(0.002)	(0.001)	(0.001)
$LSM_{5km} * GLSS5$		–0.019***		–0.012***
		(0.006)		(0.004)
$LSM_{5km} * GLSS6$		0.004		0.003
		(0.003)		(0.002)
Mean (dep. var.)	0.129	0.151	0.033	0.062
SD (dep. var.)	0.336	0.358	0.179	0.241
Observations	6117	12389	6117	12389
R^2	0.201	0.207	0.115	0.132
Household controls	Yes	Yes	Yes	Yes
Distance brackets <= 50km	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Time FE	No	Yes	No	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at GLSS cluster level in parentheses. All regressions include sample weights. Absolute and extreme poverty headcount measures based on national poverty lines. Additional distance brackets of the number of artisanal mines and large-scale gold production include 5-10, 10-20, 20-30, 30-40, 40-50km.

available across time and space. As noted in chapter 2, nightlights are commonly used as a proxy for local economic development. In this section I will use a grid-cell panel approach to compare areas with and without artisanal and industrial mining over time to infer effects on local economic activity.

Figure 3.5 shows growth in cell-level nighttime lights from 1992 to 2013. Growth is weak overall, as most cells remain at zero DN from 1992 to 2013. Strong increases only appear in urban centres such as Accra and Kumasi. At first glance, cells with ASM and LSM seem to have increased in luminosity over this period, although this is more apparent for LSM.

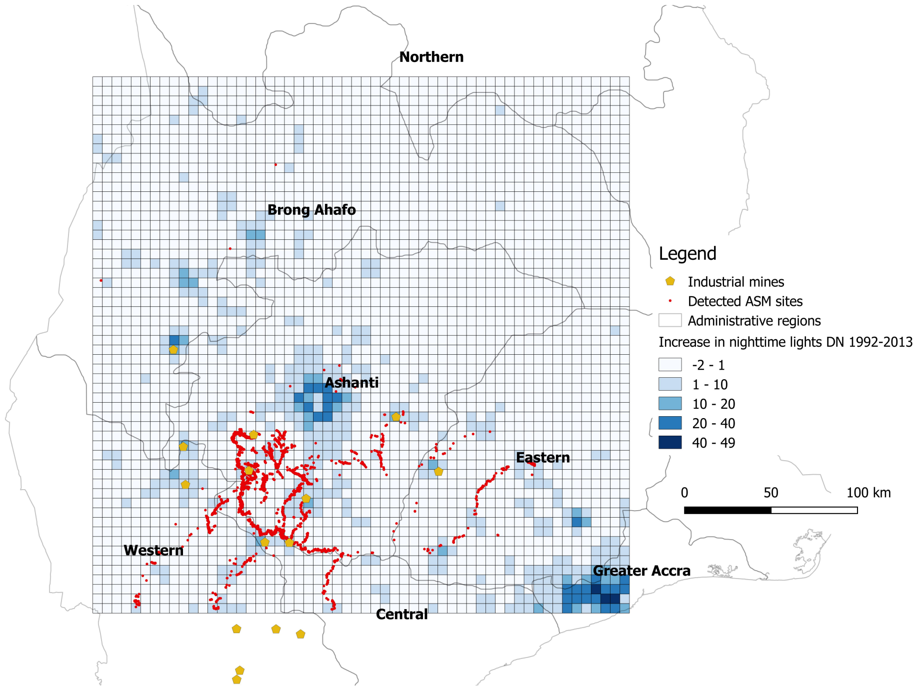


Figure 3.5: Change in nighttime lights from 1992 to 2013

Source: author's calculation based on NOAA data.

3.4.1 Identification

The main identification in this section comes from changes in outcome variables in cells with and without artisanal and industrial mining over time. The baseline estimating equation takes the form:

$$\begin{aligned} Nightlights_{it} = \beta_0 + \beta_1 ASM_i + \beta_2 (ASM_i * year_t) \\ + \beta_3 LSM_{it} + \beta_4 (LSM_{it} * year_t) \\ + \gamma_i + \delta_t + \varepsilon_{it}, \end{aligned} \tag{3.2}$$

where $Nightlights_{it}$ is the level of nighttime lights expressed in digital number units (DN), ranging from 0-165.[8] ASM_i is a dummy that is equal to one if there is at least one detected artisanal mine in cell i according to the satellite images from 2014. As before this dummy is time-invariant. LSM_{it} is a time-varying dummy that takes one if there is an active large-scale mine in cell i in year t. $ASM_i * year_t$ and $LSM_{it} * year_t$ are year interactions of ASM and LSM. As more of the detected artisanal mines become active over time, due to increases in gold price and improvements in production techniques (see figures 2.1 and 2.8), the years after 2010 in the interaction $ASM_i * year_t$ will show whether ASM affected economic development.

Importantly, γ_i are cell-fixed effects to control for constant unobserved differences between ASM and non-ASM cells, such as soil fertility, altitude, proximity to water and initial human capital and education levels. The regressor ASM_i is then omitted because of perfect collinearity with the cell fixed effects. δ_t are year-fixed effects and are used in all specifications. They account for overall changes in unobserved characteristics. In the context of satellite-based measures such as nighttime lights this is especially important, since they fluctuate strongly between years. The error term ε_{it} is clustered at the cell level.

A simplified way of showing this relationship is comparing mining and non-mining cells before and after 2010. The estimating equation is then:

$$\begin{aligned} Nightlights_{it} = &\beta_0 + \beta_1 ASM_i + \beta_2 (ASM_i * Post2010_t) \\ &+ \beta_3 LSM_{it} + \beta_4 (LSM_{it} * Post2010_t) \\ &+ \gamma_i + \delta_t + \varepsilon_{it}, \end{aligned} \tag{3.3}$$

where $ASM_i * Post2010_t$ and $LSM_{it} * Post2010_t$ are now interactions of ASM and LSM with an indicator that is one for years after 2010 and zero otherwise.

Both versions of the baseline model indicate whether cells with artisanal or industrial mining become brighter relative to all sample cells without mining, while accounting for cell- and time fixed effects. However, it may be that the full set of non-mining cells is not a suitable control group. For example, cells further away are more likely have different time trends in absence of the mining *treatment*.

Therefore, two additional, more demanding specifications are used. First, only ASM and LSM cells and their first-degree neighbours are included, which cuts the sample from 68,967 observations to 12,882 for the nighttime lights estimation (see "neighbour-pair sample" in figure 2.11).

A second variant of this is the same neighbour sample as above, but non-ASM neighbours can appear more than once to form neighbour pairs with ASM cells. The cell-fixed effects γ_i in this case are replaced by neighbour-pair fixed effects. The identification then

[8] This is the range that results from correcting nightlights for systematic under- and overestimation and top-coding as described in chapter 2. The original range as provided by NOAA (2019) is from 0-63 DN.

only comes from the relative changes in the outcome between ASM cells and their non-ASM neighbours over time. This strategy is based on Acemoglu et al. (2012), Buonanno et al. (2015) and Berman et al. (2017). This estimation also makes use of spatial heteroskedasticity and autocorrelation consistent (HAC) standard errors, developed by Conley (1999) and Hsiang et al. (2011). The HAC standard errors allow for cross-sectional spatial autocorrelation as well as location-specific serial correlation. Note that HAC standard errors in this case are only marginally larger than the cell-clustered standard errors from the previous estimations.[9]

3.4.2 Results

This section presents results of the grid-level analysis on nighttime lights. The estimations are based on equations 3.2 (figure 3.6) and 3.3 (tables 3.6 and 3.7).

Figure 3.6 illustrates the main result of the nightlight estimation: Artisanal mining does not make cells brighter, but industrial mining does. Panels 1 and 2 depict the point estimates of the interaction between the year indicators δ_t and indicators ASM_i and LSM_{it} respectively. If artisanal mining had an impact on local development, then we would expect increasing point estimates after 2010. However, if anything these cells become dimmer over time. Large-scale mining in panel 2 also does not show a clear trend, but the interaction coefficients are significantly above zero in most years. Panels 3 and 4 confirm this result by using continuous ASM and LSM measures.

Tables 3.6 and 3.7 support the nightlight result further. First, table 3.6 shows results from equation 3.3 for the full (columns (1) and (2)) and neighbour sample ((3) and (4)). The coefficient on $ASM_i * Post2010_t$ is slightly negative and statistically insignificant in all four specifications. LSM_{it} on the other hand is positive throughout, which means that areas with active industrial mines are significantly brighter at night. The size of the LSM coefficients indicates increases in luminosity between 1.9 and 7.4 DN.

Perhaps it is not surprising that areas with industrial mining emit more light at night. It is very likely that at least part of this light comes from the operating mines themselves. In this case it would not be clear if there is economic development on top of the mining activity. Column (2) therefore includes the mining status of neighbouring cells to see if these cells also benefit in the form of more night light. Again, while ASM neighbours are not brighter, LSM ones are. This suggests that industrial mining leads to more lighting even outside the production area.

Columns (3) and (4) use the reduced sample of only mining cells and their first degree neighbours. Comparing ASM cells and non-ASM neighbours shows very similar results to the full sample in column (1) with only LSM positively contributing to bright-

[9]Similar to Berman et al. (2017) I use a spatial kernel of 500 kilometres and a basically infinite time horizon (1000 years) for the temporal lag.

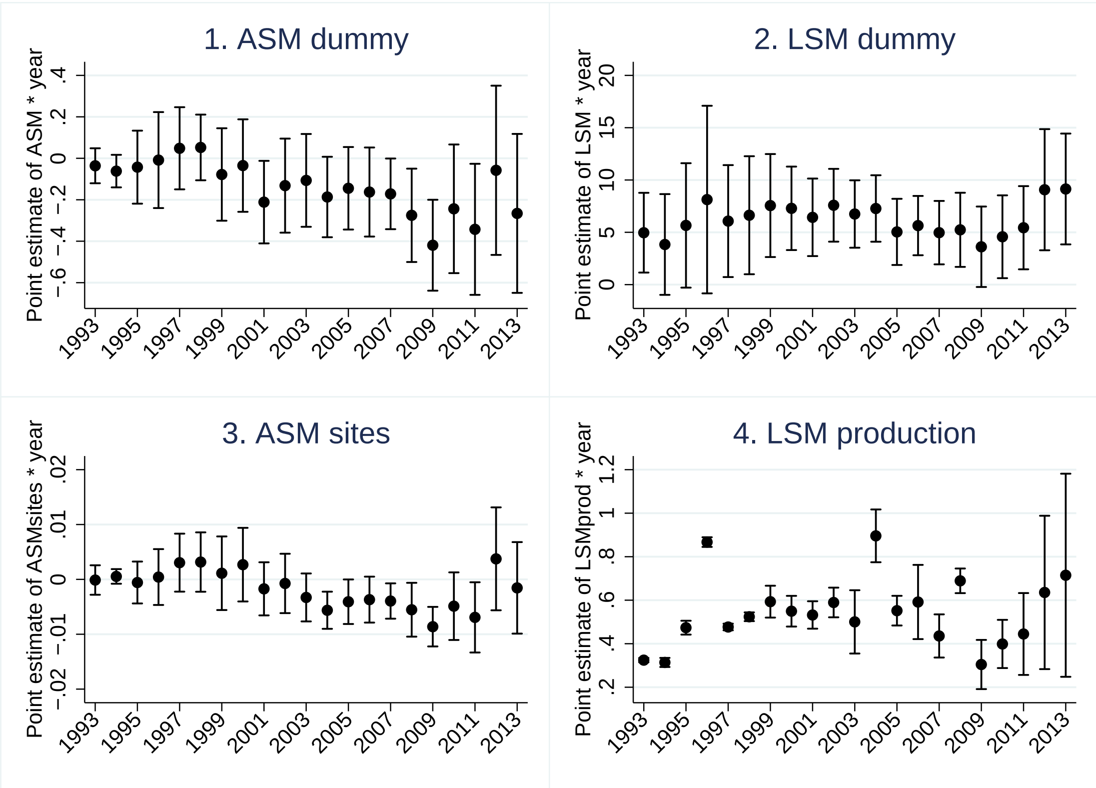


Figure 3.6: ASM, LSM and nightlight intensity

Notes: Based on equation 3.2. Depicted are point estimates of interaction terms of ASM or LSM and year dummies.

Table 3.6: Effects of ASM and LSM on nighttime luminosity

	Nighttime luminosity (in DN)			
	(1)	(2)	(3)	(4)
$ASM * Post2010$	−0.106	−0.185	−0.204	−0.200
	(0.121)	(0.177)	(0.183)	(0.200)
LSM	2.473***	2.479***	1.921***	7.441***
	(0.738)	(0.738)	(0.730)	(1.390)
$LSM * Post2010$	2.451	2.442	2.708	6.384***
	(2.176)	(2.166)	(2.582)	(1.704)
$ASM_n * Post2010$		0.095		
		(0.197)		
LSM_n		1.023***		
		(0.291)		
$LSM_n * Post2010$		1.341***		
		(0.438)		
Mean (dep. var.)	1.163	1.163	1.117	0.947
SD (dep. var.)	8.447	8.447	3.788	2.859
Observations	68967	68967	12882	42768
R^2	0.020	0.021	0.084	0.084
Controls	No	No	No	No
Full sample	Yes	Yes	No	No
Neighbour sample	No	No	Yes	Yes
Cell FE	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes
Neighbour-pair FE	No	No	No	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Based on equation 3.3. Clustered standard errors at cell level in parentheses for columns (1) to (3). Columns (1) and (2) use the full grid sample. Column (3) includes only ASM and LSM cells and their first-degree neighbours. Column (4) is the same sample as in (3), but non-ASM neighbours can appear more than once to form neighbour pairs with more than one ASM cell. Column (4) uses Conley standard errors.

Table 3.7: Effects of active ASM and LSM on nighttime luminosity

	Nighttime luminosity (in DN)		
	(1)	(2)	(3)
Active *ASM*	–0.227***	–0.214**	–0.102
	(0.087)	(0.092)	(0.119)
LSM	2.850***	2.021**	8.044***
	(0.865)	(0.922)	(1.402)
Mean (dep. var.)	1.163	1.117	0.947
SD (dep. var.)	8.447	3.788	2.859
Observations	68967	12882	42768
R^2	0.020	0.081	0.079
Controls	No	No	No
Full sample	Yes	No	No
Neighbour sample	No	Yes	Yes
Cell FE	Yes	Yes	No
Year FE	Yes	Yes	Yes
Neighbour-pair FE	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at cell level in parentheses for columns (1) and (2). Column (1) uses the full grid sample. Column (2) includes only ASM and LSM cells and their first-degree neighbours. Column (3) is the same sample as in (2), but non-ASM neighbours can appear more than once to form neighbour pairs with more than one ASM cell. Column (3) uses Conley standard errors.

ness. When controlling for spatial autocorrelation and using neighbour-pair fixed effects in column (4) the coefficient on ASM remains slightly negative, but is still insignificant. LSM is very strongly associated with nighttime lights with a 7.4 DN increase overall and an additional 6.3 DN after 2010. Note that the results are similar when applying HAC standard errors to the full sample or when using a continuous measure of ASM (number of sites) and LSM (production in tonnes), shown by table B.3 in the appendix.[10]

Finally, table 3.7 uses the time-varying measure of artisanal mining, identified by the combination of the machine-learning detection and an increase in red light (see chapter 2). The results here are qualitatively identical to the previous ones. If anything, ASM reduces nighttime luminosity, whereas LSM is associated with nightlight increases.

Overall, large-scale mining cells are significantly brighter than non-LSM cells. Since 2010 they have become even brighter and so have neighbouring non-mining cells. ASM cells on the other hand show no such effects. If nighttime lights are interpreted as a proxy for local economic development this clearly shows that the gains from large-scale mining are stronger than from artisanal mining. One caveat discussed is that large-scale mining cells might simply be brighter because the operation of these mines emits more light at night than small-scale mining. However, the fact that neighbouring cells of LSM cells are also brighter suggests that this is not the only reason.

3.5 Channels and discussion

Clearly, the presented evidence points to zero effects of ASM on economic household outcomes. This holds true in different model specifications, explanatory variable definitions and choices of outcomes. Different distances also don't seem to matter. This is a surprising result in the light of, as was outlined in the beginning, the importance of ASM as an employer, numerous qualitative studies and personal field research in Ghana.

But there are various reasons that may explain why artisanal and small-scale mining don't exhibit positive economic benefits to the local economy. The first set of reasons has to do with measurement problems. While the Ghana living standards survey is representative at the country and regional level it may not draw an accurate picture at the cluster, i.e. village or neighbourhood, level with too few observations available. On the other hand the measure of artisanal mining applied here is likely only a crude proxy of ASM activity. The identified mines can be either active or inactive during the time of the household survey. Even when a big share of ASM sites is likely to be active during 2012/13, it is not clear for how long this has been the case. Further, little can be said about the intensive margin of these operations. An attempt is made to account for this by counting the number of detected mines close to household locations instead of simply taking the distance to the closest artisanal mining site. Still, how much gold is produced and how many workers

[10]The results are also not sensitive to the choice of spatial or temporal lag.

are employed cannot be answered with this type of data. Finally, the linear distance between households and mining sites may be misleading. Mountain terrain, forests or rivers between them can possibly block the access and thus prevent economic linkages. Villages and ASM sites that appear close through my measure of artisanal mining may in reality be far from each other in travelling time. This is especially relevant, because the illegality of most ASM sites suggests that they operate in remote or somewhat hidden areas to face less risk of prosecution.[11]

There are also a number of economic reasons to explain the limited effectiveness of ASM to raise living standards. Viewed from a local demand shock perspective as outlined in Aragón and Rud (2013) and Chuhan-Pole et al. (2017), the increase in artisanal mining activity should increase the demand and thus the price for local inputs. This means that both labour prices, i.e. nominal wages, as well as prices for goods and services are increasing. Depending on whether workers are mobile or not, real wages may then either remain relatively stable or increase. However, even if the impact on wages was stronger than on prices and if immigration of workers was low, the increase in real wages may still be offset by pollution that is associated with health costs, and thereby decreased human capital and labour supply, and reduced agricultural productivity.

Several of these predictions can be tested. As already known, the overall effect on real incomes, encompassing income from agriculture, non-agricultural enterprises and wages, is zero (see table 3.2). But is there any evidence for an increased labour demand from a higher number of hours worked? This is tested in table 3.8. It shows that hours worked are higher with more ASM sites nearby (column 1), especially for young (2) and young male (3), but less for female workers (4). This can be interpreted as an increase in local labour demand due to ASM.

If labour demand in fact goes up, does the zero effect on real incomes then come from rising prices and or immigration of workers? To see if prices are affected, I look at the local price index, computed by the GSS. Since this variable is generated at the cluster level, I use this as the unit of observation. I then regress the cluster-level price index on the same explanatory variables as before and on region-survey wave fixed effects. Results are shown in table 3.9. By this measure, prices are not higher in ASM than in non-ASM areas. Rising prices therefore can not explain the zero income result.

While prices do not seem to drive down the real income effects, immigration might. Figure 3.7 shows responses from community officials surveyed during GLSS 2 to 6. They were asked whether more people had moved in or out of their community over the last ten years. In non-ASM areas, the share of communities with more arrivals than departures remained stable at around 52 percent from 1988 to 2013. Communities with artisanal

[11] One way to test this could be to use roads data to compute travelling distances between two locations. However, the travelling time between many survey and ASM locations here cannot be calculated due to missing or incomplete roads data.

Table 3.8: Effects of ASM and LSM on hours worked

	Hours worked (week)			
	(1)	(2)	(3)	(4)
ASM_{5km}	−0.144***	−0.183***	−0.224**	−0.114***
	(0.043)	(0.058)	(0.101)	(0.041)
$ASM_{5km} * GLSS5$	0.166***	0.194***	0.150	0.146***
	(0.045)	(0.061)	(0.110)	(0.044)
$ASM_{5km} * GLSS6$	0.111**	0.147**	0.227**	0.080*
	(0.046)	(0.060)	(0.109)	(0.047)
LSM_{5km}	0.074	0.004	−0.057	0.141
	(0.101)	(0.135)	(0.170)	(0.110)
$LSM_{5km} * GLSS5$	0.948***	1.544***	1.501***	0.342
	(0.270)	(0.323)	(0.549)	(0.297)
$LSM_{5km} * GLSS6$	0.396	0.605	1.662***	0.225
	(0.336)	(0.386)	(0.499)	(0.323)
Mean (dep. var.)	38.435	40.499	33.474	36.473
SD (dep. var.)	21.207	21.491	22.658	20.744
Observations	22065	10736	4146	11329
R^2	0.226	0.272	0.403	0.191
Household controls	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes
Distance brackets <= 50km	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at GLSS cluster level in parentheses. All regressions include sample weights. Total number of hours worked in last 7 days, conditional of being employed. Young workers defined as under the age of 30. Additional distance brackets of the number of artisanal mines and large-scale gold production include 5-10, 10-20, 20-30, 30-40, 40-50km.

Table 3.9: Effects of ASM and LSM on local prices

	Cluster prices (1)	Cluster prices (2)
ASM_{5km}	–0.028	–0.004**
	(0.027)	(0.002)
$ASM_{5km} * GLSS5$		0.002
		(0.003)
$ASM_{5km} * GLSS6$		–0.027
		(0.025)
LSM_{5km}	0.055	0.003
	(0.039)	(0.010)
$LSM_{5km} * GLSS5$		–0.320***
		(0.090)
$LSM_{5km} * GLSS6$		0.105
		(0.078)
Observations	488	892
R^2	0.276	0.993

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at GLSS cluster level in parentheses. Additional distance brackets of the number of artisanal mines and large-scale gold production include 5-10, 10-20, 20-30, 30-40, 40-50km.

mining on the other hand show an increase from 55 percent in 1988 (GLSS 2) to 79 percent in 2012/13 (GLSS 6).[12]

A second indicator of increased immigration to ASM areas is shown in table 3.10. Using the same control variables as in the main specification (age, education, sex, religion amongst others) the probit regression in column 1 shows that individuals in ASM areas in 2012/13 are 0.5 percent more likely to have immigrated. This effect does not come from return migration (column 2), but is likely driven by males (3) and young males (4). Immigration therefore might be one reason why incomes do not increase.

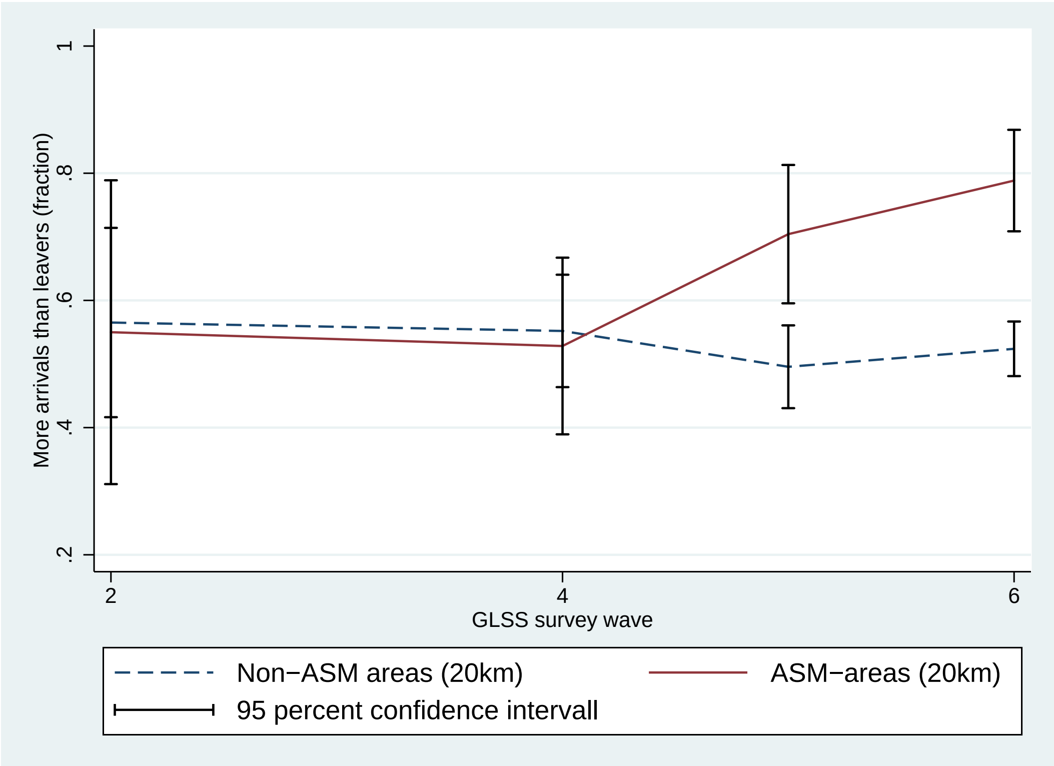


Figure 3.7: Share of communities reporting more arrivals than departures

ASM area if at least one detected ASM site within 20 kilometres, non-ASM area otherwise. Unconditional mean of cluster officials reporting more arrivals than departures in the past ten years.

Another explanation from workers that migrate to mining areas is that income from artisanal mining may be sent back to workers' homes in the form of remittances. The income gains generated in areas next to artisanal mines therefore may be diluted and spread out across the country.[13] If this was the case, expenditure on remittances should be higher in areas with ASM compared to those without.

Fortunately, GLSS collects remittance expenditure and income at the household level. I transform these values in 2010 real terms, as before with total income and expenditure,

[12]Note that 20 kilometres is used to define an ASM area here due to the low number of community-level observations.

[13]Given the large number of foreign individuals engaged in Ghana's ASM sector, it is also likely that a sizeable share of incomes is sent abroad. However to test this is outside of the scope of this thesis.

Table 3.10: Effects of ASM and LSM on immigration

	Moved here (1)	Returned here (2)	Moved here (males only) (3)	Moved here (males < 35 y.) (4)
ASM_{5km}	-0.001	0.004**	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.003)
$ASM_{5km} * GLSS5$	-0.001	-0.004*	0.001	0.002
	(0.002)	(0.002)	(0.002)	(0.003)
$ASM_{5km} * GLSS6$	0.005**	-0.001	0.005**	0.005*
	(0.002)	(0.002)	(0.002)	(0.003)
LSM_{5km}	0.008	-0.010**	0.017**	0.012
	(0.007)	(0.004)	(0.008)	(0.009)
$LSM_{5km} * GLSS5$	0.005	0.007	0.004	0.002
	(0.010)	(0.011)	(0.010)	(0.014)
$LSM_{5km} * GLSS6$	-0.029*	0.011	-0.042**	-0.041**
	(0.017)	(0.013)	(0.017)	(0.016)
Observations	41080	41080	19418	12739
Pseudo R^2	0.108	0.134	0.122	0.104

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Unit of observation are individuals. Clustered standard errors at GLSS cluster level in parentheses. Based on probit estimation. All regressions include sample weights.

and regress them within the previous framework. Figure 3.8 shows the interacted coefficients from GLSS wave 6 and the different ASM distance bands. Both remittance income and expenditure are essentially flat and therefore not affected by the presence of ASM.

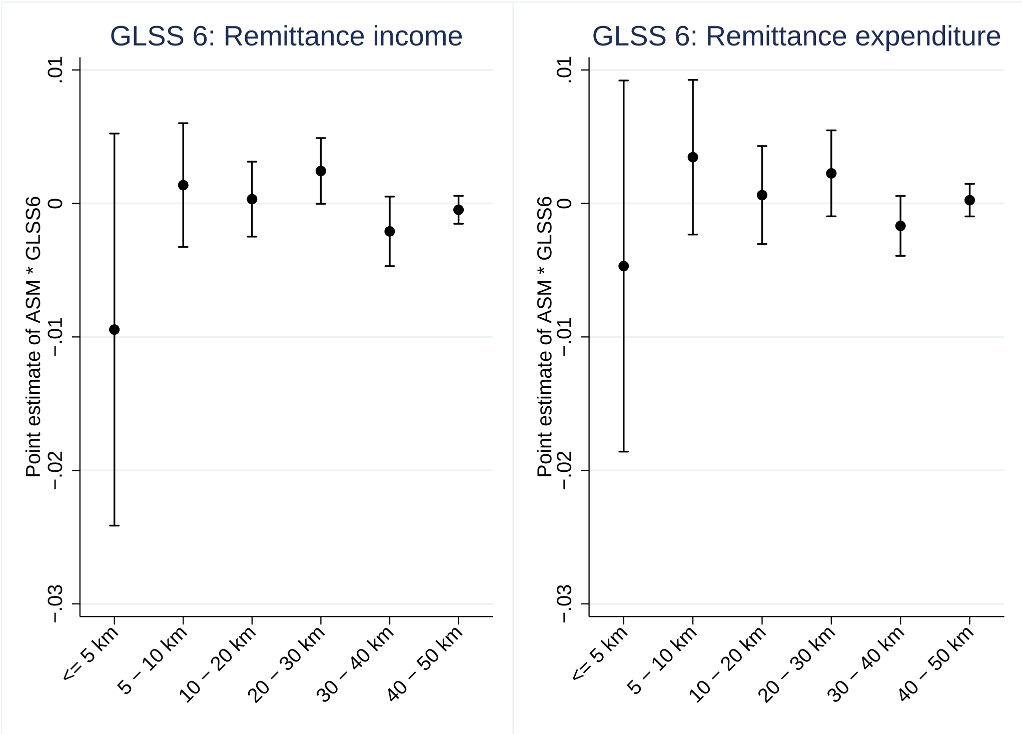


Figure 3.8: Effect of ASM on remittances

95 percent confidence intervalls are used. Point estimates are the coefficients of $ASM * GLSS6$ on remittance income and expenditure for different distance bands from the specification described in equation (3.1).

Overall, only immigration seems to contribute to the zero real income effect in ASM areas, while prices and remittances show no responses. An alternative view on these results is that the overall zero income effects masks heterogeneity of outcomes. To explore if different individuals or sectors benefit from ASM, I differentiate by industry of employment, income quantile, level of education and region.

First, I look at different income effects by industry of employment to see if income gains in one sector potentially negate reductions in others to make for a zero overall effect. Figure 3.9 shows the effect of ASM in 2012/13 on per-adult incomes by distance band and primary industry of employment of the household head. This is essentially the main specification, while cutting the sample by four broad industries agriculture, manufacturing, trade and other services. Here, the results show that the zero income effect occurs within each industry. Different sectors therefore do not seem to be affected differently by artisanal mining.

Next I check whether the income gains occur for poorer or richer households exclusively. For this I re-run the main specification in the form of a quantile regression using

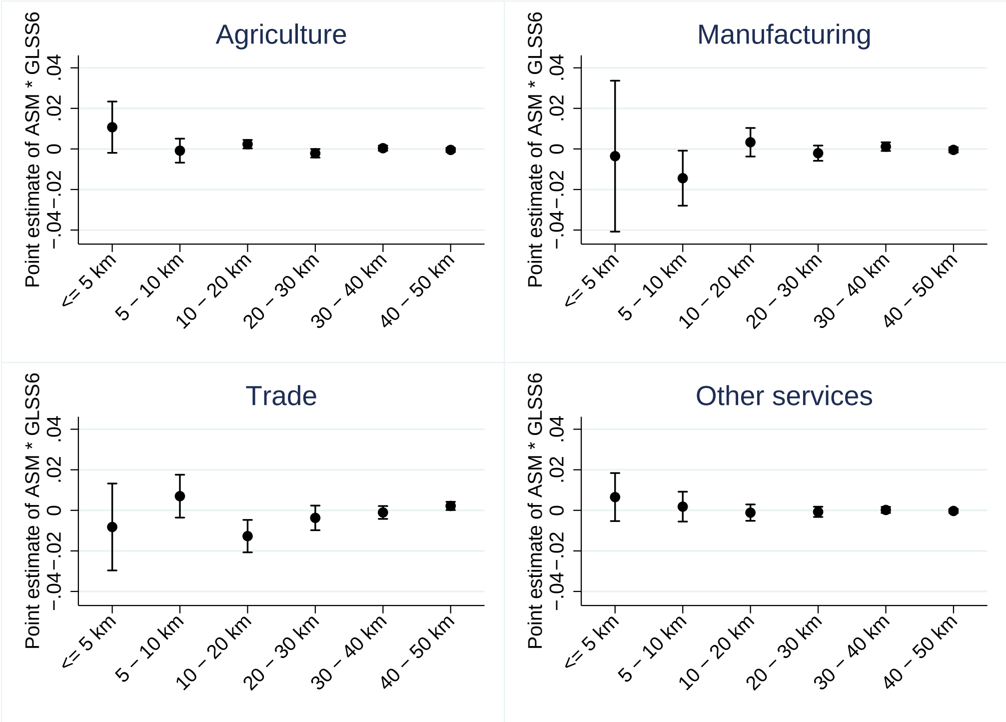


Figure 3.9: Effect of ASM on net incomes by distance and industry

Notes: Net income in log real per-adult equivalent values. The plotted coefficients show the interaction effect between ASM and GLSS6.

the 10th, 25th, 50th, 75th and 90th percentile. As with the segmentation by industry of employment, there is no ASM effect at either income quantile.

In table 3.12 I test whether the mining effects differ by educational attainment. Splitting the sample between household heads with primary education or below (1) and above primary education (2) shows that there is no income gain for neither education level.

Finally, table 3.13 splits the sample between the largest region Ashanti, which has most of the ASM activity, and the remaining ones combined. Interestingly, in the Ashanti region there is weak support for a positive ASM, along with LSM, effect on incomes, while the effect is reverse for the other regions. However, note that this sample split reduces the sample size considerably, such that there is not enough variation in column 2 to identify the effects of ASM and LSM separately.[14]

[14]As an additional test, I estimate the ASM effect on incomes in areas with and without LSM activity. The results are identical to the previous ones.

Table 3.11: Quantile regression results for income

	(1) OLS	(2) Q(0.10)	(3) Q(0.25)	(4) Q(0.50)	(5) Q(0.75)	(6) Q(0.90)
ASM_{5km}	-0.002	-0.002	-0.003	-0.002	-0.003	-0.002
	(0.004)	(0.006)	(0.003)	(0.004)	(0.003)	(0.004)
$ASM_{5km} * GLSS5$	0.004	0.007	0.005	0.003	0.004	0.002
	(0.003)	(0.006)	(0.003)	(0.005)	(0.003)	(0.004)
$ASM_{5km} * GLSS6$	0.002	-0.003	-0.000	0.004	0.003	0.004
	(0.004)	(0.007)	(0.003)	(0.004)	(0.003)	(0.004)
LSM_{5km}	-0.008	-0.013	-0.003	-0.011	-0.013	-0.002
	(0.007)	(0.015)	(0.007)	(0.007)	(0.012)	(0.013)
$LSM_{5km} * GLSS5$	0.102***	0.122***	0.088***	0.084***	0.103***	0.079
	(0.019)	(0.027)	(0.020)	(0.019)	(0.018)	(0.060)
$LSM_{5km} * GLSS6$	0.009	0.000	0.005	0.023	0.003	-0.012
	(0.015)	(0.015)	(0.015)	(0.016)	(0.017)	(0.024)
Mean (dep. var.)	6.818	6.818	6.818	6.818	6.818	6.818
SD (dep. var.)	1.486	1.486	1.486	1.486	1.486	1.486
Observations	12389	12389	12389	12389	12389	12389
R^2	0.289					
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Distance brackets <= 50km	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Net income in log real per-adult equivalent values. Clustered standard errors at GLSS cluster level in parentheses for OLS, heteroscedasticity robust standard errors for quantile regressions. All regressions include sample weights. Additional distance brackets of the number of artisanal mines and large-scale gold production include 5-10, 10-20, 20-30, 30-40, 40-50km.

Table 3.12: Effects of mining on income by education

	Primary education or below (1)	Above primary education (2)
ASM_{5km}	0.002	–0.005
	(0.004)	(0.004)
$ASM_{5km} * GLSS5$	–0.000	0.006*
	(0.005)	(0.004)
$ASM_{5km} * GLSS6$	–0.006	0.005
	(0.006)	(0.005)
LSM_{5km}	0.015	–0.021***
	(0.010)	(0.007)
$LSM_{5km} * GLSS5$	0.096***	0.099***
	(0.029)	(0.022)
$LSM_{5km} * GLSS6$	–0.019	0.022
	(0.025)	(0.014)
Mean (dep. var.)	6.389	7.086
SD (dep. var.)	1.412	1.468
Observations	4981	7408
R^2	0.241	0.274
Household controls	Yes	Yes
Distance brackets <= 50km	Yes	Yes
District FE	Yes	Yes
Time FE	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at GLSS cluster level in parentheses. All regressions include sample weights. Net income in 2010 log per-adult equivalent household values. Additional distance brackets of the number of artisanal mines and large-scale gold production include 5-10, 10-20, 20-30, 30-40, 40-50km.

Table 3.13: Effects of mining on income by region

	Ashanti (1)	All other regions (2)
ASM_{5km}	–0.008	0.019*
	(0.006)	(0.011)
$ASM_{5km} * GLSS5$	0.008	–0.017*
	(0.006)	(0.010)
$ASM_{5km} * GLSS6$	0.017*	–0.019*
	(0.009)	(0.011)
LSM_{5km}	–0.029**	–0.009
	(0.012)	(0.008)
$LSM_{5km} * GLSS5$	0.129***	0.154
	(0.028)	(0.152)
$LSM_{5km} * GLSS6$	0.085**	0.000
	(0.033)	(.)
Mean (dep. var.)	7.029	6.685
SD (dep. var.)	1.600	1.394
Observations	4308	8081
R^2	0.341	0.250
Household controls	Yes	Yes
Distance brackets <= 50km	Yes	Yes
District FE	Yes	Yes
Time FE	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at GLSS cluster level in parentheses. All regressions include sample weights. Net income in 2010 log per-adult equivalent household values. Additional distance brackets of the number of artisanal mines and large-scale gold production include 5-10, 10-20, 20-30, 30-40, 40-50km.

3.6 Conclusion

This chapter has analysed if and how artisanal and small-scale mining affects economic outcomes at the local level. For this I have looked at individual and household level evidence and nighttime lights in combination with the novel ASM data introduced in chapter 2. The results clearly show that local economic outcomes, in particular income, are not affected by nearby artisanal mining operations, while there is some evidence for beneficial effects of large-scale operations instead. The zero result is not driven by heterogeneity in education levels, industries and income level. It seems likely however that immigration, which is larger in ASM areas than outside, contributes to stagnant incomes.

Chapter 4

Health and environmental effects of artisanal and industrial mining

4.1 Introduction

Mineral extraction can yield material wealth, but may also bring about environmental destruction and health hazards along with it. These latter, negative externalities are the focus of this chapter. Case study evidence to date suggests negative health and environmental effects of ASM[1]. However, similar to the estimation of economic outcomes, these studies are small in scope, as they are lacking a good measure for artisanal mining. The evidence of large-scale mining is mixed. Benshaul-Tolonen (2018) finds that child mortality falls after industrial mines open in Africa. Von der Goltz and Barnwal (2019) on the other hand find increased stunting in children as a consequence of lead pollution.

The first section documents child health effects of artisanal and industrial mining. Three common types of early childhood diseases are considered: malaria (proxied by fever), respiratory diseases (cough) and diarrhoea. These are among the main causes of early childhood deaths with 20 percent attributed to malaria, 13 percent to acute respiratory diseases and 8 percent to diarrhoea (World Health Organisation, 2015).

Child health may be affected by ASM through different channels. First, as the mining pits fill up with rainwater over time they become non-moving water bodies and thereby ideal breeding spots for malaria-carrying mosquitoes (Mantey et al. (2016) and World Health Organization (2016)). As a result, fever and malaria infection rates may be higher near ASM sites. Further, absorption of mercury may lead among many other things to diarrhoea. Finally cyanide, often used alongside mercury in the ASM process, when improperly stored or disposed can negatively affect respiration (World Health Organization, 2016). These negative effects on respiration may also be reinforced through the emission of dust particles during the preparation and production part of ASM. The incidence of

[1] See Rajaee et al. (2015), Long et al. (2015) on health effects and Cobbina (2012) on environmental effects of ASM.

fever, diarrhoea and cough, which are collected in DHS survey data, are therefore good indicators of health effects of ASM.[2] Note that many of these negative health externalities also extend to industrial large-scale mining. In particular, respiratory diseases may be related to air pollution from LSM operations (Aragón and Rud, 2016).

On the other hand, child health outcomes might be improved in the presence of ASM and LSM if the additional income improves nutrition and allows parents to purchase better health care. Given that the previous chapter on economic effects did not show higher incomes or consumption near ASM sites, this health-improving channel does not seem very plausible.

In the second section, I analyse environmental effects of ASM and LSM, using data on forest cover loss and vegetation. Deforestation may occur at different stages of resource extraction: before operations start, during operation to extend the mining site and as a consequence of mining if increased economic activity requires additional cleared land for development.

This chapter is structured as follows. Section 4.2 studies child health effects of ASM and LSM using DHS survey data. Child mortality and child diseases are reported separately. Section 4.3 analyses health and environmental effects using a grid-level approach.

4.2 Child health effects

4.2.1 Identification

To study the potential effects of ASM and LSM operations on child health, I make use of five recent DHS survey waves between 1993 and 2014. With this data two related outcomes can be addressed: child mortality and the incidence of different child diseases. The advantage of studying child mortality is that DHS collects complete child birth and death histories for all sample mothers. This allows to link the year of birth to active ASM and LSM operations near the residence of the mother. The disadvantage of that approach is that, for (lack of) recall reasons, no information on the reason of death is collected.

To address this caveat I study three different child disease symptoms, which are based on the child survey part of DHS. While this allows to provide more detail on the type of illness, the analysis is restricted to live children in each survey wave and a relatively short recall period of two weeks. Three major types of early childhood diseases are considered: malaria (proxied by fever), respiratory diseases (cough) and diarrhoea. As outlined in the introduction, these are among the main causes of early childhood deaths (World Health Organisation, 2015).

[2] Figure C.1 illustrates some of the proposed health channels. Images (a) and (b) show standing water bodies as mosquito breeding grounds that result from ASM operations, (c) shows the unfiltered release of mercury into ground water.

The identification strategy is similar to the one developed in Kotsadam et al. (2018), with the estimating equation taking the form:

$$\begin{aligned} Outcome_{it} = \beta_0 + \beta_1 ActiveASM_{it} + \beta_2 InactiveASM_{it} \\ + \beta_3 LSM_{it} + \beta_4 \mathbf{X}_{it} + \delta_{rt} + \gamma_m + \varepsilon_{it}, \end{aligned} \tag{4.1}$$

where, in the child mortality analysis, $Outcome_{it}$ is whether child i, born in year t died within 12 months. In this case, ASM_{it} indicates whether there was at least one active ASM site within five kilometres distance from child i during its birth year. LSM_{it} is one if any large-scale gold production took place during the birth year of child i. In the case of child diseases, $Outcome_{it}$ refers to the reported incidence of the symptoms cough, fever, diarrhoea within the last two weeks of the survey date. ASM_{it} and LSM_{it} are accordingly also based on the survey year.[3]

For child mortality and child diseases two different, increasingly demanding specifications are estimated. Specification 1 includes $ActiveASM_{it}$, LSM_{it}, controls $\mathbf{X}_{it}$ and δ_{rt} as region-birth year, or region-survey year for child diseases, fixed effects. The controls $\mathbf{X}_{it}$ include the child's birth order, the mother's age and years of education, cluster altitude in metres and month of the interview. Specification 2 adds the term $InactiveASM_{it}$, which is the number of inactive ASM sites within a radius. The number of inactive ASM sites is obtained from subtracting active ASM sites from all detected ASM sites in 2014. This term controls for differences in ASM areas before mines become active. In the case of child mortality, a different way to control for pre-treatment differences is possible. Therefore specification 3 includes active ASM sites as the treatment variable and mother fixed effects γ_m. In that case, only child outcomes of the same mother are compared. This eliminates any sources of unobserved heterogeneity that are fixed for each mother, including personal characteristics and selection into mining or non-mining area. In all specifications, sample weights are used and standard errors are clustered at the DHS cluster (village/town) level.

In the case of child mortality, an underlying assumption is that sample mothers do not migrate after giving birth. In the DHS survey, only the current location of mothers, and their children, is reported.[4] If mothers give birth in one area and subsequently migrate to another this introduces measurement error, which would cause a downward bias of the point estimates.

[3]As a robustness check, the number of detected active small-scale mines and amount of gold produced by large-scale mines within five kilometres is considered.

[4]For DHS 1998 to DHS 2008, a variable indicates the migration status of women. However, this variable was not included in DHS 2014. Controlling for migration status would therefore make the sample too small to draw meaningful inferences.

4.2.2 Child mortality

This section reports results on child mortality, using the birth record of DHS sample mothers. The main results are presented in table 4.1. Columns (1) to (3) show coefficients from increasingly demanding specifications based on equation 4.1 for a five kilometre mining treatment and columns (4) to (6) for a wider 10 kilometre one. According to this, being born near an active ASM site does not significantly affect the probability of infant death. This is true for all three specifications. Large-scale mining on the other hand shows a positive and significant relationship. Being born next to a large-scale mine thereby increases the chance within the first year of life by five to seven percent. For both ASM and LSM, there are no visible effect when considering a larger treatment area in columns (4) to (6). These results are robust to an alternative specification, using post 2010 interaction instead of the active ASM measure, see tables C.1 and C.2.

Table 4.1: Effects of ASM and LSM on infant mortality

	P(Infant death (<1 year))					
	(1)	(2)	(3)	(4)	(5)	(6)
Any $ActiveASM_{5km}$	-0.014	-0.021	-0.051			
	(0.014)	(0.016)	(0.038)			
Any $InactiveASM_{5km}$		0.009				
		(0.008)				
Any LSM_{5km}	0.068***	0.067***	0.048			
	(0.023)	(0.022)	(0.033)			
Any $ActiveASM_{10km}$				0.015	0.014	0.013
				(0.010)	(0.011)	(0.033)
Any $InactiveASM_{10km}$					0.002	
					(0.008)	
Any LSM_{10km}				0.006	0.006	0.022
				(0.022)	(0.022)	(0.028)
Mean (dep. var.)	0.059	0.059	0.063	0.059	0.059	0.063
SD (dep. var.)	0.235	0.235	0.244	0.235	0.235	0.244
Observations	9878	9878	8107	9878	9878	8107
R^2	0.017	0.017	0.391	0.016	0.016	0.391
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Birthyear x region FE	Yes	Yes	Yes	Yes	Yes	Yes
Mother FE	No	No	Yes	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at DHS cluster level in parentheses.

As related outcomes, table 4.2 shows coefficients for child death (1-5 years after birth) and neonatal death (within the first month after birth). Similar to infant death, ASM does not seem to affect child death in later years, while LSM does. In the case of neonatal death, neither are significant.

Table 4.2: Effects of ASM and LSM on child and neo-natal mortality

	P(Child death (1-5 years))			P(Neo-natal death (<1 month)		
	(1)	(2)	(3)	(4)	(5)	(6)
Any $ActiveASM_{5km}$	-0.020	-0.024	-0.048	-0.009	-0.005	-0.004
	(0.015)	(0.016)	(0.038)	(0.009)	(0.011)	(0.020)
Any $InactiveASM_{5km}$		0.006			-0.005	
		(0.009)			(0.008)	
Any LSM_{5km}	0.058**	0.057**	0.030	0.019	0.020	0.006
	(0.024)	(0.023)	(0.027)	(0.015)	(0.015)	(0.036)
Mean (dep. var.)	0.070	0.070	0.076	0.037	0.037	0.040
SD (dep. var.)	0.256	0.256	0.265	0.190	0.190	0.196
Observations	9878	9878	8107	9878	9878	8107
R^2	0.018	0.018	0.381	0.014	0.014	0.399
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Birthyear x region FE	Yes	Yes	Yes	Yes	Yes	Yes
Mother FE	No	No	Yes	No	No	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.
Notes: Clustered standard errors at DHS cluster level in parentheses.

4.2.3 Child diseases

Table 4.4 shows results for the three diseases for specification 1 (columns 1, 3, 5) and the more demanding specification 2 (columns 2, 4, 6) from equation 4.1. Similar to child mortality, ASM does not affect any of the three child diseases in a significant way. LSM however again is associated with increased cough and especially fever among children. Both results are consistent with the (zero) effect of ASM and LSM on infant mortality, documented in the previous section.

Table 4.5 shows results for a wider treatment area of ten kilometres. The coefficient on LSM for fever is still positive, but now insignificant. Interestingly, diarrhoea is now positively affected by ASM. Having an active ASM site within ten kilometres increases the probability for a child to have diarrhoea by four percent, even when controlling for selection effects. This might be the result of downstream pollution from ASM sites, but lacking the exact location of children this can't be tested.[5]

[5] Recall that DHS sample locations are randomly displaced up to five kilometres.

Table 4.3: Effects of ASM and LSM on infant mortality (continuous measures)

	P(Infant death (<1 year))		
	(1)	(2)	(3)
Any $ActiveASM_{5km}$	–0.000**	–0.000	0.001
	(0.000)	(0.000)	(0.000)
Any $InactiveASM_{5km}$		–0.000	
		(0.000)	
Any LSM_{5km}	0.005*	0.005*	–0.003
	(0.003)	(0.003)	(0.015)
Mean (dep. var.)	0.059	0.059	0.063
SD (dep. var.)	0.235	0.235	0.244
Observations	9878	9878	8107
R^2	0.016	0.016	0.391
Controls	Yes	Yes	Yes
Birthyear x region FE	Yes	Yes	Yes
Mother FE	No	No	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at DHS cluster level in parentheses.

Table 4.4: Effects of ASM and LSM on child diseases (binary treatment)

	Diarrhoea		Fever		Cough	
	(1)	(2)	(3)	(4)	(5)	(6)
Any $ActiveASM_{5km}$	-0.004	-0.011	-0.037	-0.025	-0.022	-0.025
	(0.018)	(0.021)	(0.032)	(0.036)	(0.026)	(0.035)
Any $InactiveASM_{5km}$		0.010		-0.017		0.005
		(0.017)		(0.022)		(0.023)
Any LSM_{5km}	0.003	0.003	0.093**	0.094**	0.065	0.065
	(0.046)	(0.047)	(0.041)	(0.040)	(0.052)	(0.052)
Mean (dep. var.)	0.158	0.158	0.191	0.191	0.188	0.188
SD (dep. var.)	0.365	0.365	0.393	0.393	0.391	0.391
Observations	6647	6647	6651	6651	6658	6658
R^2	0.028	0.028	0.023	0.023	0.036	0.036
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region-Survey FE	Yes	Yes	Yes	Yes	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at DHS cluster level in parentheses.

Table 4.5: Effects of ASM and LSM on child health (binary treatment, 10km)

	Diarrhoea		Fever		Cough	
	(1)	(2)	(3)	(4)	(5)	(6)
Any $ActiveASM_{10km}$	0.029*	0.039**	0.002	0.015	0.014	0.020
	(0.016)	(0.018)	(0.025)	(0.029)	(0.021)	(0.024)
Any $InactiveASM_{10km}$		-0.013		-0.018		-0.008
		(0.013)		(0.018)		(0.016)
Any LSM_{10km}	-0.041	-0.041	0.037	0.037	-0.004	-0.004
	(0.028)	(0.029)	(0.034)	(0.034)	(0.032)	(0.032)
Mean (dep. var.)	0.158	0.158	0.191	0.191	0.188	0.188
SD (dep. var.)	0.365	0.365	0.393	0.393	0.391	0.391
Observations	6647	6647	6651	6651	6658	6658
R^2	0.028	0.029	0.022	0.023	0.036	0.036
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region-Survey FE	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at DHS cluster level in parentheses.

Using continuous measures for ASM and LSM in table 4.6, the main results from table 4.4 are confirmed. Both for fever and cough, LSM is positively associated. ASM has a slight negative relation with fever in this specification, but the effect size is very small. Table C.3 in the appendix confirms these results using an alternative specification, where the interaction of ASM and LSM with a post 2008 dummy is used.[6]

The results from this section show that, contrary to common belief, artisanal mining is not associated with negative child health outcomes. Large-scale mining on the other hand is linked to increased infant and child mortality, which is supported by more frequent fever and cough symptoms among under five year olds.

4.3 Grid-level analysis of health and environmental effects

This section has two goals. First, I aim to complement the survey-based child health results from the previous section with a novel malaria grid dataset. Second, using the same grid framework outlined in the data chapter 2, I use data on forest loss and vegetation to identify environmental effects of ASM and LSM.

Figures 4.1 to 4.6 show trends in malaria infection prevalence and malaria clinical

[6] 2008 is used as the cut-off date here in order to include two post treatment waves. Results are robust to only using DHS wave 2014 as post treatment.

Table 4.6: Effects of ASM and LSM on child health (continuous treatment)

	Diarrhoea		Fever		Cough	
	(1)	(2)	(3)	(4)	(5)	(6)
$ActiveASMsites_{5km}$	0.000	0.000	-0.001*	-0.001**	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$InactiveASMsites_{5km}$		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)
$LSMprod_{5km}$	0.003	0.003	0.005***	0.005***	0.005**	0.005**
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Mean (dep. var.)	0.158	0.158	0.191	0.191	0.188	0.188
SD (dep. var.)	0.365	0.365	0.393	0.393	0.391	0.391
Observations	6647	6647	6651	6651	6658	6658
R^2	0.028	0.028	0.023	0.023	0.036	0.036
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region-Survey FE	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at DHS cluster level in parentheses.

incidence rates, as well as forest cover loss and Normalised Difference Vegetation Index (NDVI). Malaria rates (figures 4.1, 4.2 and 4.3) for 2-10 year old children have decreased strongly since 2000, from more than 60 to below 40 percent. However, both the geographical distribution and the distinction between ASM and non-ASM cells clearly show that this decrease hasn't been ubiquitous. Whereas the capital Accra and northern regions recorded malaria prevalence reductions of over 60 percent, the rate was only 30 to 45 percent in many rural and urban areas in the centre and south-west. The latter are also the areas where most ASM and LSM activity takes place. The time trend shown in figure 4.3 supports the notion that mining, and in particular artisanal and small-scale mining, is related to malaria. Until 2011, average malaria prevalence was lower in ASM areas, but has been higher since.

In terms of environmental damage, figure 4.4 suggests a similar pattern as in the case of malaria. Areas with ASM and LSM operations have over time experienced more forest loss than non-mining areas. This is also true in general when looking at annual development in figure 4.5. However, there are two surprising insights from figure 4.5. First, deforestation has always been slightly higher in ASM than non-ASM areas, even before most of the operations started around 2010. This counters the intuition that small-scale miners clear forest in order to operate. That this is not the case was confirmed qualitatively during field visits in Ghana, where miners reported that ASM mostly takes place outside forests. In their view, clearing forest is related to significant expenses, along with noise and attention. Most of the easily accessible material is further said to be found near riverbanks, where vegetation is naturally low.

The big spike in deforestation in 2013 is the second surprising insight from this figure. One potential explanation is that by that time ASM operations had become so large, that clearing forest to continue operating was necessary. While plausible, there is no qualitative or quantitative evidence of this channel. A technical reason for this spike could also be the change in methodology in producing the forest loss data from Hansen et al. (2013), which implemented an improved detection after 2011.

As an alternative measure of environmental effects figure 4.6) shows the percentage change in Normalised Difference Vegetation Index (NDVI) from 2000 to 2014. The general trend in vegetation degradation from the forest loss data can also be seen here. While the South of the country has deteriorated in vegetation, this also occurs in cells without ASM or LSM activity.

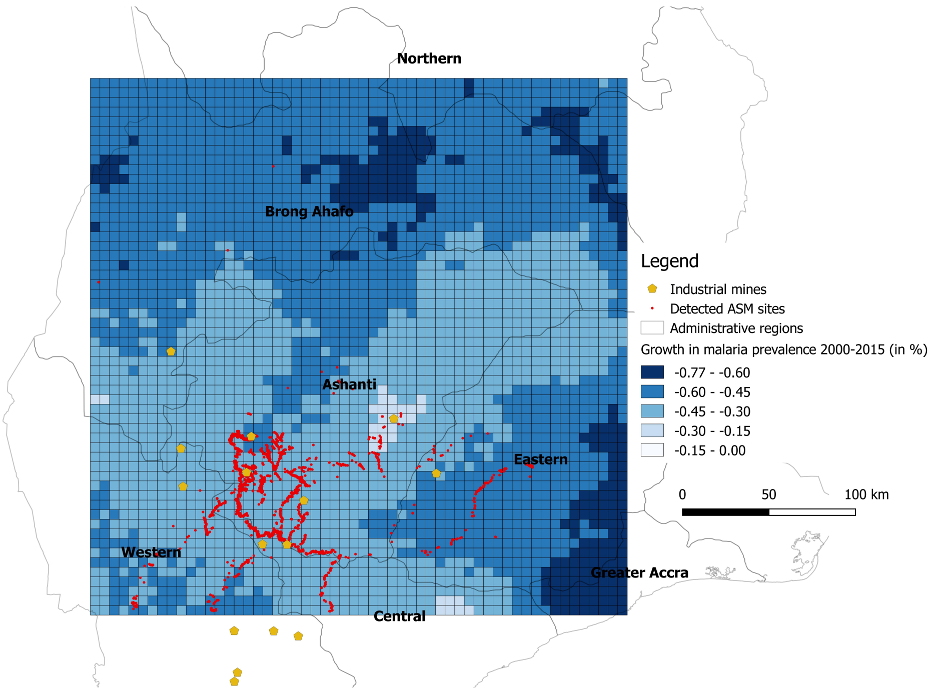


Figure 4.1: Reduction in malaria prevalence from 2000 to 2015

Source: author's visualisation based data from Bhatt et al. (2015). Notes: Growth is percentage change in malaria prevalence between 2000 and 2015.

4.3.1 Identification

The identification strategy in this section is identical to the one described for nighttime lights in section 3.4 of the previous chapter. The estimating equation takes the form:

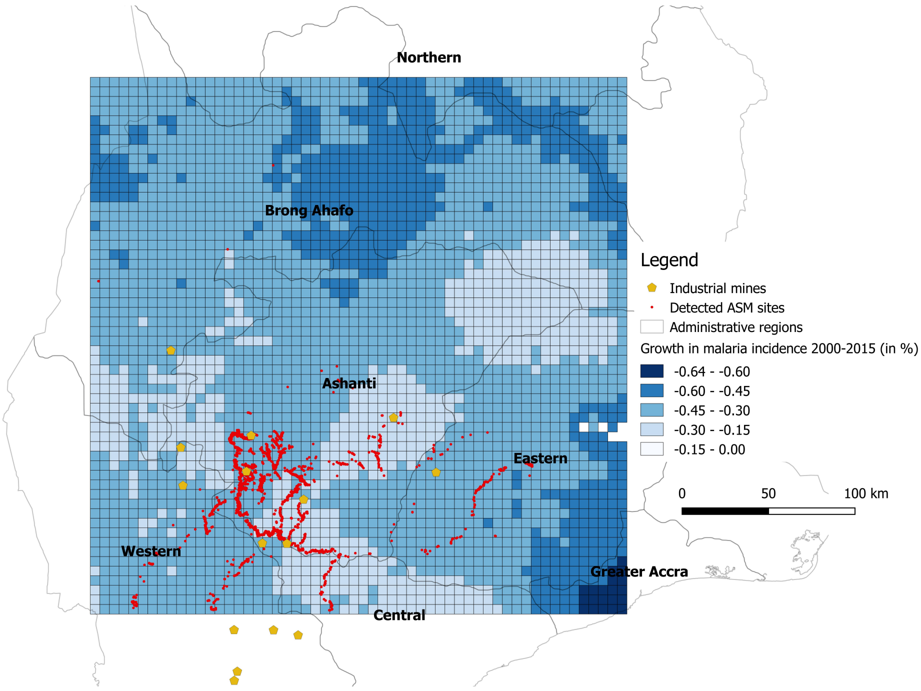


Figure 4.2: Reduction in malaria incidence from 2000 to 2015

Source: author's visualisation based data from Bhatt et al. (2015). Notes: Growth is percentage change in malaria incidence between 2000 and 2015.

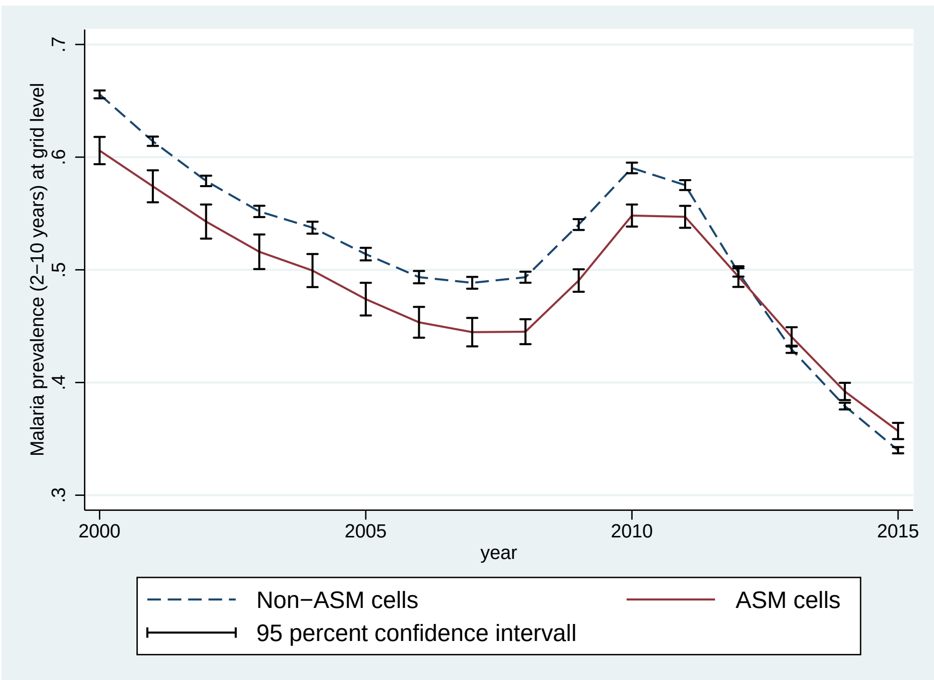


Figure 4.3: Malaria prevalence rates from 2000 to 2015

Source: author's visualisation based data from Bhatt et al. (2015). Unconditional averages of malaria prevalence are shown.

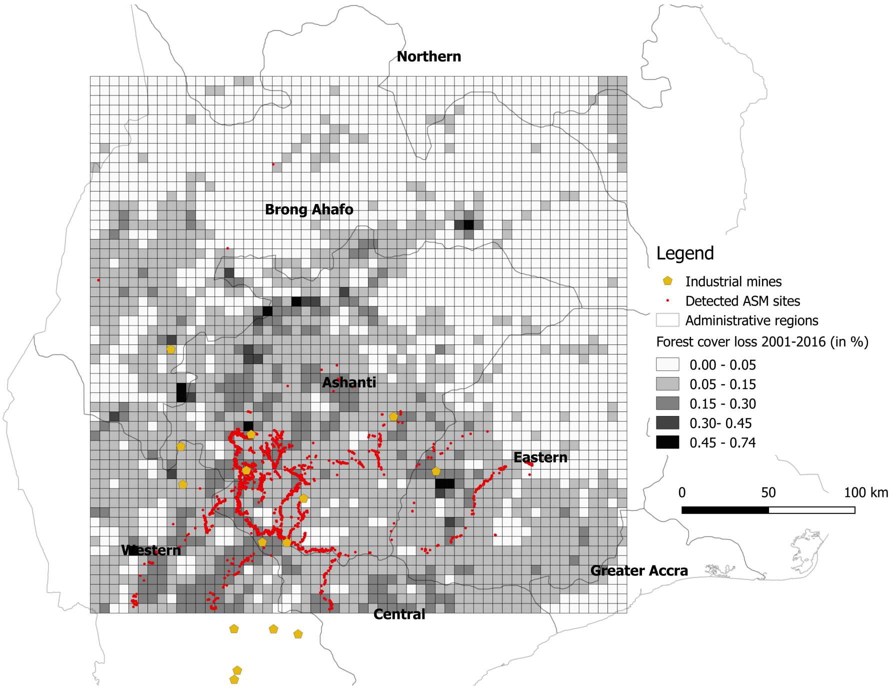


Figure 4.4: Total forest cover loss from 2001 to 2016

Source: author's visualisation based on Hansen et al. (2013). Notes: The forest cover loss number here is the cumulative fraction of each grid cell where forest was lost between 2001 and 2016.

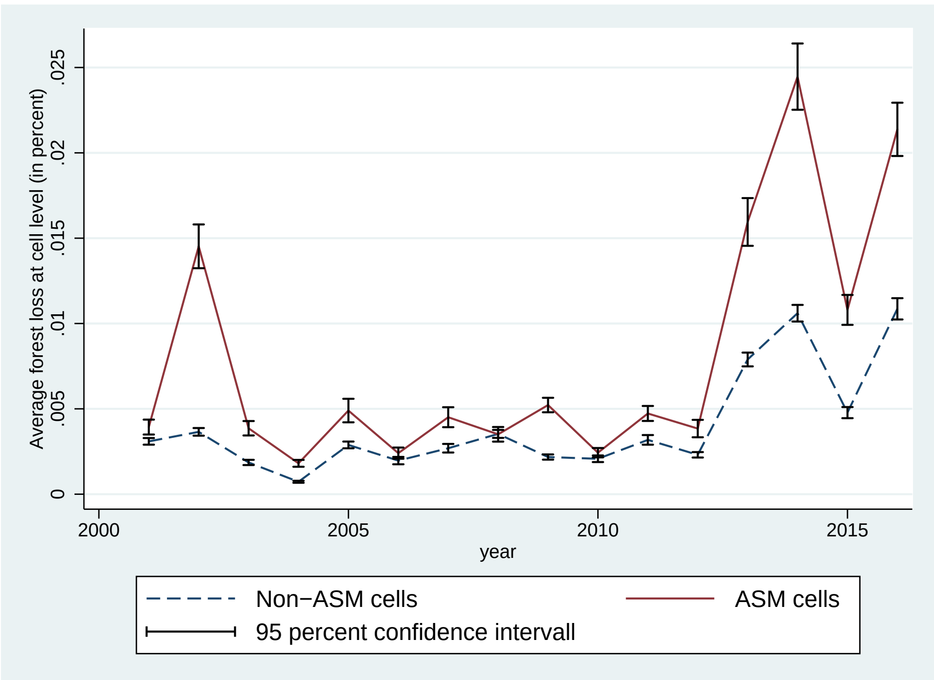


Figure 4.5: Annual mean forest cover loss from 2001 to 2016

Source: author's visualisation based on Hansen et al. (2013). Unconditional averages of forest cover loss are shown for cells with and without ASM.

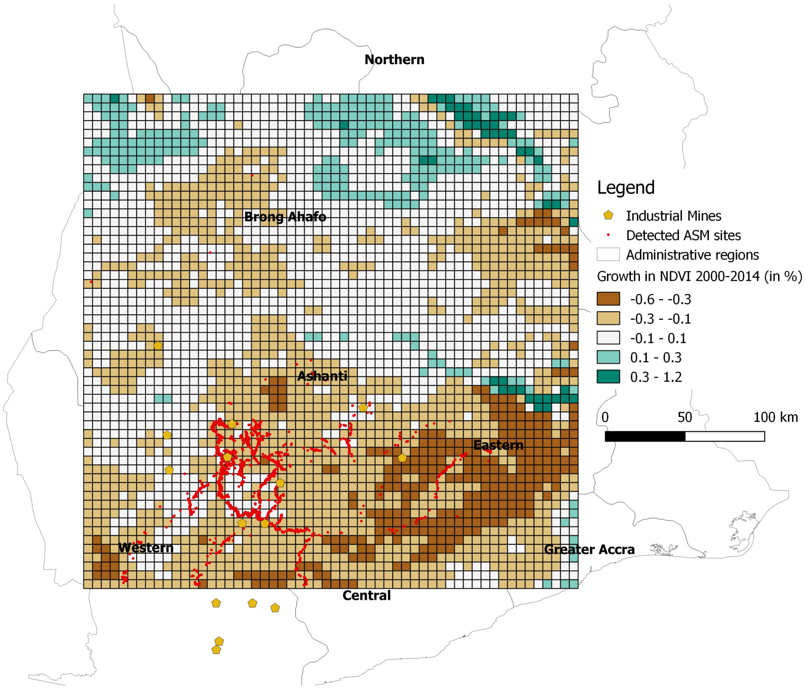


Figure 4.6: Growth in NDVI from 2000 to 2014

Source: author's visualisation based on AidData. Notes: Growth is percentage change in NDVI between 1982 and 2014.

$$\begin{aligned} Outcome_{it} = \beta_0 + \beta_1 ASM_i + \beta_2 (ASM_i * year_t) \\ + \beta_3 LSM_{it} + \beta_4 (LSM_{it} * year_t) \\ + \gamma_i + \delta_t + \varepsilon_{it}, \end{aligned} \tag{4.2}$$

where $Outcome_{it}$ is either the level of forest loss, NDVI or malaria prevalence. As before, ASM_i is a time-invariant dummy that is equal to one if there is at least one detected artisanal mine in cell i according to the satellite images from 2014. Similarly LSM_{it} is a dummy that takes one if there is an active large-scale mine in cell i in year t. $ASM_i * year_t$ and $LSM_{it} * year_t$ are year interactions of ASM and LSM. As artisanal mining increases in production over time due to increases in gold price and production techniques (see figure 2.1), the years after 2010 in the interaction $ASM_i * year_t$ will show whether ASM and non-ASM cells developed differently over time. δ_t are year-fixed effects and are used in all specifications.

As in the previous chapter, a simplified way of showing this relationship is comparing mining and non-mining cells before and after 2010. The estimating equation is then:

$$\begin{aligned} Outcome_{it} = \beta_0 + \beta_1 ASM_i + \beta_2(ASM_i * Post2010_t) \\ + \beta_3 LSM_{it} + \beta_4(LSM_{it} * Post2010_t) \\ + \gamma_i + \delta_t + \varepsilon_{it}, \end{aligned} \tag{4.3}$$

where $ASM_i * Post2010_t$ and $LSM_{it} * Post2010_t$ are now interactions of ASM and LSM with an indicator that is one for years after 2010 and zero otherwise.

4.3.2 Results

First, I test whether malaria infection prevalence is affected by the presence of ASM and LSM. Table 4.7 shows results for this. Columns (1) to (4) show increasingly demanding specifications. Comparing ASM and LSM cells against the full sample (column 1), shows that malaria prevalence in ASM cells after 2010 is significantly higher than in non-ASM cells, while controlling for cell fixed effects. The same is true for LSM post 2010, but not before. Column (3) shows that even in when only looking at ASM and LSM cells and their direct neighbours, these effects remain statistically significant, although smaller in size. Column (2) shows further that also neighbouring cells are affected.

Figure 4.7 illustrates the timing of the effect of artisanal mining on malaria infection rates. Here, only the reduced sample, consisting in ASM and LSM cells and their direct neighbours is used. From 2001 to around 2005 ASM and non-ASM areas follow the same trend in malaria prevalence, but from 2010 onwards ASM cells have consistently higher rates.

Overall, this suggests that the standing water bodies that remain during and after ASM operations take place do indeed provide breeding spots for malaria carrying mosquitoes, which in turn increase infection rates. A caveat of this estimation however is that the outcome is a modelled variable, based on many geographic covariates of derivations thereof (see figure A.6 in the appendix). These covariates, for example land cover and surface brightness can in turn be correlated with the ASM measure. This may introduce omitted variable bias with an unknown direction because of the many different covariates and derivations.

Next, I evaluate environmental effects of artisanal and large-scale mining in the form of forest loss and plant healthiness. Data on forest loss at the grid level is available for years 2001 to 2016. The coefficient on $ASM_i * Post2010_t$ is positive at 0.2 to 0.5 percentage points in all four specifications. Given the small average of forest cover loss of 0.4 percent per cell per year in the full sample, this represents a sizeable reduction. This suggests that artisanal mining is in fact associated with increased deforestation. Large-scale mining has larger coefficients than ASM, especially when considering post 2010, but the standard errors are also larger. Overall, both ASM and LSM seem to contribute

Table 4.7: Effects of ASM and LSM on malaria prevalence

	Malaria prevalence (in %)			
	(1)	(2)	(3)	(4)
$ASM * Post2010$	0.044***	0.013***	0.013***	0.008**
	(0.004)	(0.005)	(0.005)	(0.004)
LSM	–0.031	–0.029	–0.028	–0.049***
	(0.019)	(0.019)	(0.018)	(0.011)
$LSM * Post2010$	0.057***	0.048***	0.021**	0.025**
	(0.008)	(0.009)	(0.009)	(0.011)
$ASM_n * Post2010$		0.035***		
		(0.004)		
LSM_n		–0.025***		
		(0.006)		
$LSM_n * Post2010$		0.041***		
		(0.004)		
Mean (dep. var.)	0.516	0.516	0.496	0.488
SD (dep. var.)	0.145	0.145	0.121	0.109
Observations	50176	50176	9680	31104
R^2	0.603	0.607	0.683	0.006
Controls	No	No	No	No
Full sample	Yes	Yes	No	No
Neighbour sample	No	No	Yes	Yes
Cell FE	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes
Neighbour-pair FE	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at cell level in parentheses. Columns (1) and (2) use the full grid sample. Column (3) includes only ASM cells and their first-degree neighbours. Column (4) is the same sample as in (3), but non-ASM neighbours can appear more than once to form neighbour pairs.

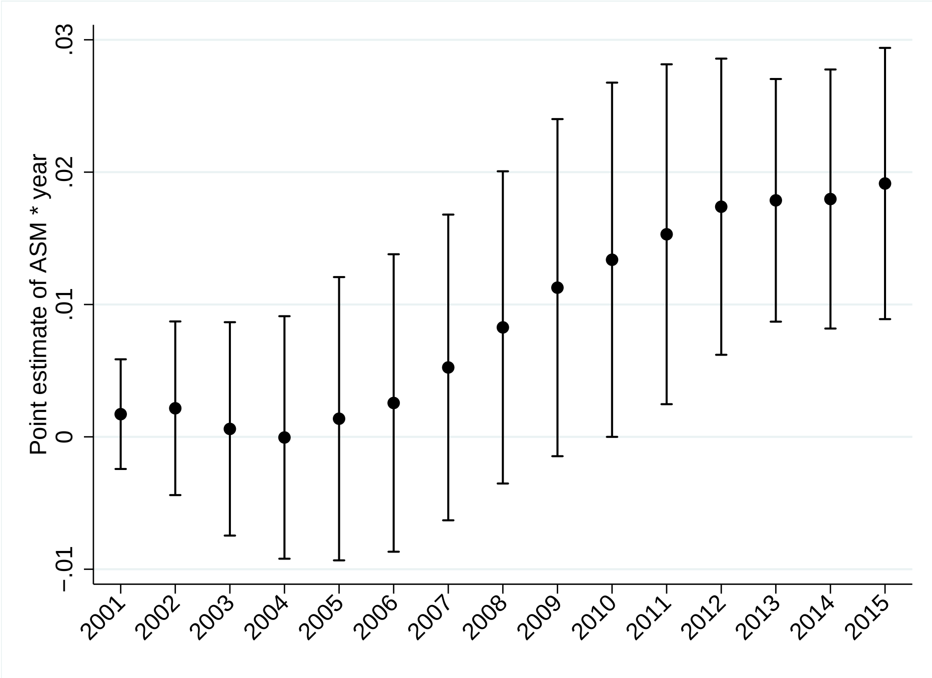


Figure 4.7: ASM and malaria prevalence - year interaction

Notes: Based on neighbour sample and equation 4.2

to deforestation. This is confirmed in an alternative specification in table C.4, which instead of annual loss shows results for forest cover based on the initial forest cover in 2000 (in percent of the cell) minus annual forest loss. In line with the previous table, the interaction between ASM and years after 2010 is negative in all four columns. Large-scale mining is insignificant in all specifications, except for the neighbour-pair fixed effect one in column (4), where it is large and negative. The LSM coefficients are also again quite large compared to ASM.

To see whether the effect on forest loss extends to an overall reduction in vegetation healthiness, I look at levels of normalised difference vegetation index (NDVI). Table 4.9 shows results from equation 4.3 for years 1981 to 2014. Past 2010, ASM cells seem to have become less green (column 1), but this effect disappears in the more demanding specifications. LSM is insignificant in all specifications. Overall, the NDVI results do not support the view that plant healthiness is reduced as a consequence of either ASM or LSM.

Table 4.8: Effects of ASM and LSM on forest cover loss

	Forest Cover Loss (in %)			
	(1)	(2)	(3)	(4)
$ASM * Post2010$	0.005***	0.002***	0.002***	0.002**
	(0.000)	(0.000)	(0.001)	(0.001)
LSM	0.004	0.005	0.005*	0.001
	(0.003)	(0.003)	(0.003)	(0.002)
$LSM * Post2010$	0.005	0.004	0.002	0.009
	(0.004)	(0.004)	(0.003)	(0.009)
$ASM_n * Post2010$		0.003***		
		(0.000)		
LSM_n		0.003**		
		(0.001)		
$LSM_n * Post2010$		0.003**		
		(0.001)		
Mean (dep. var.)	0.004	0.004	0.007	0.007
SD (dep. var.)	0.009	0.009	0.011	0.010
Observations	49698	49698	9680	29160
R^2	0.179	0.184	0.343	0.010
Controls	No	No	No	No
Full sample	Yes	Yes	No	No
Neighbour sample	No	No	Yes	Yes
Cell FE	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes
Neighbour-pair FE	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at cell level in parentheses. Columns (1) and (2) use the full grid sample. Column (3) includes only ASM cells and their first-degree neighbours. Column (4) is the same sample as in (3), but non-ASM neighbours can appear more than once to form neighbour pairs.

Table 4.9: Effects of ASM and LSM on NDVI

	Normalised Difference Vegetation Index			
	(1)	(2)	(3)	(4)
$ASM * Post2010$	−48.348***	12.353	9.579	−2.639
	(9.706)	(12.504)	(12.385)	(15.852)
LSM	−25.965	−30.974	−40.273	25.677
	(37.165)	(36.614)	(42.336)	(26.900)
$LSM * Post2010$	−10.841	12.717	65.293	6.285
	(63.369)	(60.536)	(63.421)	(56.486)
$ASM_n * Post2010$		−77.617***		
		(8.920)		
LSM_n		−51.318***		
		(10.864)		
$LSM_n * Post2010$		65.142***		
		(24.272)		
Mean (dep. var.)	2793.585	2793.585	3075.340	3078.426
SD (dep. var.)	565.648	565.648	451.606	427.330
Observations	84672	84672	16335	52488
R^2	0.496	0.497	0.623	0.000
Controls	No	No	No	No
Full sample	Yes	Yes	No	No
Neighbour sample	No	No	Yes	Yes
Cell FE	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes
Neighbour-pair FE	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at cell level in parentheses. Columns (1) and (2) use the full grid sample. Column (3) includes only ASM cells and their first-degree neighbours. Column (4) is the same sample as in (3), but non-ASM neighbours can appear more than once to form neighbour pairs.

4.4 Conclusion

This chapter has analysed health and environmental effects of artisanal and industrial mining. From a health perspective, child mortality and the incidence of child diseases are unaffected by ASM. LSM on the other hand has been shown to be associated with both increased child mortality and the incidence of fever and cough among children under five. Using a separate malaria dataset, areas increase in infection rates after ASM operations become active. However, because this connection is not supported by increased fever rates, the claim of ASM as a malaria driver remains only suggestive. From an environmental perspective, ASM and LSM both contribute to deforestation. Taken as a whole, these results clearly show negative externalities of resource extraction at the local level. However, because this is mostly driven by large-scale industrial mining, many of the proposed health hazards of ASM can not be supported.

Chapter 5

Conclusion

This thesis is among the first to provide empirical evidence on the economic, environmental and health effects of artisanal and small-scale mining. Using a novel data set on the exact GPS location of artisanal mines in connection with household survey data, I estimate, somewhat surprisingly, zero income gains in areas close to ASM operations. While not linked to increased prices, higher immigration into artisanal mining areas likely prohibits income increases. Large-scale mining on the other hand shows positive economic effects both using survey and nighttime lights data.

To estimate health and environmental costs of artisanal and industrial mining I use demographics and health survey data, gridded malaria data as well as forest loss and vegetation indicators. While child health is not affected, malaria incidence is relatively higher in ASM areas. Both artisanal and industrial mining contribute to forest loss, but overall plant healthiness is not affected. These results illustrate the trade-off between economic gains and health and environmental costs that natural resource extraction poses.

Better data on the intensive margin of ASM, for example in the form of production volumes, number of workers and type of artisanal extraction, can help to establish the precise channels through which artisanal mining affects different outcomes. Future work may also more accurately determine opening and activity of ASM sites from a panel of satellite images. Nonetheless, the empirical evidence presented here allows for a first examination of the local effects of artisanal and small-scale mining in an otherwise data-sparse environment.

Appendix A

Appendix to chapter 2

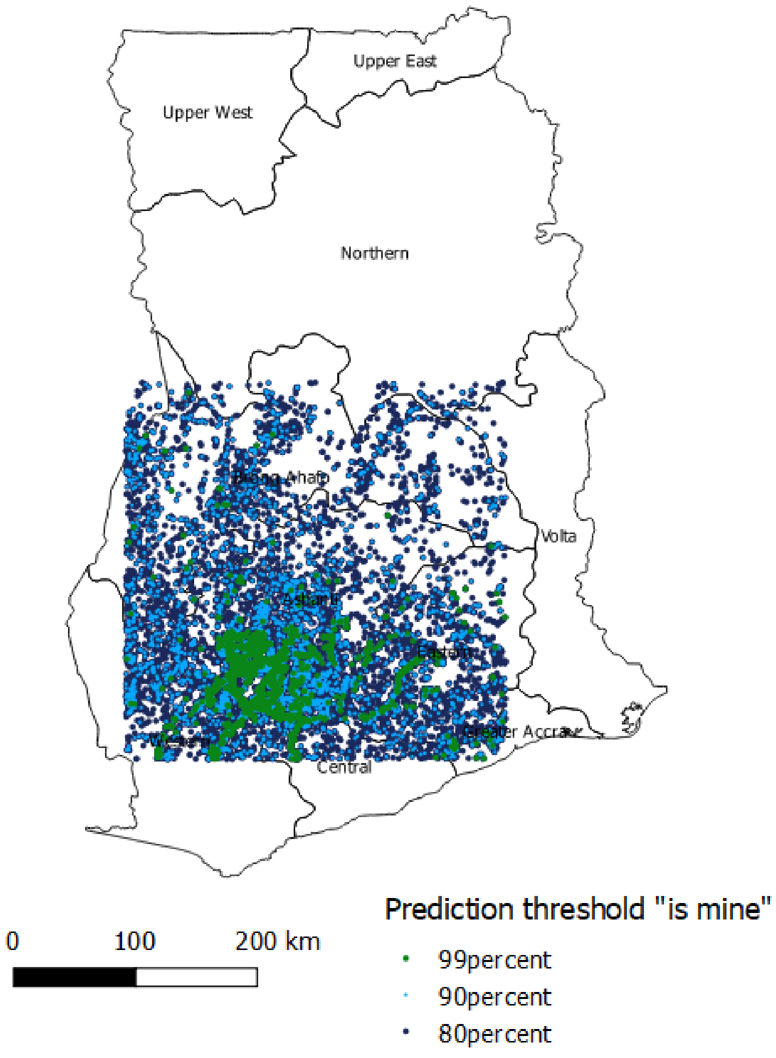


Figure A.1: Detected mines by prediction threshold

Source: Microsoft, author's visualisation.

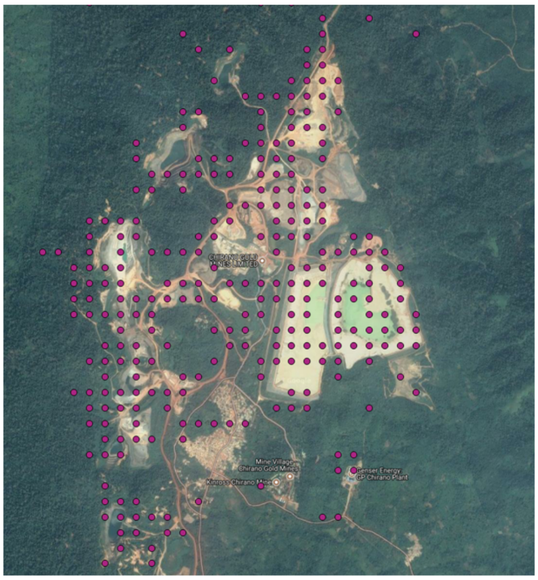


Figure A.2: Example of misprediction - large-scale mine

Source: Microsoft, author's visualisation.

Figure A.3: Example of misprediction - forest

Source: Microsoft, author's visualisation.

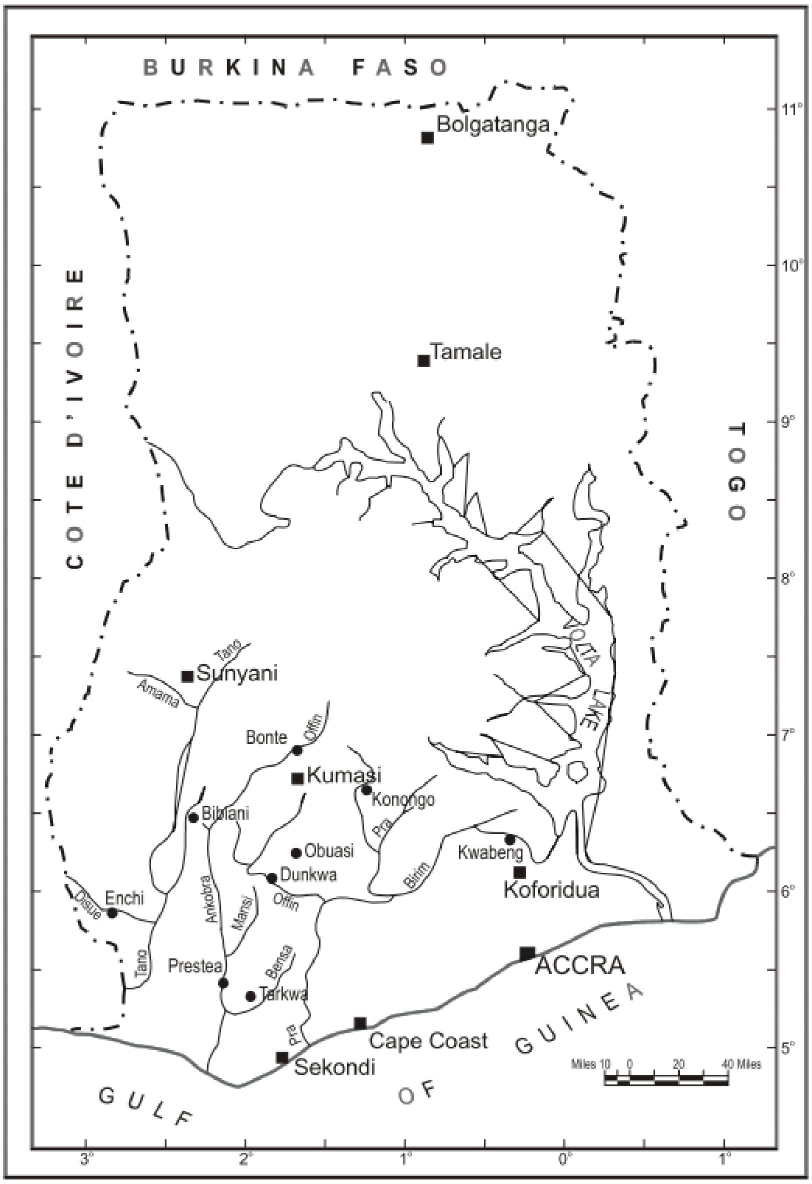


Figure A.4: Rivers in Ghana associated with placer gold deposits

Source: Hilson (2001).

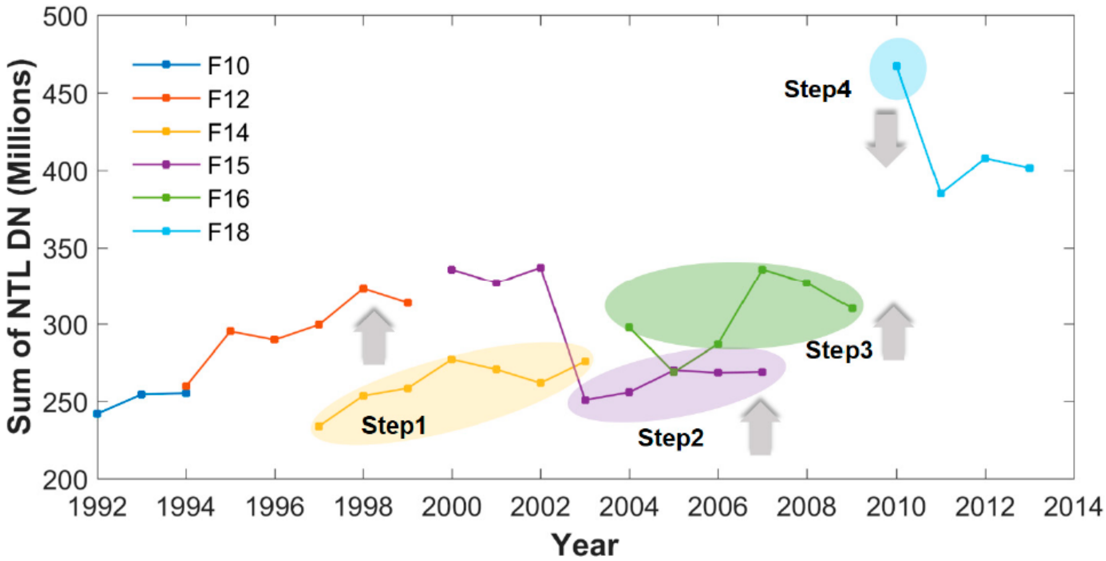


Figure A.5: Calibration of nighttime lights according to Li and Zhou (2017)

Source: Li and Zhou (2017).

Variable Class	Variable(s)	Source	Type
Temperature	Land Surface Temperature (daytime, night-time, and diurnal-flux)[7]	MODIS Product	Dynamic Monthly
Temperature Suitability	Temperature Suitability for *Plasmodium falciparum*[8]	Modelled Product	Dynamic Monthly
Precipitation	Mean Annual Precipitation[9]	WorldClim	Synoptic
Vegetation Vigour	Enhanced Vegetation Index[10]	MODIS Derivative	Dynamic Monthly
Surface Wetness	Tasselled Cap Wetness[11]	MODIS Derivative	Dynamic Monthly
Surface Brightness	Tasselled Cap Brightness[11]	MODIS Derivative	Dynamic Monthly
IGBP Landcover	Fractional Landcover[12]	MODIS Derivative	Dynamic Annual
IGBP Landcover Pattern	Landcover Patterns[12]	MODIS Derivative	Dynamic Annual
Terrain Steepness[13]	Slope Angle & Slope Angle Thresholds	SRTM Derivatives	Static
Topographically Redistributed Water[13]	Flow Accumulation & Topographic Wetness Index	SRTM Derivatives	Static
Elevation	Digital Elevation Model[14]	SRTM	Static
Human Population	AfriPop[15] ,GRUMP, and GPW[16]	Modelled Products	Dynamic Annual
Infrastructural Development	Accessibility[17] to Urban Centres and Night-time Lights	Modelled Product and VIIRS	Static
Moisture Metrics	Aridity and Potential Evapotranspiration[18]	Modelled Products	Synoptic

Figure A.6: Covariates used to model malaria prevalence by Bhatt et al. (2015)

Source: Bhatt et al. (2015).

Appendix B

Appendix to chapter 3

Table B.1: Sample summary statistics

	(1)	(2)	(3)	(4)
	mean	sd	min	max
Number of ASM sites (5 km)	4.16	18.75	0.00	152.00
Large-scale gold prod. (5 km) in tonnes	0.35	3.31	0.00	50.65
Log net income p.a. in GHC (real)	6.76	1.48	-3.79	13.38
Log expenditure p.a. in GHC (real)	6.11	1.30	2.11	11.32
Share any HH member in agriculture	0.56	0.50	0.00	1.00
Share any HH member in extractives	0.02	0.14	0.00	1.00
Share HH urban	0.44	0.50	0.00	1.00
Share HH with electricity	0.56	0.50	0.00	1.00
Share male	0.50	0.30	0.00	1.00
Average age	28.52	14.69	6.75	99.00
Education of HH head on scale 0-4	2.16	1.45	0.00	4.00
Share HH head literate	0.55	0.50	0.00	1.00
Share HH head moved here	0.33	0.47	0.00	1.00
Share HH head always here	0.67	0.47	0.00	1.00
Share HH head returned here	0.21	0.41	0.00	1.00
Cluster price index	1.00	0.04	0.92	1.10
Observations	12389	12389	12389	12389

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Based on main sample from GLSS 4-6. p.a. refers to per-adult equivalent. Log income and expenditure in real 2010 local GHC.

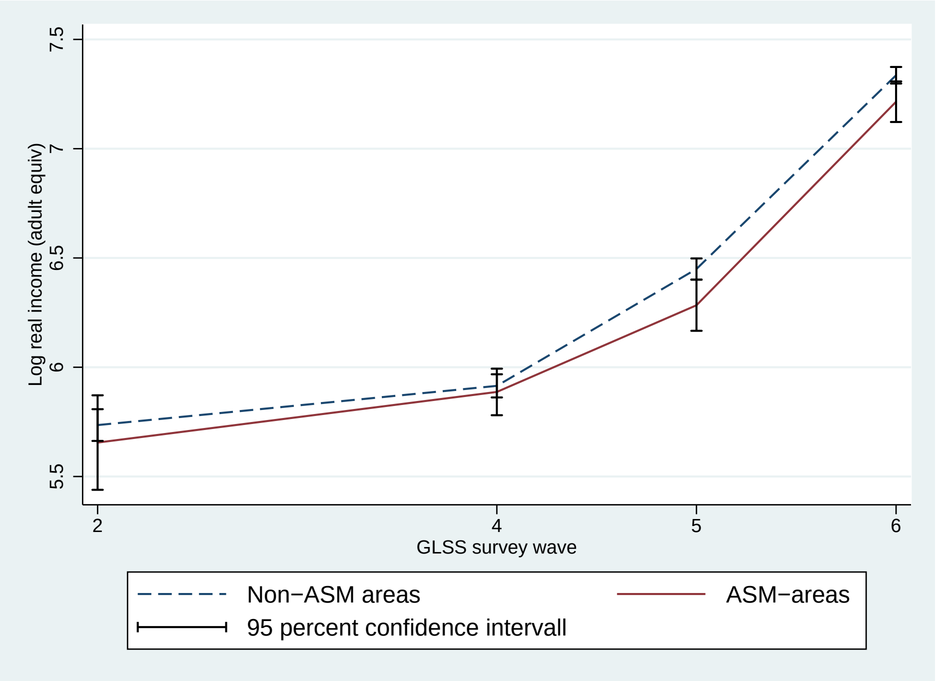


Figure B.1: Pre-trends for income, 0-5km distance band

Notes: Unconditional mean of log real per-adult equivalent net income for ASM areas (at least one ASM site) and non-ASM areas.

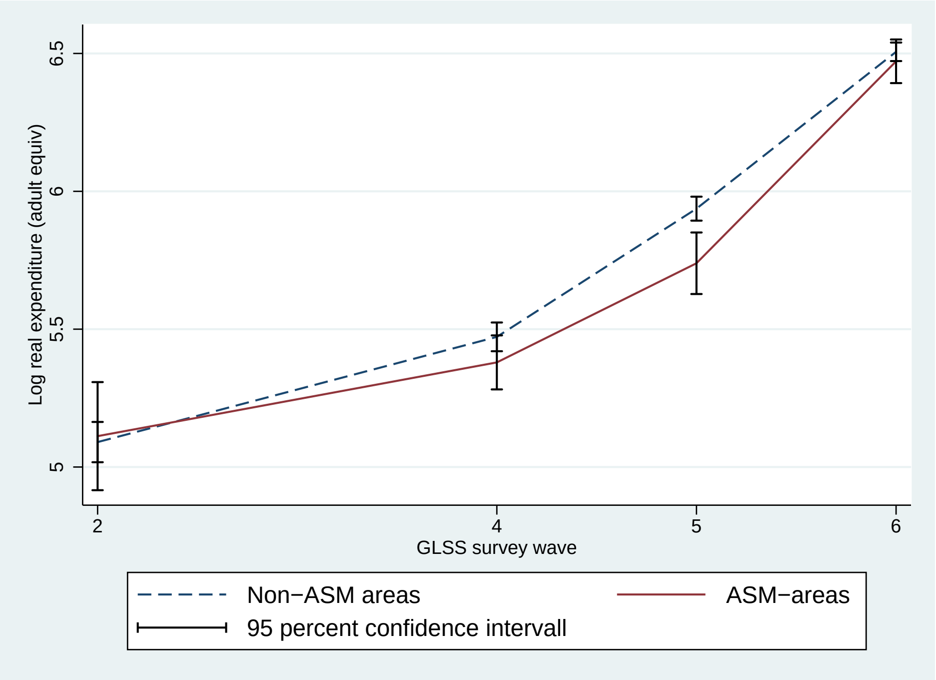


Figure B.2: Pre-trends for expenditure, 0-5km distance band

Notes: Unconditional mean of log real per-adult equivalent expenditure for ASM areas (at least one ASM site) and non-ASM areas.

Table B.2: Effect of ASM on income using inverse hyperbolic sine transformation of the dependent variable

	Income (1)	Income (2)	Income (3)	Income (4)
ASM_{5km}	–0.002	–0.010	–0.007	–0.010
	(0.007)	(0.006)	(0.008)	(0.008)
$ASM_{5km} * GLSS5$	0.016**	0.023***	0.015*	0.018**
	(0.008)	(0.007)	(0.008)	(0.008)
$ASM_{5km} * GLSS6$	–0.000	0.010	0.000	0.004
	(0.010)	(0.010)	(0.009)	(0.009)
LSM_{5km}	–0.054**	–0.056***	–0.034	–0.034
	(0.022)	(0.021)	(0.023)	(0.026)
$LSM_{5km} * GLSS5$	0.149***	0.142***	0.163***	0.158***
	(0.037)	(0.035)	(0.042)	(0.046)
$LSM_{5km} * GLSS6$	0.089**	0.096**	0.090*	0.093*
	(0.044)	(0.046)	(0.048)	(0.055)
Mean (dep. var.)	6.47	6.47	6.47	6.46
SD (dep. var.)	4.01	4.01	4.01	4.11
Observations	13413	13413	13413	12213
R^2	0.051	0.056	0.069	0.085
Household controls	Yes	Yes	Yes	Yes
Industry controls	No	No	No	Yes
Distance brackets <= 50km	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Region FE	Yes	No	No	No
Region-Time FE	No	Yes	No	No
District FE	No	No	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at GLSS cluster level in parentheses. All regressions include sample weights. Net income in 2010 per-adult equivalent household values, expressed as inverse hyperbolic sine transformation to include negative net incomes.

Table B.3: Effects of continuous ASM and LSM on nighttime luminosity

	Nighttime luminosity (in DN)			
	(1)	(2)	(3)	(4)
$ASMsites * Post2010$	0.000	0.002	0.001	0.002
	(0.003)	(0.003)	(0.003)	(0.002)
$LSMprod$	0.434***	0.433***	0.426***	1.092***
	(0.059)	(0.060)	(0.093)	(0.071)
$LSMprod * Post2010$	0.124	0.127	0.276	0.744*
	(0.155)	(0.155)	(0.256)	(0.447)
$ASMsites_n * Post2010$		–0.001		0.001
		(0.001)		(0.001)
$LSMprod_n$		0.154***		0.399***
		(0.050)		(0.064)
$LSMprod_n * Post2010$		0.137**		0.291**
		(0.062)		(0.144)
Mean (dep. var.)	1.163	1.163	0.990	0.947
SD (dep. var.)	8.447	8.447	3.033	2.859
Observations	68967	68967	12882	42768
R^2	0.020	0.022	0.149	0.233
Controls	No	No	No	No
Full sample	Yes	Yes	No	No
Neighbour sample	No	No	Yes	Yes
Cell FE	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes
Neighbour-pair FE	No	No	No	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at cell level in parentheses for columns (1) to (3). Columns (1) and (2) use the full grid sample. Column (3) includes only ASM and LSM cells and their first-degree neighbours. Column (4) is the same sample as in (3), but non-ASM neighbours can appear more than once to form neighbour pairs with more than one ASM cell. Column (4) uses Conley standard errors.

Table B.4: Effects of continuous active ASM and LSM on nighttime luminosity

	Nighttime luminosity (in DN)		
	(1)	(2)	(3)
Active $ASMsites$	0.001	0.001	0.008**
	(0.004)	(0.004)	(0.004)
$LSMprod$	0.456***	0.441***	0.933***
	(0.073)	(0.110)	(0.075)
Mean (dep. var.)	1.163	1.117	0.947
SD (dep. var.)	8.447	3.788	2.859
Observations	68967	12882	42768
R^2	0.020	0.089	0.107
Controls	No	No	No
Full sample	Yes	No	No
Neighbour sample	No	Yes	Yes
Cell FE	Yes	Yes	No
Year FE	Yes	Yes	Yes
Neighbour-pair FE	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at cell level in parentheses for columns (1) and (2). Column (1) uses the full grid sample. Column (2) includes only ASM and LSM cells and their first-degree neighbours. Column (3) is the same sample as in (2), but non-ASM neighbours can appear more than once to form neighbour pairs with more than one ASM cell. Column (3) uses Conley standard errors.

Appendix C

Appendix to chapter 4

(a) Recent ASM pit I

(b) Recent ASM pit II

(c) Chamfi crushing machine

(d) Mechanised ASM

Figure C.1: Small-scale mining operations in Ghana during field visit

Notes: Photographs taken during field visit in Ghana in March 2019.

Table C.1: Effects of ASM and LSM on infant mortality (alternative identification, binary measure)

	P(Infant death) (1)	P(Child death) (2)	P(Neonatal death) (3)
Any ASM_{5km}	0.009	0.006	–0.002
	(0.008)	(0.009)	(0.007)
Any $ASM_{5km} * Post2010$	–0.027*	–0.031*	–0.016
	(0.016)	(0.016)	(0.014)
Any LSM_{5km}	0.066***	0.055**	0.019
	(0.021)	(0.022)	(0.018)
Any $LSM_{5km} * Post2010$	0.010	0.019	0.005
	(0.061)	(0.061)	(0.044)
Mean (dep. var.)	0.059	0.070	0.037
SD (dep. var.)	0.235	0.256	0.190
Observations	9878	9878	9878
R^2	0.017	0.018	0.014
Controls	Yes	Yes	Yes
Birthyear x region FE	Yes	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at DHS cluster level in parentheses.

Table C.2: Effects of ASM and LSM on infant mortality (alternative identification, continuous measures)

	P(Infant death) (1)	P(Child death) (2)	P(Neonatal death) (3)
$ASMsites_{5km}$	–0.000	–0.000	–0.000
	(0.000)	(0.000)	(0.000)
$ASMsites_{5km} * Post2010$	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
$LSMprod_{5km}$	0.004	0.003	0.001
	(0.003)	(0.003)	(0.001)
$LSMprod_{5km} * Post2010$	0.011*	0.011*	0.004
	(0.006)	(0.006)	(0.003)
Mean (dep. var.)	0.059	0.070	0.037
SD (dep. var.)	0.235	0.256	0.190
Observations	9878	9878	9878
R^2	0.017	0.018	0.014
Controls	Yes	Yes	Yes
Birthyear x region FE	Yes	Yes	Yes

* p < 0.1, ** p < 0.05, *** p < 0.01.

Notes: Clustered standard errors at DHS cluster level in parentheses.

Table C.3: Effects of ASM and LSM on child diseases (alternative specification)

	Diarrhoea (1)	Fever (2)	Cough (3)
$ASMsites_{5km}$	0.000	0.000	0.000**
	(0.000)	(0.000)	(0.000)
$ASMsites_{5km} * Post2008$	−0.000	−0.001**	−0.001***
	(0.000)	(0.000)	(0.000)
$LSMprod_{5km}$	0.007**	0.006***	0.006***
	(0.003)	(0.001)	(0.002)
$LSMprod_{5km} * Post2008$	−0.031***	−0.009	−0.012
	(0.004)	(0.010)	(0.008)
Mean (dep. var.)	0.158	0.191	0.188
SD (dep. var.)	0.365	0.393	0.391
Observations	6647	6651	6658
R^2	0.030	0.023	0.037
Controls	Yes	Yes	Yes
Region-Survey FE	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at DHS cluster level in parentheses.

Table C.4: Effects of ASM and LSM on forest cover

	Forest Cover (in %)			
	(1)	(2)	(3)	(4)
$ASM * Post2010$	–0.027***	–0.008***	–0.005*	–0.020***
	(0.002)	(0.002)	(0.003)	(0.004)
LSM	–0.010	–0.011	–0.013	–0.126***
	(0.012)	(0.013)	(0.012)	(0.016)
$LSM * Post2010$	–0.039	–0.034	–0.014	–0.016
	(0.024)	(0.025)	(0.023)	(0.016)
$ASM_n * Post2010$		–0.021***		
		(0.002)		
LSM_n		–0.011		
		(0.009)		
$LSM_n * Post2010$		–0.039***		
		(0.010)		
Mean (dep. var.)	0.315	0.315	0.470	0.488
SD (dep. var.)	0.172	0.172	0.109	0.109
Observations	52834	52834	10285	31104
R^2	0.472	0.496	0.660	0.029
Controls	No	No	No	No
Full sample	Yes	Yes	No	No
Neighbour sample	No	No	Yes	Yes
Cell FE	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes
Neighbour-pair FE	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Clustered standard errors at cell level in parentheses. Columns (1) and (2) use the full grid sample. Column (3) includes only ASM cells and their first-degree neighbours. Column (4) is the same sample as in (3), but non-ASM neighbours can appear more than once to form neighbour pairs.

Bibliography

Acemoglu, D., C. García-Jimeno, and J. A. Robinson (2012). Finding Eldorado: Slavery and long-run development in Colombia. *Journal of Comparative Economics 40*(4), 534–564.

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

Aragón, F. M. and J. P. Rud (2016). Polluting Industries and Agricultural Productivity: Evidence from Mining in Ghana. *Economic Journal 126*(November), 1980–2011.

Aryeetey, E., O. Bafour, and D. Twerefou (2007). Impact of mining sector reforms on output, employment and incomes in Ghana, 1980–2002. Technical report, Institute of Statistical, Social and Economic Research (ISSER).

Asner, G. P., W. Llactayo, R. Tupayachi, and E. R. Luna (2013). Elevated rates of gold mining in the Amazon revealed through high-resolution monitoring. *Proceedings of the National Academy of Sciences of the United States of America 110*(46), 18454–9.

Basu, N., E. Clarke, A. Green, B. Calys-Tagoe, L. Chan, M. Dzodzomenyo, J. Fobil, R. N. Long, R. L. Neitzel, S. Obiri, E. Odei, L. Ovadje, R. Quansah, M. Rajaee, and M. L. Wilson (2015). Integrated Assessment of Artisanal and Small-Scale Gold Mining In Ghana—Part 1: Human Health Review. *International Journal of Environmental Research and Public Health 12*(8), 5143–5176.

Bazillier, R. and V. Girard (2017). The gold digger and the machine. Evidence on the distributive effect of the artisanal and industrial gold rushes in Burkina Faso. *Working Paper*, 1–47.

Benshaul-Tolonen, A. (2018). Local Industrial Shocks and Infant Mortality. *Economic Journal 129*(August), 1561–1592.

Berazneva, J. and T. S. Byker (2017). Does Forest Loss Increase Human Disease? Evidence from Nigeria. *American Economic Review: Papers & Proceedings 107*(5).

Berman, N., M. Couttenier, D. Rohner, and M. Thoenig (2017). This mine is mine! How minerals fuel conflicts in Africa. *American Economic Review 107*(6), 1564–1610.

Bhatt, S., D. J. Weiss, E. Cameron, D. Bisanzio, B. Mappin, U. Dalrymple, K. E. Battle, C. L. Moyes, A. Henry, P. A. Eckhoff, E. A. Wenger, O. Briët, M. A. Penny, T. A. Smith, A. Bennett, J. Yukich, T. P. Eisele, J. T. Griffin, C. A. Fergus, M. Lynch, F. Lindgren, J. M. Cohen, C. L. Murray, D. L. Smith, S. I. Hay, R. E. Cibulskis, and P. W. Gething (2015). The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature 526*(7572), 207–211.

Bluhm, R. and M. Krause (2018). Top Lights - Bright Cities and their Contribution to Economic Development. *CESifo Working Paper* (7411).

Bruederle, A. and R. Hodler (2017). Nighttime Lights as a Proxy for Human Development at the Local Level. *CESifo Working Paper* (6555).

Brunnschweiler, C. N. (2008). Cursing the Blessings? Natural Resource Abundance, Institutions, and Economic Growth. *World Development 36*(3), 399–419.

Brunnschweiler, C. N. and E. H. Bulte (2008). The resource curse revisited and revised: A tale of paradoxes and red herrings. *Journal of Environmental Economics and Management 55*(3), 248–264.

Bulte, E. H., R. Damania, and R. T. Deacon (2005). Resource intensity, institutions, and development. *World Development 33*(7), 1029–1044.

Buonanno, P., R. Durante, G. Prarolo, and P. Vanin (2015). Poor Institutions, Rich Mines: Resource Curse in the Origins of the Sicilian Mafia. *Economic Journal 125*(586), F175–F202.

Burgess, R., M. Hansen, B. a. Olken, P. Potapov, and S. Sieber (2012). The political economy of deforestation in the tropics. *Quarterly Journal of Economics 2001*(June), 1–48.

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

Chuhan-Pole, P., A. L. Dabalen, and B. C. Land (2017). *Mining in Africa: Are Local Communities Better Off?* Africa Development Forum.

Cobbina, S. J. (2012). Effect of Illegal Small-Scale Mining Operations on Vegetation Cover of Arid Northern Ghana. *Research Journal of Environmental and Earth Sciences 4*(6), 674–679.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1–45.

Cooke, E., S. Hague, and A. McKay (2016). The Ghana Poverty and Inequality Report: Using the 6th Ghana Living Standards Survey. Technical report.

Coulter, L. L., D. A. Stow, Y. H. Tsai, N. Ibanez, H. chien Shih, A. Kerr, M. Benza, J. R. Weeks, and F. Mensah (2016). Classification and assessment of land cover and land use change in southern Ghana using dense stacks of Landsat 7 ETM+ imagery. *Remote Sensing of Environment 184*, 396–409.

Crawford, G., C. Agyeyomah, G. Botchwey, and A. Mba (2015). The impact of Chinese involvement in small-scale gold mining in Ghana. Technical Report May, IGC.

Cust, J. and S. Poelhekke (2015). The Local Economic Impacts of Natural Resource Extraction. *Annual Review of Resource Economics 7*(1), 251–268.

Elvidge, C. D., D. Ziskin, K. E. Baugh, B. T. Tuttle, T. Ghosh, D. W. Pack, E. H. Erwin, and M. Zhizhin (2009). A fifteen year record of global natural gas flaring derived from satellite data. *Energies 2*(3), 595–622.

Financial Times (2017). Ghana crackdown on illegal gold mining inflames tensions with Beijing.

Fisher, E., R. Mwaipopo, W. Mutagwaba, D. Nyange, and G. Yaron (2009). "The ladder that sends us to wealth": Artisanal mining and poverty reduction in Tanzania. *Resources Policy 34*(1-2), 32–38.

Ghana Statistical Service (1998). Ghana Living Standards Survey (GLSS4) 1998, version 2.0.

Ghana Statistical Service (2008). Ghana Living Standards Survey (GLSS5) 2005, version 2.0.

Ghana Statistical Service (2012). Ghana Living Standards Survey (GLSS6) 2012, version 1.0.

GHEITI (2015). GHEITI report on the mining sector 2014. Technical Report December, Ministry of Finance, GHEITI.

GHEITI (2017). Data Analysis of Ghana's Extractive Sector.

Goodman, S., A. BenYishay, and D. Runfola (2017). Overview of the geo Framework. Technical report, AidData.

Google Earth Engine (2019). Earth Engine Data Catalog: USGS Landsat 7 Surface Reflectance Tier 1.

Gruendler, K. and A. Link (2019). Ethnic Inequality and the Wealth of Nations. *Working Paper*.

Gylfason, T., T. H. Thor, and G. Zoega (1999). A mixed blessing. Natural Resources and Economic Growth. *Macroeconomic Dynamics 3*, 204–225.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. a. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, J. R. G. Townshend, P. Patapov, R. Moore, M. Hancher, S. a. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommaredy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science 342*(November), 850–854.

Henderson, J. V., A. Storeygard, and D. N. Weil (2012). Measuring Economic Growth from Outer Space. *American Economic Review 102*(2), 994–1028.

Hilson, G. (2001). A contextual review of the Ghanaian small-scale mining industry. *Mining, Minerals and Sustainable Development* (76), 29.

Hilson, G. (2002). The environmental impact of small-scale gold mining in Ghana: identifying problems and possible solutions. *Geographical Journal 168*(1), 57–72.

Hilson, G. (2004). Structural Adjustment in Ghana: Assessing the Impacts of Mining-Sector Reform. *Africa Today 51*(2), 52–77.

Hilson, G. (2010). Child labour in African artisanal mining communities: Experiences from Northern Ghana. *Development and Change 41*(3), 445–473.

Hilson, G. (2016). Farming, small-scale mining and rural livelihoods in Sub-Saharan Africa: A critical overview. *Extractive Industries and Society 3*(2), 547–563.

Hilson, G., A. Hilson, and E. Adu-Darko (2014). Chinese participation in Ghana's informal gold mining economy: Drivers, implications and clarifications. *Journal of Rural Studies 34*(January), 292–303.

Hilson, G. and J. McQuilken (2014). Four decades of support for artisanal and small-scale mining in sub-Saharan Africa: A critical review. *Extractive Industries and Society 1*(1), 104–118.

Hilson, G. and C. Potter (2003). Why Is Illegal Gold Mining Activity so Ubiquitous in Rural Ghana? *African Development Review 15*(2-3), 237–270.

Hilson, G. and C. Potter (2005). Structural adjustment and subsistence industry: Artisanal gold mining in Ghana. *Development and Change 36*(1), 103–131.

Hirons, M. (2011). Locking-in carbon, locking-out livelihoods? Artisanal mining and REDD in Sub-Sahara Africa. *Journal of International Development 23*(1), 1140–1150.

Hsiang, S. M., K. C. Meng, and M. A. Cane (2011). Civil conflicts are associated with the global climate. *Nature 476*(7361), 438–441.

Jaques, E., B. Zida, M. Billa, C. Greffié, and J. Thomassin (2006). Artisanal and small-scale gold mines in Burkina Faso: today and tomorrow. In G. Hilson (Ed.), *Small-scale mining, rural subsistence and poverty in West Africa*. Practical Action.

Kitula, A. G. N. (2006). The environmental and socio-economic impacts of mining on local livelihoods in Tanzania: A case study of Geita District. *Journal of Cleaner Production 14*(3-4), 405–414.

Knutsen, C. H., A. Kotsadam, E. H. Olsen, and T. Wig (2017). Mining and Local Corruption in Africa. *American Journal of Political Science 61*(2), 1–15.

Kotsadam, A., G. Østby, S. A. Rustad, A. F. Tollefsen, and H. Urdal (2018). Development aid and infant mortality. Micro-level evidence from Nigeria. *World Development 105*, 59–69.

Lederman, D. and W. F. Maloney (2006). *Natural resources, neither curse nor destiny*. World Bank Publications.

Lederman, D. and W. F. Maloney (2008). In Search of the Missing Resource Curse. *Economia 9*(1), 1–57.

Leite, C. and J. Weidmann (1999). Does Mother Nature Corrupt? Natural Resources, Corruption, and Economic Growth. *IMF Working Paper*.

Li, X. and Y. Zhou (2017). A stepwise calibration of global DMSP/OLS stable nighttime light data (1992-2013). *Remote Sensing 9*(6).

Long, R. N., K. Sun, and R. L. Neitzel (2015). Injury risk factors in a small-scale gold mining community in Ghana's upper east region. *International Journal of Environmental Research and Public Health 12*(8), 8744–8761.

Mantey, J., F. O. Nimo, K. B. Nyarko, and A. Aubynn (2016). Operational dynamics of "Galamsey" within eleven selected districts of western region of Ghana. *Journal of Mining & Environment 8*(1), 11–34.

Mantey, J., K. Nyarko, and F. Owusu-Nimo (2016). Costed reclamation and decommissioning strategy for galamsey operations in 11 selected MDAs of the Western region , Ghana. Technical Report November, International Growth Centre.

National Geographic (2017). The Dirty Secrets of Gold Mining in Senegal.

NOAA (2019). Version 4 DMSP-OLS Nighttime Lights Time Series. *National Oceanic and Atmospheric Administration's National Centers for Environmental Information, Ashville*.

Ofosu-Mensah, E. A. (2011). Historical overview of traditional and modern gold mining in Ghana. *International Research Journal of Library, Information and Archival Studies 1*(1), 6–22.

Owusu, E. E. and G. Dwomoh (2012). The Impact of Illegal Mining on the Ghanaian Youth: Evidence From Kwaebibirem District In Ghana. *Research on Humanities and Social Sciences 2*(6), 86–93.

Pedelty, J., S. Devadiga, E. Masuoka, M. Brown, J. Pinzon, C. Tucker, E. Vermote, S. Prince, J. Nagol, C. Justice, D. Roy, J. Ju, C. Schaaf, J. Liu, J. Privette, and A. Pinheiro (2007). Generating a long-term land data record from the AVHRR and MODIS instruments. *International Geoscience and Remote Sensing Symposium (IGARSS)* (May), 1021–1024.

Rahm, M., B. Jullian, A. Lauger, R. de Carvalho, L. Vale, J. Totaram, K. Cort, M. Djojodikromo, M. Hardjoprajitno, S. Neri, R. Vieira, E. Watanabe, M. do Carmo Brito, P. Miranda, C. Paloeng, V. Moe Soe Let, S. Crabbe, and M. Calmel (2015). Monitoring the impact of gold mining on the forest cover and freshwater in the Guiana Shield. Technical report.

Rajaee, M., B. N. Sánchez, E. P. Renne, and N. Basu (2015). An investigation of organic and inorganic mercury exposure and blood pressure in a small-scale gold mining community in Ghana. *International Journal of Environmental Research and Public Health 12*(8), 10020–10038.

Saavedra, S. and M. Romero (2017). Local incentives and national tax evasion: The response of illegal mining to a tax reform in Colombia. *Working Paper*, 1–50.

Sachs, J. D. and A. M. Warner (1995). Natural Resource Abundance and Economic Growth. *NBER Working Paper Series*.

Sachs, J. D. and A. M. Warner (2001). The curse of natural resources. *European Economic Review 45*(4-6), 827–838.

Sala-i Martin, X. and A. Subramanian (2013). Addressing the natural resource curse: An illustration from Nigeria. *Journal of African Economies 22*(4), 570–615.

Schueler, V., T. Kuemmerle, and H. Schröder (2011). Impacts of surface gold mining on land use systems in Western Ghana. *Ambio 40*(5), 528–539.

Simoes, A. and C. Hidalgo (2011). The Economic Complexity Observatory: An Analytical Tool for Understanding the Dynamics of Economic Development. *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*.

Stelios Michalopoulos and E. Papaioannou (2014). National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics* (January), 151–213.

Teschner, B. A. (2012). Small-scale mining in Ghana: The government and the galamsey. *Resources Policy 37*(3), 308–314.

Unitar (2016). UNITAR's Operational Satellite Applications Programme Satellite: Mapping of Artisanal and Small Scale Gold Mining in Central Kalimantan, Indonesia. Technical Report February, UNITAR.

US Geological Service (2014). 2014 Minerals Yearbook: Ghana [Advance Release]. Technical report.

USGS (2019). Landsat 7. *U.S. Geological Survey*.

Van Straaten, P. (2000). Human exposure to mercury due to small scale gold mining in northern Tanzania. *The Science of the Total Environment 259*(1-3), 45–53.

Von der Goltz, J. and P. Barnwal (2019). Mines - The Local Wealth and Health Effects of Mineral Mining in Developing Countries. *Journal of Development Economics 139*, 1–16.

World Bank (2015). Socioeconomic Impact of Mining on Local Communities in Africa. Technical Report June 25.

World Health Organisation (2015). Ghana: WHO statistical profile. Technical report.

World Health Organization (2016). Artisanal and small-scale gold mining and health - Environmental and Occupational Health Hazards Associated with Artisanal and Small-scale Gold Mining. Technical report.

Young, N. E., R. S. Anderson, S. M. Chignell, A. G. Vorster, R. Lawrence, and P. H. Evangelista (2017). A survival guide to Landsat preprocessing. *Ecology 98*(4), 920–932.