

Gold Mining and Proto-Urbanization

Recent Evidence from Ghana

Marcel Fafchamps

Michael Koelle

Forhard Shilpi



WORLD BANK GROUP

Development Research Group

Agriculture and Rural Development Team

June 2015

Abstract

Central place theory predicts that agglomeration can arise from external shocks. This paper investigates whether gold mining is a catalyst for proto-urbanization in rural Ghana. Using cross-sectional data, the analysis finds that locations within 10 kilometers from gold mines have more night light and proportionally higher employment in industry and services and in the wage sector. Non-farm employment decreases at 20–30 kilometers distance to gold mines. These findings are consistent with agglomeration effects that induce non-

farm activities to coalesce in one particular location. This paper finds that, over time, an increase in gold production is associated with more wage employment and apprenticeship, and fewer people employed in private informal enterprises. It also finds that the changes arising from increasing gold production are not reversed when large gold mines shrink. However this pattern cannot be ascribed unambiguously to agglomeration effects, given an increase in informal mining after formal mines decrease output is also observed.

This paper is a product of the Agriculture and Rural Development Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at fshilpi@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Gold Mining and Proto–Urbanization: Recent Evidence from Ghana*

Marcel Fafchamps
Stanford University, BREAD and NBER

Michael Koelle
Oxford University

Forhad Shilpi
The World Bank

Keywords: Urbanization, natural resources

JEL Code: R12

*We would like to express our sincere thanks to Dr. Phylomena Nyarko of Ghana Statistical Services for making the census data (10 percent sample) available to us, and to Mrs. Jacqueline Anum and Harold Coulombe for their generous help with the data. We would also like to thank Ryan Engstrom and Richard Hinton from George Washington University, and Brian Blankespoor from World Bank for their assistance with various GIS data. This research was funded by the Knowledge for Change Program of the World Bank.

1 Introduction

Much debate remains around the question of where towns and cities locate. There is general agreement that the existence of agglomeration externalities is the reason why towns and cities exist. But the factors that predict the emergence of new towns and cities seldom predict their precise location, especially in the absence of strong geographical differentiation (e.g., large mountains, long rivers, presence of sea). In such cases, some historical accident is often presented as the reason why a particular town is where it is.

Many different types of historical accidents have been associated with the birth of a town – e.g., the site of a Roman fort, the intersection of two ancient roads, or the existence of a shrine or sacred place. The discovery of mining deposits or other extractive resources is known to have led to the creation of new towns, notably in the American West. But these towns do not always survive when mines or resources are exhausted or lose their market value, and many become ghost towns. There are important exceptions, such as Johannesburg, the South African megacity that grew out of gold mining. Other towns, such as Manaus, shrink once resource extraction declines (in the case of Manaus, natural rubber), but they do not disappear altogether.

In this paper, we look at whether the recent gold mining boom in Ghana contributes to the process of formation of urban areas. Given the short historical time frame, we are only able to observe initial phases of such a process – we call this ‘proto-urbanization’.¹ We use the opening and closure of formal gold mines and the rapid changes in mining output as a quasi-natural experiment to examine what mining does to the organization of economic activity across space and over time. Central place theory predicts a reconfiguration not only of the place that directly experiences an external shock – in our case, gold mining – but also of the surrounding hierarchical urban network. We identify proto-urbanization in the context of a developing country by measures of market activity and economic specialization, as well as consumption patterns.

There is a growing recent literature examining the effect of road and railway construction on

¹We use the word by analogy to ‘proto-star’. A proto-star is a star in the formation process. it begins with a cloud of matter that slowly coalesces onto a single point. At some point fusion starts and light is emitted in the middle of the cloud, although it may be difficult to see because the star is surrounded by dust and is still small. Gravity attracts further matter towards the star. By analogy, proto-urbanization is a process of agglomeration of economic activity which has begun, but which may not be clearly distinguishable yet for the observer.

the organization of economic activity across space. Donaldson (2014) shows how placement of railway in colonial India led to market integration and improved welfare significantly by allowing regions to exploit gains from trade. Trunk road and railway construction in China reinforced concentration of economic activity, and increased levels of GDP (Faber 2014; Banerjee, Duflo and Qian 2012) while local radial roads and railways along with ringroads led to some deconcentration (Baum-Snow et al 2014). Bird and Straub (2014) document a dual pattern in the organization of economic activity following the establishment of Brasilia as capital city, and its connection with trunk roads. Improved transport connections increased concentration of economic activity and population around the main centers in the South of the country, and spurred the emergence of secondary economic centers in the less developed North, in line with theoretical predictions regarding agglomeration effects. In the United States, Chanda and Thompson (2000) exploit the construction of interstates to document agglomeration effects in rural counties crossed by the highways. Michaels (2008) finds that interstates facilitate demand for high skilled labor where it is abundant.

There is less statistical work on the role of natural advantages and primary production in the rise of cities. Glaeser, Kerr and Kerr (forthcoming) provide evidence on how proximity to historical mining deposits led US cities to specialize in activities with significant economies of scale (e.g. larger firm sizes) at the cost of fewer start-ups. Jedwab (2014) documents how, in Ghana and Cote d'Ivoire, cocoa production spawned many small towns that survived the gradual westward shift of cocoa plantations that took place over the last century. Bleakley and Lin (2012) show how historic barriers along fluvial trade routes act as a coordination device for placement of new settlements, and how path dependence shapes the long-term evolution of these cities. Work on the socioeconomic impacts exhibits evidence of positive and negative effects. Demand spillovers of a large mine in Peru increase incomes in non-mining sectors in surrounding areas (Aragon and Rud 2013) while in Ghana, Gold mining reduces agricultural productivity by almost 40% via pollution, and negatively affects poverty, child nutrition and respiratory health (Aragon and Rud 2012). Focusing on female employment near mines in Sub-Saharan Africa, Kotsadamen and Tolonen (2013) find that women move from agriculture into the service sector when a new mine opens, but this reallocation does not survive mine closures. For a larger, global set of developing countries, von der Goltz and Barnwal (2014) find that mining communities

enjoy a substantial medium term rise in asset wealth but incur negative health consequences due to exposure to various pollutants.

We ignore informal gold mines that are responsible for much of the pollution discussed in the literature, and look at the short-term effects of a recent revival in mining. Gold has long been mined in Ghana, so much so that the colonial name for the country was the Gold Coast. Gold exports dropped significantly after independence as a result of an over-valued exchange rate, under-investment in primary production, and stagnating gold prices after the end of the Bretton Woods agreements. But in the last two decades, gold mining has experienced a revival in Ghana. The sector has been liberalized and new entry has taken place. As a result, in the recent past several new formal mines have opened while others have closed down following a depletion of their economically viable ore deposits. Using recent highly disaggregated location-specific data from a combination of sources, we investigate whether variation in gold mining between 2000 and 2010 is associated with tell-tale signs of proto-urbanization, such as an increase in consumption and a shift in employment composition towards non-farm production and wage employment. Ghana is a good country to study urbanization, given the rapid growth of both the population and the economy over the last two decades. Rapid urbanization is less likely to be observed in a country with a stagnant population or economy.

Using cross-sectional data, we find that locations in the immediate vicinity (up to 10 km) of gold mines have more night light and proportionally higher employment in industry and services and in the wage sector. Non-farm employment decreases at distances of 20-30 km from mining locations. These findings are consistent with agglomeration effects predicted by central place theory that induce non-farm activities to coalesce in one particular location. Exploiting variations in mining intensity over time and space, we find that an increase in gold production is associated with more wage employment and apprenticeship, and fewer people employed in private informal enterprises. We also find that the changes arising from increasing gold production are not reversed when large gold mines shrink. However this pattern cannot unambiguously be ascribed to agglomeration effects given that we also observe an increase in informal mining after formal mines decrease output. Although we do not have detailed information on their activities, we suspect that the informal miners flock to the sites of abandoned mines to seek gold residues in the mines themselves and in the tailings. Our results tend to suggest that this activity is

not particularly remunerative – incomes proxied by nightlights indeed seem to fall locally in spite of the population influx. This pattern is not entirely different from that documented by Jedwab (2014) for cocoa: the temporary prosperity brought by cocoa triggered the creation of small towns, but once cocoa production fell, these same towns attracted informal sector workers no longer employed in cocoa cultivation. What we observe is not identical but it shares one common feature, namely, that the *end* of primary production is associated with an influx of people and a rise in informal employment.

The rest of the paper is organized as follows. In Section 2 we briefly describe the conceptual framework. We present the gold mining sector in Ghana and various data sources used in the paper in Section 3, together with descriptive statistics of the variables of interest. Section 4 outlines the empirical strategy. and Section 5 presents estimation results.

2 Conceptual framework

The over-arching concept behind our analysis is central place theory: urban centers arise because certain economic activities benefit from agglomeration externalities. The sectoral mix of cities indirectly reveals, for a given point in space and time, where these externalities are strongest. For instance, in the 19th and early 20th century, most of the manufacturing activities were concentrated in cities. But since the 1950s, manufacturing moved out of urban centers to make room for finance and other services (e.g., Desmet and Fafchamps 2007). Agriculture, in contrast, requires a lot of land and thus tends to be displaced by urbanization – even though peri-urban agriculture often is the most productive because of its emphasis on high value crops and dairy production (Jacobs 1969). The urban and peri-urban share of agricultural employment, however, is low because of the importance of non-farm production (e.g., Fafchamps and Shilpi 2003). Based on this, we expect urban centers to be characterized by less employment in agriculture and more in non-primary sectors. Mining, by definition, follows sub-soil mineral deposits, the location of which often bears no relationship with the location of cities. Whether or not mineral deposits are exploited does, however, depend on accessibility and profitability. Formal mines create a significant amount of industrial employment, typically mostly for well-educated and highly specialized workers that migrate to the mining areas. However, miners may demand

local, non-tradable services and perhaps source some local production inputs.

For countries at low levels of economic development, agglomeration externalities arise in part because spatial concentration of demand makes job specialization possible (Fafchamps and Shilpi 2005, Fafchamps 2012). Many goods and services that are provided outside of the market realm in rural areas fall within the market domain in cities (Fafchamps 2011). As a result, many individuals are able to avail non-farm jobs as primary occupation, a fact that is reflected in occupational data. Because specialization allows the emergence of new and better goods and services, towns end up serving their surrounding rural hinterland in non-agricultural products and services – including trade and transformation of agricultural products.

Specialization, market provision, and concentration of demand also mean that firm size can increase to capture economies of scale. It follows that the share of people working for a wage (or running a firm with employees) is larger in towns and cities than in the surrounding countryside, where most people are self-employed or family workers, typically on a farm. Based on this, we expect urban centers to be characterized by a larger share of formal employment and a large proportion of employees or employers and a lower proportion of self-employed and family workers. There should also be more public sector employment because schools, health facilities, and government bureaucracies are often located in small urban centers from which they serve the surrounding countryside.

Since towns and cities have more people in specialized employment (either for a wage or on own account) as opposed to the range of self-provision activities that farmers undertake, we also expect them to have more people looking for employment, that is, a higher proportion of unemployed. In contrast, in a country like Ghana, people living on a farm are often underemployed, at least seasonally, but they rarely report themselves as formally unemployed. For similar reasons, in our dataset, women in rural Ghana often listed farming as their primary occupation while married women in urban areas are more likely to describe themselves as homemaker.²

We expect the number and size of cities to grow with economic development. The intuition is that higher levels of income shift demand away from basic necessities such as food and shelter, towards more specialized goods and services best provided in an urban environment. In Africa, urbanization seems to also have been fueled by population growth, with much urbanization

²For similar evidence from South Asia, see Fafchamps and Shilpi (2005).

occurring even in periods with little economic growth. Based on this, we expect new towns to arise in Africa, particularly in countries with rapid economic growth.

Viewed in a dynamic spatial context, central place theory predicts that towns should arise at more or less regular intervals that follow a honeycomb shape, and that they should structure into a hierarchy of small and large towns (e.g., Christaller 1933, Isard 1956).³ According to this theory, towns surrounded by a rural hinterland arise as a part of a self-emerging relational structure. The position of each town relative to other matters, but not their absolute location. It follows that, on a relatively undifferentiated terrain, the precise location of newly emerging towns is not known beforehand, as there are typically several equally suitable potential sites. Given this, the location of new towns can be influenced by historical accident – such as a temporary gold mine. This observation forms the basis of our investigation.

Our key identifying assumption is that new gold mines locate where there is gold and are not attracted by existing towns and cities. Entry in gold mining takes place when gold prices are high and the domestic business climate is sufficiently attractive, and exit takes place when deposits are exhausted or gold prices fall. Gold prices are determined internationally and the location of new mines is, if anything, negatively correlated with urban centers: gold mining is a source of pollution and new mines do not tend to locate too close to already established towns and cities. We can furthermore control for time-invariant local attributes by creating a panel of localities, and we also control to a certain degree for differential trends

Central place theory predicts the emergence of new towns in countries with a rapid population and economic growth, but not their precise location. It follows that we may be able to observe whether gold mining affects the precise placement of new embryonic towns. We call this process proto-urbanization.

To locate embryonic towns, we do not rely on population density because it can be affected by the size of gold deposits and by local agricultural productivity conditions – and thus may not indicate urbanization in the central place theory sense. We rely instead on their employment structure, i.e., we identify proto-urbanization by looking for tell-tale signs of urban-like activity, along the lines of the predictions outlined above.

³The hierarchical structure of towns and cities has been used to account for the fact that the distribution of town sizes follows an exponential distribution – often called Zipf Law (e.g., Henderson 1985). The evidence on towns following a honeycomb shape is less clear, but towns are typically not located right next to each other.

Finding that gold mining leads to the emergence of urban-like activity does not, by itself, guarantee that the embryonic town will survive once the gold deposits are exhausted. There are indeed many examples of towns that were abandoned in this manner – so called ghost towns. The critical test of whether proto-urbanization is taking place due to gold mining is whether the embryonic town survives – and thrives – after gold mining stops. Only then can we conclude that gold mining can be a catalyst for town formation.

3 Data

To test our theoretical predictions, we select a country with rapid population and economic growth, Ghana, and a time-period during which international gold prices increased and much entry took place in gold mining, i.e., the period between 2000 and 2010. We start by briefly describing the gold mining sector in Ghana. We then detail the data sources we use to study proto-urbanization.

3.1 Gold Mining in Ghana

During colonial times, Ghana was known as the Gold Coast. Gold mining dates back a long time in the pre-colonial era. The Ashanti kingdom, ruled from its historic capital Kumasi, controlled mining activities in many parts of the kingdom. At the end of the 19th century, French and British entrepreneurs obtained concessions and introduced modern, imported machinery to the industry (Taylor 2006). Mining was largely nationalized after independence until liberalization in the early 1990s. In the last two decades, the Ghanaian government has gradually reduced its stake in existing mining companies, and has awarded concessions to new prospective entrants, mostly foreign companies.

Gold mining is heavily concentrated in the southwest of the country, mostly between the cities of Sekondi/Takoradi on the coast, and Kumasi in the interior of the country. The maps in Figure 1 and Figure 2 show the spatial distribution of mines and cities in 2000 and 2010, respectively. The location of mines is determined by the geology. The vast majority of open pit and underground mines lie along the Ashanti Gold Belt, a geological formation that stretches from Sekondi/Takoradi to the East of Kumasi (Carranza *et al.* 2009). Other mines are located

in the Sefwi belt, parallel to the Ashanti belt (Pigois *et al.* 2003). Alluvial mining takes places along the rivers that run through this mineral-rich area: the Pra river and its tributary, and the Birim river. We focus on formal mines, for which production data are available.

We obtain mine-level production data from annual country reports published by the U.S. Geological Survey (USGS) for the years 1991-2003, and from corporate annual reports for the years 2004-2010. We cross-validate USGS and company data for the years preceding 2004. We exclude gold mined by artisans and traded by the Precious Minerals Marketing Corporation, which cannot be traced geographically. Our measure of mining activity covers all formal enterprises, which are medium and large-scale mines. We code mining location using a variety of sources, including annual reports, other corporate publications, technical industry reports, and high-resolution satellite imagery.

Some 18 formal gold mines operated in Ghana at some point in time between 1991 and 2010. Of these, 16 were operating in the decade preceding the year 2000; 12 of them were still operating in 1999. In the 2000s, there were 12 operating mines, of which 8 were still active in the year 2010. Table 1 lists the average yearly production volume by decade for each mine. Some alluvial placer operations in the Eastern region are quite small, producing on average 26 kg of gold per year. Mid-size mines have an average yearly production of about 2,000-5,000 kg. The largest mines in our study produce between 8,000 and 20,000 kg of gold per year. In the 1990s, the mine at Obuasi produced on average as much gold as all the other mines together. The production volume of Obuasi subsequently halved, and overall mine production decreased until the opening of new mines. In more recent years, a high level of production comes out of a small number of mines, each of which is fairly large. The largest is the Tarkwa mine, with about a quarter of overall production in 2010. The latest mine is the Ahafo mine, which opened only in 2008 and immediately became the second-largest mine. Figure 1 shows details on the evolution of output for the 10 largest mines over our study period. Overall production increased over time, as a function of enlargement of existing and opening of new mines. However, there is substantial variation in the evolution of individual mines. This is the variation we exploit.

Gold mining provides for a very small share of employment in Ghana as a whole. The sector as a whole – including formal and informal, large and small mines– employed only 0.72 percent of the country’s labor force in 2000, and 0.93 percent in 2010. But in the immediate vicinity

of mines, gold mining accounts for a substantial share of local employment. For example, 15 percent of all employed males who live within 10 km of a large-scale mine worked in gold mining in 2000. This figure rises to almost 20 percent in 2010.

In Figure 2, we present a map of Ghana in which the shaded area shows all the areas located within 100 km of any mine in the year 2000. Figure 3 does the same thing for the year 2010. The two maps show that urbanization in a large proportion of the country could potentially be affected by proximity to gold mining. We see that the shaded area includes all the regions West of the mines up to Ivorian border and it stretches to the North past Kumasi. In 2000, it spread past Accra and Lake Volta to the East due to the presence of two small alluvial mines there. Because these mines were discontinued in 2010, there is less shaded area in the Eastern part of the country in Figure 3.

3.2 Data sources

In addition to the mining data mentioned in the previous section, we make use of two data sources to construct proxies for proto-urbanization: the 2000 and 2010 population censuses; and night-time illumination data from satellite imagery. We discuss them in turn.

The census data were made available to us at the enumeration area (EA) level. There were 26,708 EAs in the 2000 Ghana census and 37,625 EAs in the 2010 census. In addition, we have a complete map of the EAs from the 2000 census. In this map, EAs in urban localities are grouped together into a single polygon, a town. This affects 64 % of the EAs. We extract information about the GPS coordinates of the center of each of these EAs and towns. In total, we obtain 12,472 such coordinates. For the enumeration of the 2010 census, some of the EAs were split up, but none were merged. This allows us to link them to our map. We are able to link 11,698 of EAs across the two censuses.⁴ They cover 96.17 % of all households in 2000, and 98.51 % in 2010.

Within each EA we have data on 10% of all households. In addition to gender, marital status, and relationship to the household head, the two censuses contain, for each adult, a broad characterization of their sector of activity (e.g., mining, agriculture, trade), labor market

⁴In the following, we refer to the set of polygons that we have mapped as ‘EAs’, acknowledging that in some instances, these polygons group multiple enumeration areas as defined for Census purposes.

participation (e.g. employed, unemployed, homemaker), form of employment, if any (e.g., self-employed, wage worker, apprentice), and type of employer (e.g. private formal, private informal, public). Information on occupation and sector of activity is the focus of our empirical analysis.

The night-time illumination data come from satellite imagery. They were obtained from the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS). The dataset covers most of the earth at a 30 arc-second grid (approximately 1 square kilometer). The luminosity measure for each cell, or pixel, is a digital number in the range 0 to 63 that increases with luminosity. There are 280,000 pixels for Ghana. Recent studies have made use of these data as a high-resolution, local measure of economic development (Bleakley and Lin 2012; Henderson, Storeygard, and Weil 2012; Michalopoulos and Papaioannou 2013). We map the nightlight data into our EAs, and are able to construct a panel dataset for 12,390 EAs and 19 years (1992-2010).⁵ For most years, we have observations from 2 different satellite readings. To reduce measurement error and noise, we use the average of the two as nightlight measure for that year. We also use year dummies to correct for differences in satellite calibration over time.

3.3 Descriptive statistics

Administratively, Ghana consists of 10 regions which are subdivided into districts. There were 110 districts in 2000. We keep the 2000 district boundaries throughout the analysis. The average (median) district has a size of about 2,200 (1,350) square km. According to the census data, the population of Ghana was 18.85 million in 2000 and 24.66 million in 2010, corresponding to an annual rate of population growth of 2.7% – close to the African average. Urbanization is equally rapid: in 2000, about 44 percent of the population lived in urban areas, and 51 percent in 2010. These two features make Ghana an ideal candidate to study town formation.

Since the main focus of our analysis is occupational data, we focus on the working-age population, which we define as all adults aged 20 to 60 years. These make up 42.5 percent of the total population in 2000, and 45.4 percent in 2010. The composition of the labor force, labor supply, employment status, and type of employer are summarized in Table 2.

In line with the rapid urbanization of the country, we observe a rapid transition of the

⁵We additionally construct a grid where each cell contains 3×3 pixels. This yields 31,334 grid cells for each of the years.

economy out of agriculture: in 2000, more than half of total employment was in agriculture but by 2010, the employment share of agriculture had fallen to 38 percent. The difference was absorbed in the tertiary sector, with trade (from 17 to 20 percent), services (from 10 to 16.5 percent) and food and accommodation (from 2.4 to 5.7 percent) all experiencing increasing employment shares. Unemployment decreased from 7.7 to 4.6 percent while the employment share of manufacturing remained stable. As documented in Fafchamps and Shilpi (2003, 2005), a higher share of employment in services is consistent with a more urbanized economy with higher levels of occupational specialization.

The bottom part of Table 2 documents the types of occupations and kinds of firms and organizations in which Ghanaians are employed. The largest occupational group is self-employed workers without employees, who accounted for 69 percent of employment in 2000 and 61 percent in 2010. Around 5 percent of the self-employed are employers, that is, they have employees. Wage employment increased between 2000 and 2010 from 17 to 21 percent. The rest are family workers, apprentices, and other kinds of workers. The large majority of employment is in informal private businesses, and the proportion is rising over time – from 83% in 2000 to 85% in 2010. Over the same period, employment in formal private sector jobs fell slightly from 8.8 to 7.5 percent. Employment in the public sector accounts for a stable 7 to 8 percent of total employment.

The distribution of the night light measure over several time periods is documented in Table 3. This measure is zero when a pixel is unlit at night and positive otherwise. Over our study period, about 93 percent of the country is unlit. This is in line with what is documented for poor countries with a medium population density (Henderson et al. 2012). The percentage unlit decreases from 97 percent in 1992 to 90 percent in 2010. Over time, we observe the entire night light distribution shifting upwards. This is in line with the general pattern of economic and urban growth in Ghana over this time period.

Nights light is unevenly distributed across space. Maps with night light in the mining region of southern Ghana are shown in Figure 4 for the years 1992, 2000, and 2010. In 1992, the first year for which data are available, the main cities of Accra and Kumasi, as well as a number of smaller cities, are very brightly lit. About the only lit places between Kumasi and the coast are all near mines. Over time, more and more places along on the coastline and along major highways

emit lights at night, documenting the spatio-temporal pattern of economic development and the growth of cities. Places near mines are consistently brightly lit.

4 Empirical specification

Our objective is to examine how employment and income patterns vary with proximity to a (formal) gold mine. Let y_{it} denote an outcome of interest measured at the level of locality i at time t . For example, y_{it} can be the proportion of male adults that are employed in agriculture in location i . Let vector $\{m_{jt}\}$ represent all known gold mines in Ghana, indexed by j . Variable m_{jt} is defined as the average amount of gold produced by mine j in a relevant time interval up to year t . In our analysis, this time interval is typically 10 years, given that most of our outcome variables y_{it} are only available in 2000 and 2010.

We are interested in estimating $E[y_{it}|\{m_{jt}\}]$, that is, in estimating how the average value of y_{it} systematically varies with proximity to various gold mines. To this effect, we wish to estimate a model of the form:

$$y_{it} = F(\{m_{jt}\}) + \text{controls} + u_{it} \quad (1)$$

where function $F(\cdot)$ is our object of interest. Controls are detailed later. We cluster standard errors at the district level throughout.

We think of the flexible function $F(\cdot)$ as aggregating the predictive effect of gold mines located at various distances from i . To capture this idea in a compact way, we follow Fafchamps and Shilpi (2005) and let:

$$F(\{m_{jt}\}) = \int_0^\infty \gamma(\delta) m_t(\delta) d\delta \quad (2)$$

where $\gamma(\delta)$ is a parameter that varies with distance δ and $m_t(\delta)$ is the set of mines in $\{m_{jt}\}$ that are located at distance δ from location i . In practice, we discretize $\gamma(\delta)$ into a fixed number of distance intervals or ‘bins’ indexed by k . To be more specific, $k = 1$ corresponds to $0 \leq \delta \leq 10$ km, $k = 2$ corresponds to $10 < \delta \leq 20$ km, etc. Since we do not expect employment patterns to be affected by gold mines more than 100 km away, we assume that beyond 100 km, $F(\cdot)$ is

zero. With this parametric assumption, function $F(\cdot)$ can be rewritten as:

$$F(\{m_{jt}\}) = \sum_{k=1}^K \gamma_k m_{ikt} \quad (3)$$

where m_{ikt} is the sum of m_{jt} located in distance interval k from location i . So, for instance, if mine m_{3t} is located 5 km from i , mines m_{7t} and m_{9t} are located 12 km from i , and all other mines are more than 100 km from i , we have:

$$F(\{m_{jt}\}) = \gamma_1 m_{3t} + \gamma_2(m_{7t} + m_{9t})$$

We estimate three specifications of model (1). The first specification is the simplest, as it omits controls:

$$y_{it} = F(\{m_{jt}\}) + u_{it} \quad (4)$$

In the second specification, we control for proximity to towns and cities. The estimated model is:

$$y_{it} = F(\{m_{jt}\}) + G(\{c_{jt}\}) + u_{it} \quad (5)$$

where function $G(\cdot)$ controls for how y_{it} varies with proximity to various urban centers j of more than 10,000 inhabitants, where the size of each city is proxied by its population c_{jt} . Conditioning on urban proximity identifies the pattern of correlation between employment pattern y_{it} and proximity to gold mines, net of any correlation between proximity to mines and proximity to urban centers. Function $G(\cdot)$ is constructed in a manner similar to $F(\cdot)$, i.e., we write:

$$G(\{c_{jt}\}) = \sum_{k=1}^N \lambda_k c_{ikt} \quad (6)$$

where c_{ikt} is the sum of city population c_{jt} located in distance interval k from location i and N is the number of urban centers of more than 10,000 inhabitants in Ghana.

The third model we estimate adds district fixed effects v_d to control for persistent geographical differences due to variation in natural endowment and historical legacy:

$$y_{it} = F(\{m_{jt}\}) + G(\{c_{jt}\}) + v_d + u_{it} \quad (7)$$

With model (7), we estimate the correlation pattern between y_{it} and proximity to gold mines that is not due to a correlation between gold mining and persistent geographical differences.

Model (7) includes district fixed effects but not fixed effects at the level of the location i itself. This may affect inference if the location of gold mines is correlated with location specific invariant characteristics that also affect urbanization. To allow for this possibility, we estimate a dynamic version of regression models (5) and (7):

$$\begin{aligned}\Delta y_{it} &= \Delta F(\{m_{jt}\}) + \Delta G(\{c_{jt}\}) + \Delta u_{it} \\ &= \sum_{k=1}^K \gamma_k \Delta m_{ikt} + \sum_{k=1}^N \lambda_k \Delta c_{ikt} + \Delta u_{it}\end{aligned}\tag{8}$$

In model (8) identification of the γ_k 's and λ_k 's relies solely on variation between 2000 and 2010. Gold mining output m_{ikt} and city population c_{ikt} do vary between 2000 and 2010 so that the coefficients are point identified. But the amount of variation is limited so that what we gain in terms of potential bias, we may lose in terms of efficiency. We estimate model (8) with and without controls for changing city population – controlling for changes in population size may partial out one of the potential mechanisms through which gold mining affects urbanization.

Model (8) becomes truly useful once we use it to investigate whether, over the study period, gold mines served as catalyst for urban formation. If gold mining serves as catalyst, we would expect the urbanization effect of gold mines *not* to disappear once gold mines close or shrink. Urbanization forces may even grow stronger once the pollution from gold mining (e.g., mercury, tailings) is eliminated: once mines leave, the fledging town may expand further. The alternative is that, once the mine disappears, it leaves a ghost town behind.

To investigate this possibility, we estimate a model of the form:

$$\Delta y_{it} = \sum_{k=1}^K \gamma_k^+ \Delta^+ m_{ikt} + \sum_{k=1}^K \gamma_k^- \Delta^- m_{ikt} + \sum_{k=1}^N \lambda_k \Delta c_{ikt} + \Delta u_{it}\tag{9}$$

where

$$\begin{aligned}\Delta^+ m_{ikt} &\equiv \max(0, \Delta m_{ikt}) \geq 0 \\ \Delta^- m_{ikt} &\equiv \min(0, \Delta m_{ikt}) \leq 0\end{aligned}$$

Put plainly, we have split Δm_{ikt} into positive and negative changes $\Delta^+ m_{ikt}$ and $\Delta^- m_{ikt}$. If the creation or growth of a formal gold mine in location i triggers proto-urbanization proxied by an increase in y_{it} , we expect $\gamma_k^+ > 0$. If proto-urbanization is reversed once a formal gold mine closes, we expect $\gamma_k^- > 0$ as well – remember that $\Delta^- m_{ikt} < 0$ in places where a mine closes or shrinks. If we find instead that $\gamma_k^- \leq 0$ this suggests that proto-urbanization does not stop once the mine shrinks or disappears – and even increases if $\gamma_k^- < 0$.

5 Empirical analysis

To recall, we wish to test whether gold mining fosters proto-urbanization. To do this, we examine whether the presence of a gold mine is associated with tell-tale signs of proto-urbanization. Our identifying assumption is that the location of a gold mine is exogenous to the selection of new urban centers. This assumption is reasonable given the Ghanaian context and the relatively short time frame of our study. Essential to the success of this approach is the fact that our data are available at a high level of spatial disaggregation.

What are the tell-tale signs of proto-urbanization that we can identify in our data? We can find them by looking at the sectoral composition of the economy (primary vs. nonprimary); at measures of specialization of workers in occupations; and at measures of local incomes.

First, urban activity by definition focuses on non-primary production, that is, on production other than agriculture or mining. If gold mines serve as catalyst for urbanization, we would observe that gold mines are associated with a rise in the share of non-primary employment in total employment. Second, urbanization is also associated with agglomeration effects that foster economies of scale in production and that make specialization viable. Given this, we expect to observe more employment in the (non-primary) wage sector, either as worker or as employer. For similar reasons, we expect to observe more people employed in the formal private sector. Wage employment in the public sector may similarly rise if government administration and services locate near gold mines, as this would also serve as catalyst for urbanization. Since apprentices are only observed in non-primary sectors, a rise in the share of apprentices would signal that a larger proportion of workers wish to learn skills that enable them to obtain employment in the non-primary sector.

Furthermore, agglomeration leads to the concentration of jobs in a given location. This in turn makes it more appealing for people to move to an agglomeration in search of work. For this reason, we expect to find more self-declared ‘unemployed’ individuals near gold mines. If these individuals are looking for work primarily outside mining, we would also regard this as consistent with proto-urbanization. Another pattern we expect to find is the rise in the proportion of people (mostly women) who list themselves as homemakers. If they were located in a rural village where agriculture dominates activity, these homemakers would be spending part of their time in the fields, and would typically be counted as employed in agriculture. Finally, if gold mines trigger economies of agglomeration and returns from specialization, we also expect incomes to be higher – a feature that would manifest itself as an increase in night light.

5.1 Census data

We start with the cross-sectional analysis of the census data. We report in Figures 5 and 6 the regression estimates of γ_k from model (4) estimated using the 2010 data. The two figures correspond to a different set of dependent variables from the census data. Coefficient estimates γ_k are reported graphically, together with their 95% confidence interval based on standard errors clustered at the district level. As explained earlier, we focus our attention on distances up to 100 km from a gold mine. Since distance is divided into intervals of 10 km each (see equation 3), we report ten γ_k coefficients for each regression. Coefficients represent the effect on local employment shares (in percentage points) of a one ton increase in gold production at a certain distance interval, relative to localities further away than 100 km. Mining variables represent average annual mining output in the preceding decade. Table 4 provides an overview of the distribution of our regressor of interest m_{ikt} , tons of gold produced at distance interval j from location i . By construction, the number of locations that have positive m_{ikt} increases with distance from the mine. About 60 percent of all locations in 2010 are further than 100 km away from any active gold mine; these locations constitute the omitted category in our regressions.⁶ Similarly, we express city population c_{ikt} in thousands of inhabitants of a city. Our definition of city includes all EAs that are classified as urban by the Ghana Statistical Service, and that

⁶In 2000 about 45 percent of all locations are further away than 100 km from any active mine.

have at least 10,000 inhabitants.⁷

In Figure 5 we show the proportion of adults employed in different sectors of activity. We observe that the presence of a formal gold mine in location i is associated with a significantly smaller share of agricultural employment and a higher share of employment in all other sectors, including of course gold mining. Localities more than 50 km away from a gold mine are not different from other localities. This suggests that truncating $F(\cdot)$ above 100 km is sufficiently conservative. We also note that individuals living around 30 km away from a gold mine seem to be, if anything, less likely to be employed in a non-primary sector than individuals living farther away from a mine. If confirmed by subsequent analysis, this would indicate that these occupations ‘agglomerate’ in the immediate vicinity of the gold mine – and is thus indicative of proto-urbanization.

In the top row of Figure 6 we report γ_k coefficient estimates for employment in the wage sector, employment in the non-wage sector, and the proportion of apprentices. Employment in the mining sector is omitted from the calculation.⁸ Employment in the wage sector combines salaried workers and self-employed individuals with employees. Employment in the non-wage sector combines the self-employed and family workers. Apprentices constitute a much smaller category but are of interest in their own right because they signal a desire to acquire skills outside the primary sector. We find that localities in the immediate vicinity of a gold mine have more employment in the wage sector and more apprentices. This too is consistent with proto-urbanization.

The middle row of Figure 6 splits employment into informal, private formal, and public formal.⁹ Consistent with the results in the top of the figure, we find that proximity to a gold mine is associated with significantly more formal and less informal employment in the private sector. Public sector employment is slightly higher in the vicinity of mines, suggesting that government services agglomerate there. If true, this would be another tell-tale sign of proto-urbanization. But the effect is not strong. The bottom of Figure 6 shows results for the unemployment rate and the proportion of homemakers in the adult population. We see that

⁷We try different population cutoffs, and they yield the same results.

⁸More precisely, the employment share of the wage sector is the proportion of people employed in the wage sector among all those employed excluding those employed in the mining sector (gold or other).

⁹The public formal category includes workers in the parastatal sector.

they are both higher in the vicinity of mines. As indicated earlier, this too can be interpreted as indicative of proto-urbanization.

To verify the robustness of our findings, we start by estimating Figures 5 and 6 with the 2000 data. Results, shown in the online appendix, are similar but less pronounced. Next we control for proximity to urban centers. A look at the maps in Figures 2 and 3 indeed suggests that distance to mines may be correlated with distance to large cities. There are two major mining areas: the large mine near Obuasi, accounting for more than 50 percent of all mining output in the 1990s; and the complex of mines near Bogoso-Prestea-Tarkwa. There is a town close to each of them. Right next to the Obuasi mine is the town of Obuasi proper, Ghana's eighth largest city, and Obuasi itself is located about 50-60km from Kumasi, Ghana's second largest city. There is no large town in the immediate vicinity of the Bogoso-Prestea-Tarkwa mining complex, but it is located 50-60km from Sekondi-Takoradi, Ghana's third largest city. Since localities where much of Ghanaian gold production originates happen to be located 50-60 km from a major urban center, failing to control for urban proximity may affect our estimates.

We reestimate Figures 5 and 6 with the 2000 and 2010 data using the regression model (5). It is not clear whether city population in 2000 or in 2010 is the appropriate control for pre-existing cities.¹⁰ Therefore we also estimate a mixed version of model (5) where we control for lagged city population together with 2010 mining and outcome data. Results are again very similar to those in Figures 5 and 6. They are available in the online appendix. Finally, we also include district dummies in the regression in an effort to control for geographical, institutional, and other features that are common to a district and may be correlated both with proximity to gold mines and drive urbanization. However, in our spatial analysis, districts might also pick up part of the effect of mining otherwise captured in γ_k , because locations close to mines fall within only a few districts (see Figures 1 and 2).¹¹ Comparing locations located at different distances to mines within the same district, a small geographical area, is therefore an especially conservative test of the effects of mining. Results are summarized in Figures 7 and 8 for the year 2010. For locations in the immediate vicinity of gold mines, the patterns are qualitatively similar to those

¹⁰Our mining variable is measured as a flow over the decade up to time t but the outcome variables are stocks at t .

¹¹Imagine a district that spans across locations situated 10-30 km from a mine. The dummy coefficient for this district will then pick up the average effect of mining at distances 10-30 km, while γ_2 and γ_3 will pick up deviations from this average for locations at distances 10-20 km and 20-30km, respectively.

discussed so far, but generally smaller in magnitude and less statistically significant. We do, however, find that individuals living around 30 km from a gold mine are significantly more likely to work in agriculture and less likely to be employed in the non-primary sector. They are also less likely to be listed as unemployed or homemaker, and less likely to work in the public sector.

In summary, the tell-tale signs that suggest agglomeration effects facilitated by gold mines are, on the one hand, differences in the structure of the economy at the mining sites. A smaller share of individuals in these locations are engaged in agriculture and other nonwage labor, and more dedicate the majority of their time to providing services to the market; including as wage laborers in larger organisations and increasingly so in the formal private or the public sector. Similarly, more individuals learn a trade, specialise in home production, or are dedicated to looking for work opportunities on the market. On the other hand, the opposite effects can be detected (albeit with less precision and strength) at 20-30 km from mines. This suggests that non-rural activities coalesce where mines are.

5.1.1 Dynamic analysis

Next we turn to a dynamic analysis of the census data. We first estimate model (8). Results, shown in Figures 9 and 10, are surprising: many estimates of γ_k are not significant in the immediate vicinity of a mine, but when they are, they tend to have a sign opposite to that documented in Figures 5 to 8.

The explanation for these seemingly puzzling results comes to light when we estimate model (9) and interact distance to a gold mine with the sign of the change in mining output between the 1990s and 2000s. Estimates for γ_k^+ are shown in Figures 11 and 12 while estimates for γ_k^- are shown in Figures 13 and 14. The γ_k^+ reported in Figures 11 and 12 represent the changes in economic activity associated with an expansion in formal gold mining, compared to locations with no change in gold production (typically, locations without gold mines). We see that the γ_k^+ estimates are qualitatively similar to those found in the cross-section analysis for: non-gold mining; food and accommodation; wage, non-wage, and apprenticeship sector; and private informal enterprise. Overall, this confirms that when we compare locations with and without a gold mine, the former display some tell-tale signs of proto-urbanization. There are some exceptions, however: coefficient estimates are reversed for trade and transport activities,

suggesting a crowding out of some tradable sectors. They are also reversed for unemployment and homemaking, which is probably due to the fact that gold mines expand production, they hire more workers. Other coefficients are not significant, such as agriculture. This suggests that the proportion of workers in agriculture in a given location is not affected much by temporary changes in gold production.

Estimates of the γ_k^- coefficients add a completely new dimension to the above findings. They represent the effects on outcomes of a decrease in gold mining activity. If urbanization is reversed once the gold mine disappears, we should find that γ_k^- coefficients have the same sign as γ_k^+ and as those found in Figures 5 to 8. Instead we find that, when the γ_k^- coefficient estimates are significant, they have the opposite sign to those found earlier. For instance, the γ_k^- coefficient for the share of agricultural employment in total employment – shown in the upper-left corner of Figure 13 – is significantly positive at distances 0 to 10 km. This means that locations that experienced a more dramatically shrinking gold production between 2000 and 2010 have a *smaller* share of agricultural production in 2010 than in 2000. If anything, this finding suggests that proto-urbanization continues even after the gold mine shrinks or closes.

Perhaps the most striking finding is that employment in gold mining increases once a formal mine shrinks or closes. To investigate this issue further, we split gold mining employment into formal and informal, wage and non-wage employment. Results, reported in the online appendix, indicate that a reduction in mining output is associated with an increase in informal wage and non-wage (self-) employment in gold mining. This is suggestive of a substitution by informal and small-scale mining activities once a large formal mine closes. This is likely to be a temporary phenomenon, driven by attempts to extract small amounts of gold from tailings and closed mines. This phenomenon does not, by itself, constitute evidence of a durable urbanization legacy.

5.2 Robustness

Our key identifying assumptions are that the location of gold mines is exogenous to the existence and the evolution of urban economic activity. We therefore report specifications controlling for pre-existing urban centres with the main results. The dynamic analysis constitutes a first and most important robustness check: first differencing removes any time-invariant attributes of localities that could be correlated with both mining and urbanization.

We have time-invariant measures of such attributes and we can control for them in the cross-sectional regressions. Gold deposits lie underneath chains of gentle hills in an otherwise undifferentiated landscape of South-Western Ghana. This could co-determine suitability for agriculture. Furthermore, access to the road network influences both the placement of mines, and the potential for trade and specialization. We control for slope and elevation of the terrain, and distance to the nearest primary and secondary roads of a location. This does not affect the results.¹²

Although we exploit substantial variation in mine openings and closures, we do not observe all mines from the start. Locations with more gold production in 2010 could be different from others because of a history of mining in their vicinity. In an attempt to separate the legacy effects from contemporaneous effects of mining, we reestimate Figures 5 and 6 with 2010 data and control for 2000 mining output. The γ_k coefficients that are attributable to variations in gold production over and above historical production give if anything a stronger indication of agglomeration effects.

The dynamic specifications control for time-invariant unobservable characteristics, but they rely on the assumptions of common trends in labor markets, industry composition, and other measures of agglomeration we use. We can control for differential trends given a number of observable characteristics that may mediate the impact of Ghana’s economic growth and structural transformation across locations: initial city population, terrain, and distance to roads. We find little evidence for differential trends in the first place, and our main results are unchanged.

We perform several other robustness checks. First we show that our results do not change if we let mines affect outcomes more than 100 km away. Secondly, we repeat all the analysis replacing m_{ijt} with a dummy equal to 1 if gold mine j is in operation at time t . We get similar results. Third, we demonstrate robustness to using the average of the last 5 years instead of 10 years as an alternative measure of gold production. Finally, we reestimate all the regressions weighting the data by population instead of using EA means. Results are similar.

¹²This and all of the following robustness checks are relegated to the online appendix.

5.3 Nightlight

We conduct a similar analysis on the nightlight data. As in other analyses of this kind, we take nightlight as a proxy for income. We report results from three regression specifications: (1) a cross-section specification; (2) a first-difference specification, which corresponds to model (8); and (3) a first difference specification (9) in which we allow γ_k to be different for contracting and expanding gold production. All regressions include year dummies. They control for annual differences in sensor sensitivity, as well as for overall economic and infrastructure development. The mining data, as before, comes from formal gold mines. We assign to each EA the average measure of nightlight across its area. Standard errors are spatially clustered at the district level, as before.

Coefficient estimates γ_k are again reported graphically, together with their 95% confidence interval. In Figure 15 we limit our analysis to the years 2000 and 2010. This is done so as to obtain results as comparable as possible to those reported so far. As before, mining variables represent average mining output in the preceding decade. From the upper-left corner graph, we see that, across Ghana, locations with more mining activity have more nightlight, but only in the immediate proximity ($< 10\text{km}$) of a mine. From the first-difference regressions we see that an increase in local mining activity is also associated with an increase in nightlight. The opposite is true when local mining activity decreases: we find that nightlight falls.

We also carry out a yearly analysis, using a long panel over 19 years on the EA level on annual nightlights and local gold mining output. Since we observe data at a relatively high frequency, we are concerned with serial correlation in local shocks that might influence both mining and nightlights. A prime example would be annual local fluctuations in electricity supply.

¹³ Furthermore, nightlights themselves might be serially correlated. To address these issues, we estimate a dynamic panel version of our basic model:

$$y_{it} = \alpha y_{i,t-1} + \sum_{k=1}^K \gamma_k m_{ikt} + \eta_i + u_{it} \quad (10)$$

We again take the first difference to purge out any possible correlation of time-invariant local

¹³Ghana obtains most of its electricity from hydro power plants at the Akosombo dam of Lake Volta. Electricity supply therefore partly depends on rainfall conditions. Shortages of electricity, and power cuts due to rationing, are common even in the largest cities.

characteristics η_i with both nightlights y_{it} , and gold production m_{ikt} to obtain the model

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \sum_{k=1}^K \gamma_k \Delta m_{ikt} + \Delta u_{it} \quad (11)$$

The presence of serial correlation ($\alpha \neq 0$) mechanically introduces correlation between the lagged difference $\Delta y_{i,t-1}$ and the error term Δu_{it} . A well-known remedy to this is to instrument the first-differenced equation (11) with past levels of the outcome variables. Furthermore, gold production might be endogenous to current local shocks. With sufficient serial correlation in gold production, we can also instrument current production with its past values. We report estimates of the models (10) and (11) in Table 5. In columns (1) to (3) we present OLS, fixed effects, and first-difference estimates of the static model (when $\alpha = 0$). We only find a systematic association of nightlights with mining output in the immediate proximity of mines. Reassuringly, the point estimates of this effect in columns (1) and (3) are almost identical to those reported in the top row of Figure 15. This suggests that how we periodize does not affect our estimates. In column (4) we report a basic AR(1) specification of nightlights without any covariates. The autocorrelation coefficient is close to a random walk; but it might be biased upward if there is serial correlation in the unobservables. We therefore experiment with different lags of nightlight as instruments, and find that a lag of 4 years seems to be sufficient to reject serial correlation in the Δu_{it} .¹⁴ In column (6) we add our mining covariates, and we instrument for the difference in mine production Δm_{ikt} with the second lag $\Delta m_{ik,t-2}$.¹⁵ We find that nightlights respond to mines at locations in immediate proximity of mines. We also find a statistically significant and positive effect at 40-50 km distance.

The bottom line is that the nightlight data partly confirm and partly contradict our earlier findings. At the cross-section level, gold mining is associated with more economic activity in the immediate vicinity of a mine, but these effects do not spill over to adjacent localities. A similar finding arises when we look at localities with expanding gold output, and this finding is robust to dynamic panel instrumentation. But, contrary to Figures 13 and 14, we do not find

¹⁴The appropriate test is to test for autocorrelation of second order in the differenced error term, because differencing introduces mechanical autocorrelation of first order (Arellano and Bond 1991). We cannot reject second order autocorrelation for lags smaller than order four.

¹⁵Since the evolution of gold mines is non-monotonic (see Figure 1) the lagged difference of gold production is a much more informative instruments than the lagged level. If gold mining is endogenous to contemporaneous shocks, then the second lag provides the first valid instrument.

that incomes continue to rise after the mine shrinks or disappears. This may be because some of the night light that we observe near mines is due to gold mining itself, thereby throwing doubt on the relevance of night light as a measure of incomes in our case. Similarly, we do not find an effect on nightlights at 20-30km from mines, as we did with the census data. This might partly be due to bottom-coding of the nightlights measures, which applies to most rural areas.

6 Conclusion

Central place theory indeed predicts that, when conditions are suitable for the creation of new urban centers, concentration of population driven by external shocks can determine where the new urban centers locate. We are interested in testing the idea that a permanent urban settlement can arise as a result of a shock that temporarily generates a concentration of economic activity in an arbitrary location. The external shock we examine here is gold mining, which tends to locate where minerals are found and is largely unaffected by other geographical features that make a site suitable for urban location. This is particularly true in Ghana where the geography is fairly undifferentiated with few mountains and navigable rivers.

We investigate whether gold mines in Ghana are associated with tell-tale signs of proto-urbanization. Although gold has been extracted from Ghanaian soil for centuries, the last two decades have seen a large increase in production after the sector was liberalized and foreign investors entered gold mining in Ghana. The Ghanaian economy has been growing steadily, and the population is expanding and urbanizing rapidly. These are ideal conditions to study town formation.

Using cross-sectional data we find that locations with gold mines have a larger proportion of people employed in industry and services, and a larger proportion of people employed in the wage sector or looking for work. The presence of a gold mine is also associated with more night light, which suggests higher incomes and a more urban setting with electrification. We also find evidence suggesting that the agglomeration of non-farm employment around gold mines is accompanied by a decrease of non-farm employment at distances of 20-30 km from gold mines. These findings are consistent with agglomeration effects that induce non-farm activities to coalesce in one particular location.

Slightly different results are found when we seek to identify the effect of gold mining by comparing locations over time. We find that an increase in the output of formal gold mines is associated with a higher proportion of employment in wage employment and apprenticeship, and a decrease of the proportion of people employed in private informal enterprises. These findings are broadly consistent with the cross-sectional analysis. We also find that the changes associated with an expansion in formal gold production are not reversed when formal gold mines shrink or close. The share of agricultural employment continues to shrink after formal mines decrease gold production, while employment in informal gold mining increases. This is likely to be a temporary phenomenon associated with the extraction of gold from tailings and closed mines using less technology-intensive methods. This phenomenon does not, by itself, constitute evidence of a durable urbanization legacy. Although the older and larger mines in our data are associated with large urban settlements, more time is needed to ascertain whether gold mining in Ghana can, like a grain of sand in a pearl, trigger the formation of self-sustaining agglomerations.

Bibliography

Fernando M. Aragon and Juan Pablo Rud (2013). "Natural Resources and Local Communities: Evidence from a Peruvian Gold Mine," *American Economic Journal: Economic Policy*, 5(2): 1-25, May

Fernando Aragon & Juan Pablo Rud (2012). "Mining, Pollution and Agricultural Productivity: Evidence from Ghana," Discussion Papers dp12-08, Department of Economics, Simon Fraser University

Manuel Arellano and Stephen Bond (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58(2): 277-297.

Abhijit Banerjee, Esther Duflo and Nancy Qian (2012). "On the Road: Access to Transportation Infrastructure and Economic Growth in China", MIT (mimeograph)

Julia Bird and Stéphanie Straub (2014). "The Brasília Experiment: Road Access and the Spatial Pattern of Long-term Local Development in Brazil," TSE Working Papers 14-495, Toulouse School of Economics

Nathaniel Baum-Snow, Loren Brandt, J. Vernon Henderson, Matthew A. Turner and Qinghua Zhang (2014), "Roads, Railroads and Decentralization of Chinese Cities", mimeo

Hoyt Bleakley and Jeffrey Lin (2012). "Portage and Path Dependence," *The Quarterly Journal of Economics*, 127(2):587-644

Carranza, E.J.M, Owusu, E.A., and Hale, M. (2009). "Mapping of prospectivity and estimation of number of undiscovered prospects for lode Gold, southwestern Ashanti Belt, Ghana". *Miner Deposita*, 44: 915-938

Amitabh Chandra and Eric Thompson (2000). "Does public infrastructure affect economic activity? Evidence from the rural interstate highway system. ", *Regional Science and Urban Economics*, 30(4): 457-490.

Walter Christaller (1933). *Die zentralen Orte in Süddeutschland* (The Central Places in Southern Germany), Gustav Fischer, Jena.

David Donaldson (2014). "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure", *American Economic Review*, (forthcoming)

Klaus Desmet and Marcel Fafchamps (2006). "Employment Concentration across US Counties", *Regional Science and Urban Economics*, 36(4): 482-509, July

Benjamin Faber (2014). "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System", *Review of Economic Studies* (forthcoming)

Marcel Fafchamps and Forhad Shilpi (2003). "The Spatial Division of Labor in Nepal", *Journal of Development Studies*, 39(6): 23-66

Marcel Fafchamps and Forhad Shilpi (2005). "Cities and Specialization: Evidence from South Asia", *Economic Journal*, 115(503): 477-504, April

Marcel Fafchamps (2011). "Development, Social Norms, and Assignment to Task", *Proceedings of the National Academy of Sciences (PNAS)*, 108(Supplement 4): 21308-15, December

Marcel Fafchamps (2012). "Reprint of Development, Agglomeration, and the Organization of Work", *Regional Science and Urban Economics*, 42(5): 765-78

Glaeser, Edward L., Sari Pekkala Kerr, and William R. Kerr. "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines." *Review of Economics and Statistics* (forthcoming).

Jan von der Goltz, Jan and Prabhat Barnwal (2014). "Mines: The Local Welfare Effects of Mineral Mining in Developing Countries", Discussion Paper Number 1314-19, Department of Economics, Columbia University

J. Vernon Henderson (1985). *Economic Theory and the Cities*, 2nd edition, Academic Press

J. Vernon Henderson, Adam Storeygard, and David Weil (2012). "Measuring Economic Growth from Outer Space", *American Economic Review*, 102(2): 994-1028.

Walter Isard (1956). *Location and Space-Economy: A General Theory Relating to Industrial Location, Market Areas, Land Use, and Urban Structure*, John Wiley and Sons, New York

Jane Jacobs (1969). *The Economy of Cities*, Random House, New York

Rijōemi Jedwab (2014). "Urbanization without Structural Transformation: Evidence from Consumption Cities in Africa", George Washington University (mimeograph)

Andreas Kotsadam and Anja Tolonen (2013). "Mineral Mining and Female Employment", OxCarre Research Paper 114.

Guy Michaels (2008). "The Effect of Trade on the Demand for Skill: Evidence from the

Interstate Highway System”, *The Review of Economics and Statistics*, 90(4): 683-701

Stelios Michalopoulos and Elias Papaioannou (2013). “Pre-Colonial Ethnic Institutions and Contemporary African Development”, *Econometrica*, 81(1): 113-152.

Pigois, J.P., Groves, D.I., Fletcher I.R., McNaughton N.J., and Snee, L.W. (2003). “Age constraints on Tarkwaian paleoplacer and lode-gold formation in the Tarwa-Damang district, SW Ghana”, *Mineralium Deposita*, 38: 695-714.

Taylor, A.A., 2006. *An Economic History of the Ashanti Goldfields Corporation, 1895-2004: Land, Labour, Capital, and Enterprise*. PhD Thesis, London School of Economics.

United States Geological Survey, (1994-2010). “The Mineral Industry of Ghana” in: *Minerals Year Book*, ed. United States Geological Survey, various years.

Figure 1: Gold production of 10 largest mines in Ghana, 1991-2010

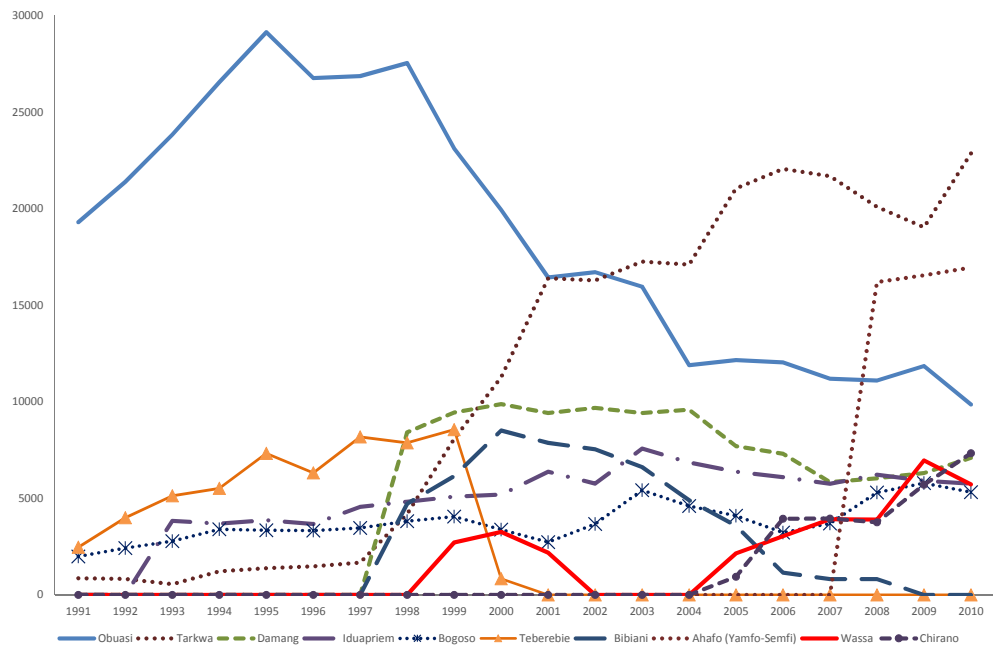


Figure 2: Ghana in 2000

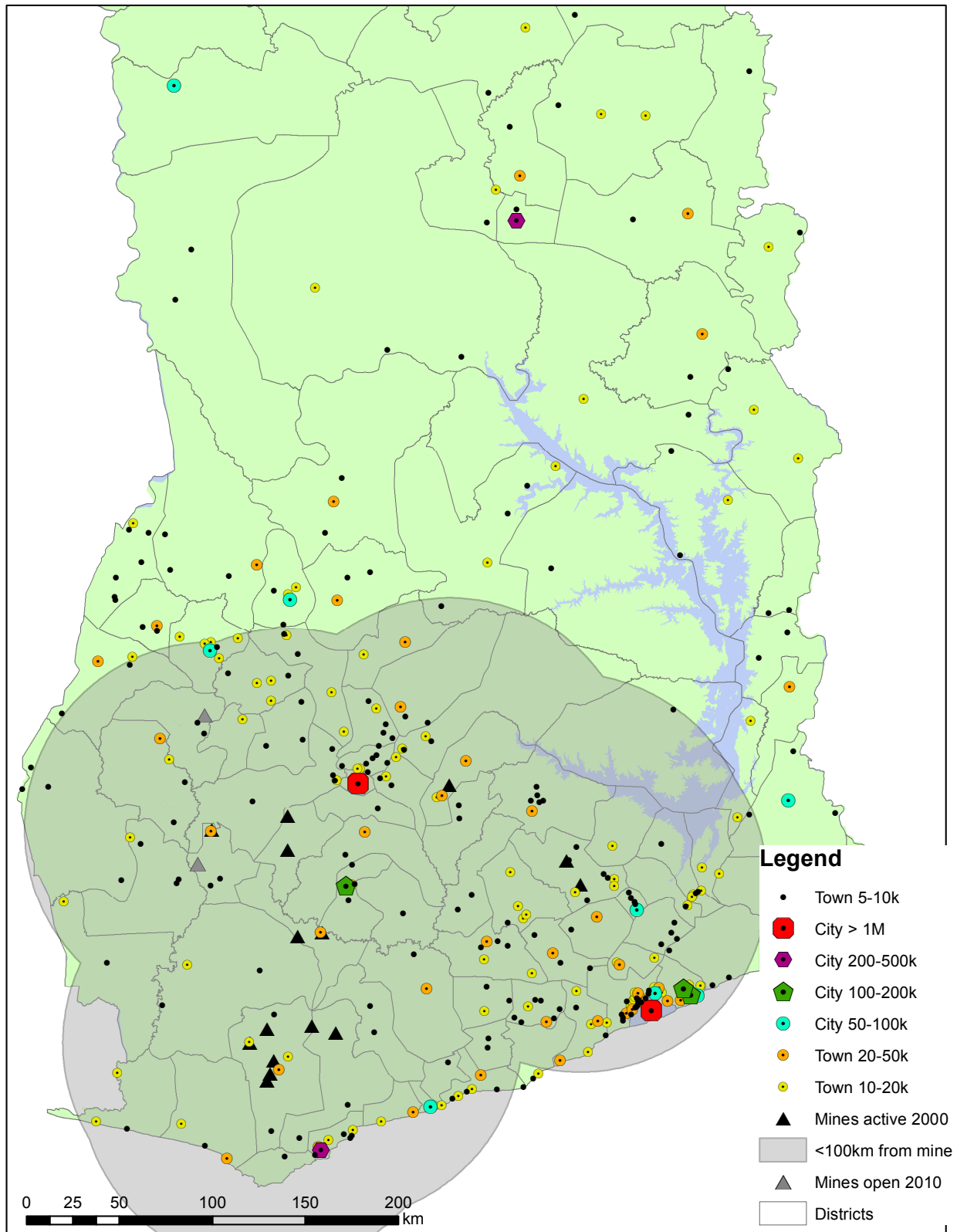


Figure 3: Ghana in 2010

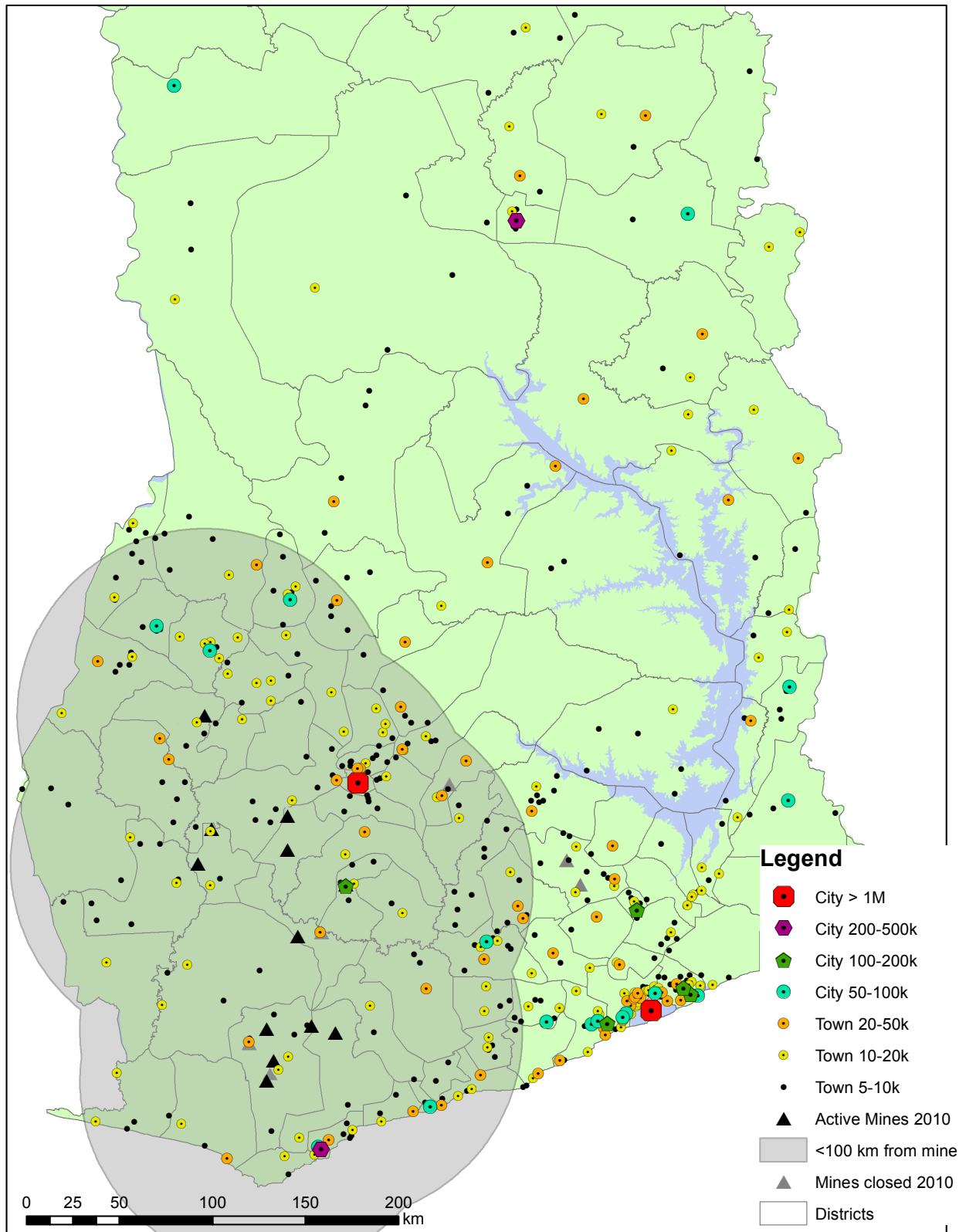


Figure 4: Night Lights in Ghana, Mining Region

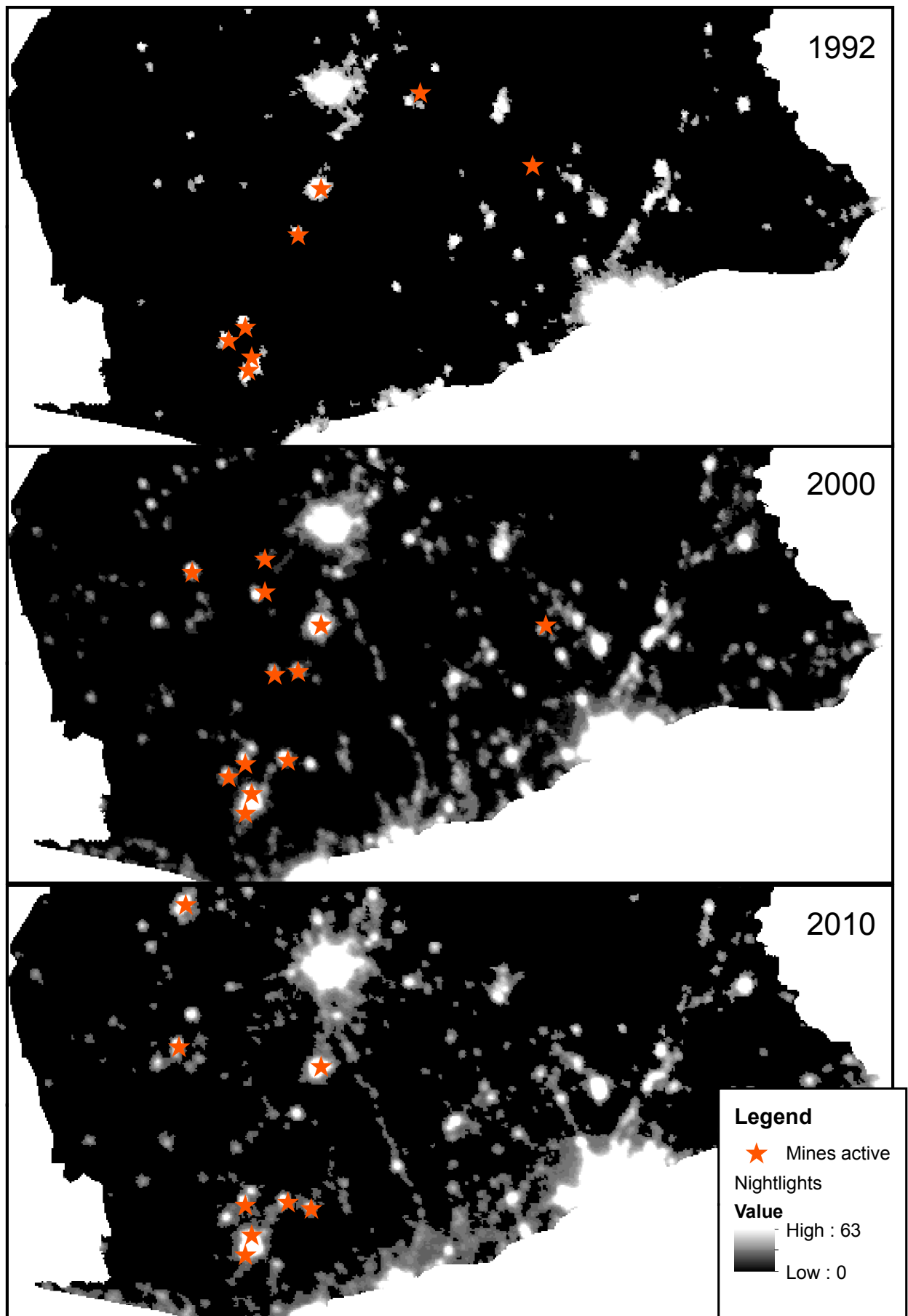


Figure 5: Gold Mining and Sector of Activity, 2010

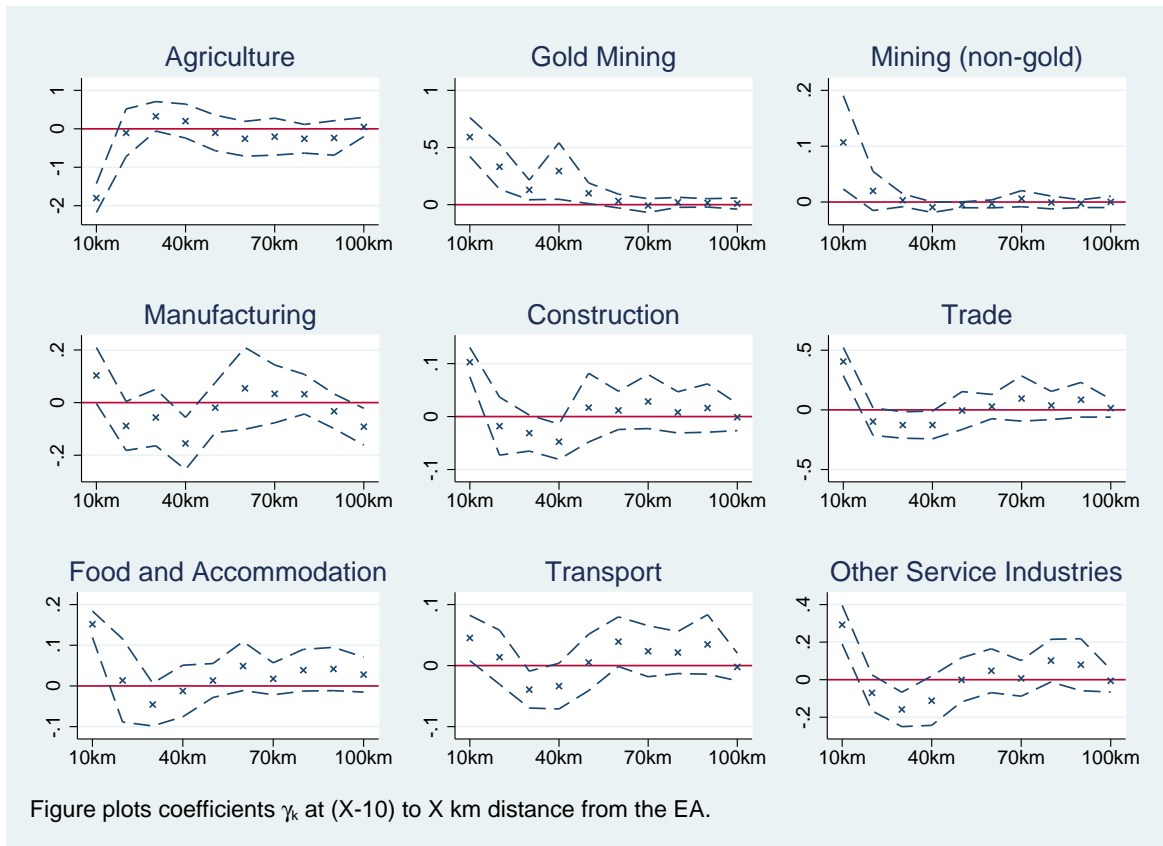


Figure 6: Gold Mining and Occupational Choice, 2010

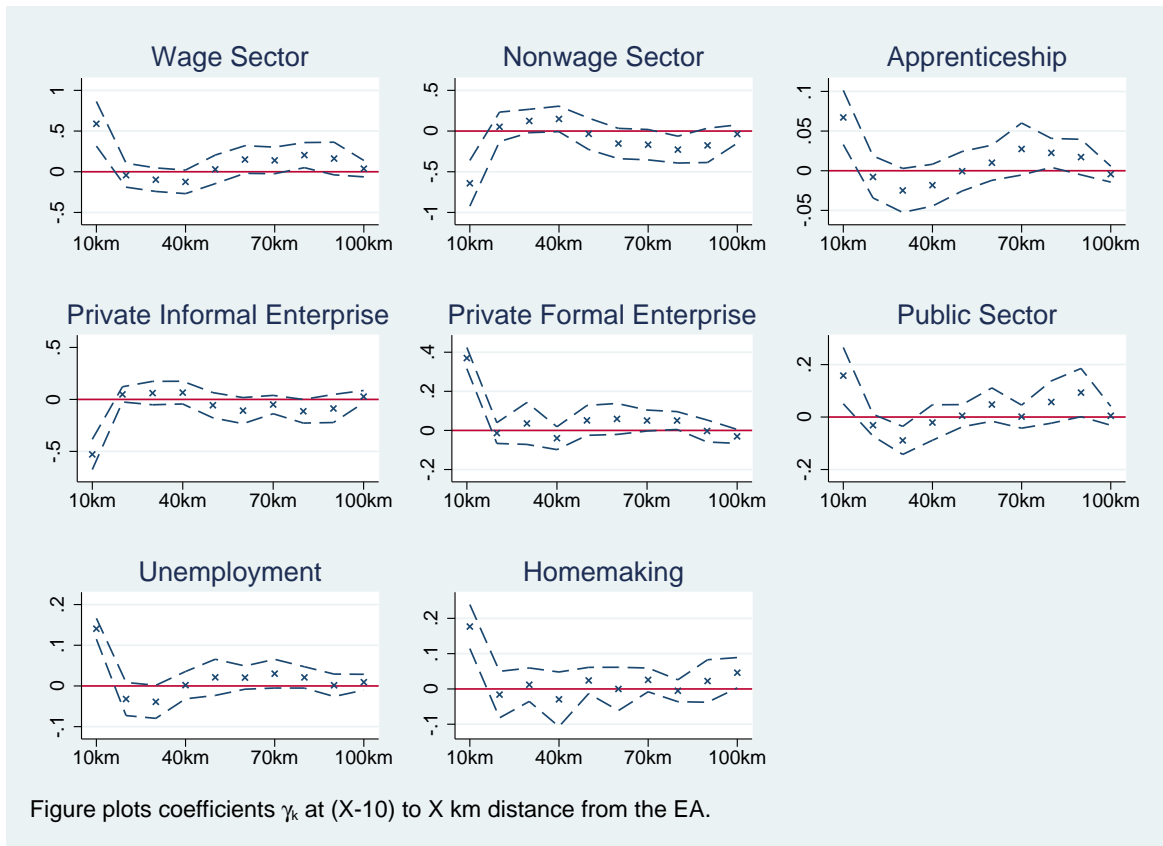


Figure 7: Gold Mining and Sector of Activity, 2010 - City and district controls

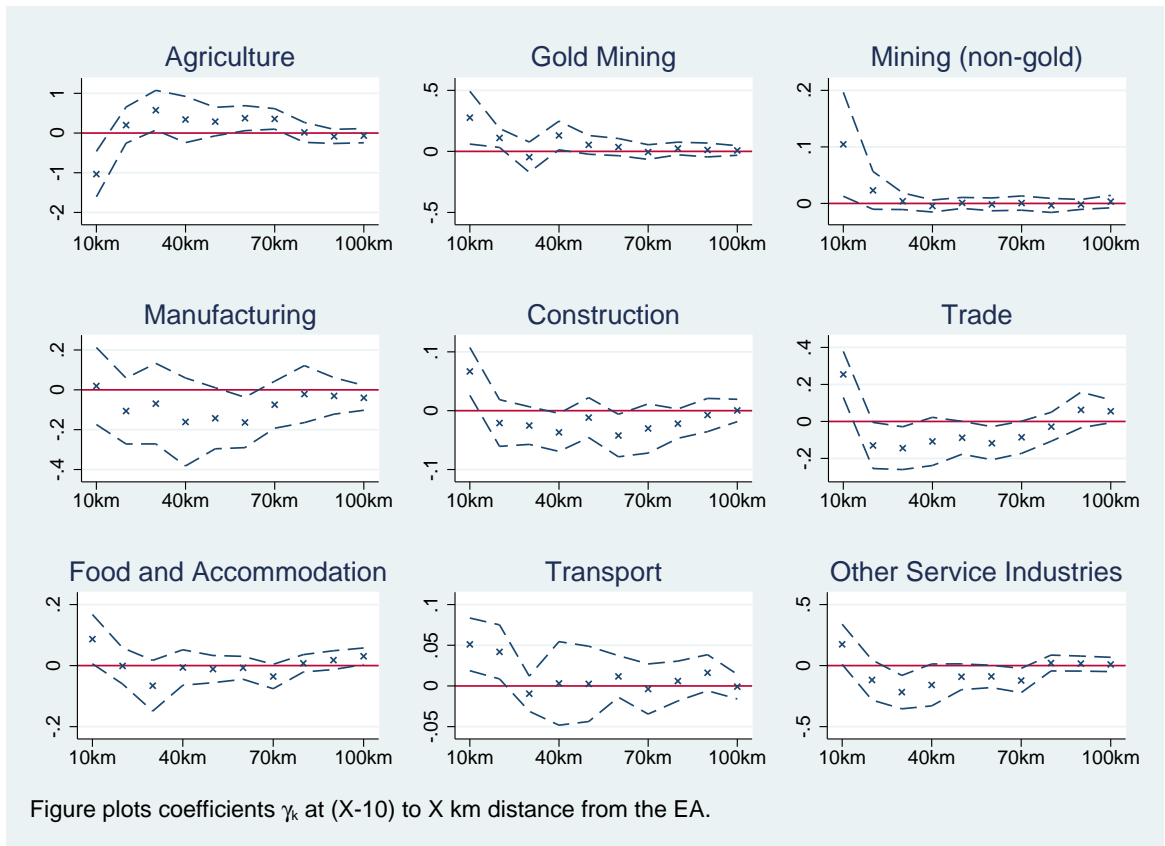


Figure 8: Gold Mining and Occupational Choice, 2010 - City and district controls

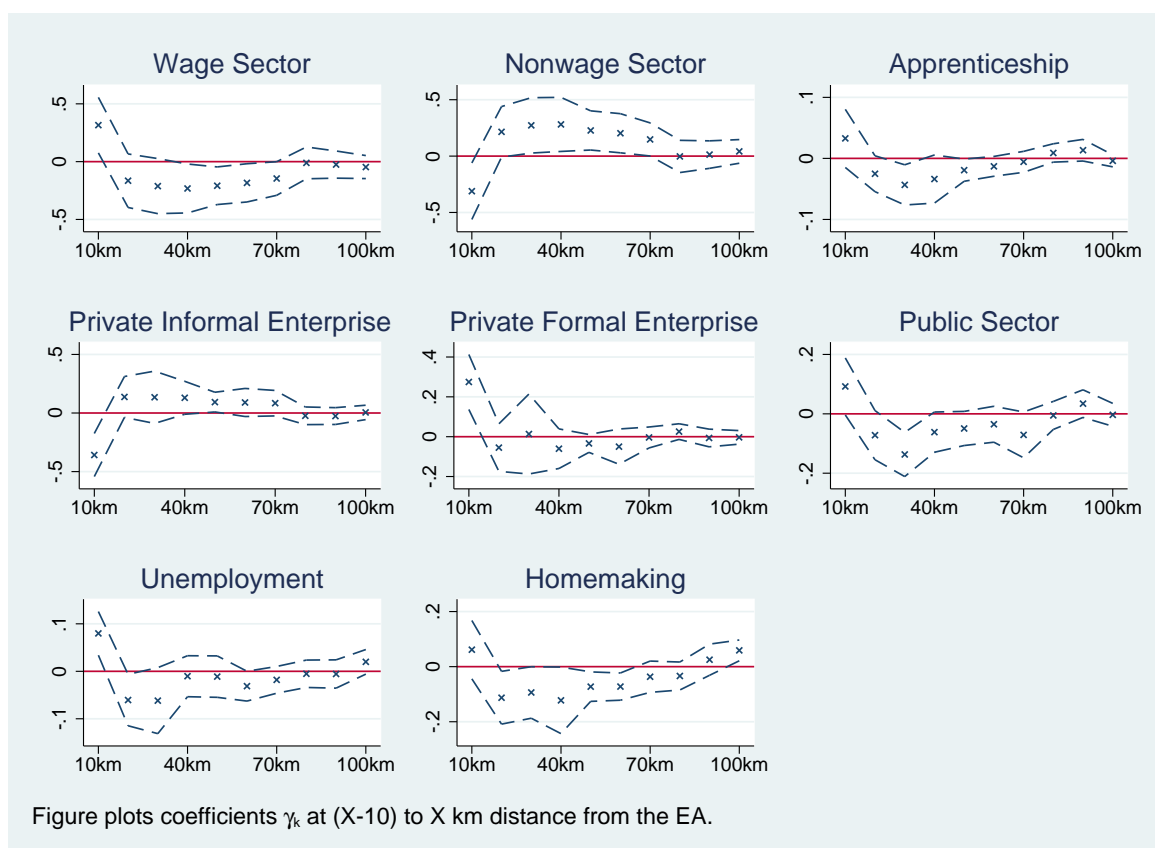


Figure 9: Gold Mining and Sector of Activity, First Differences

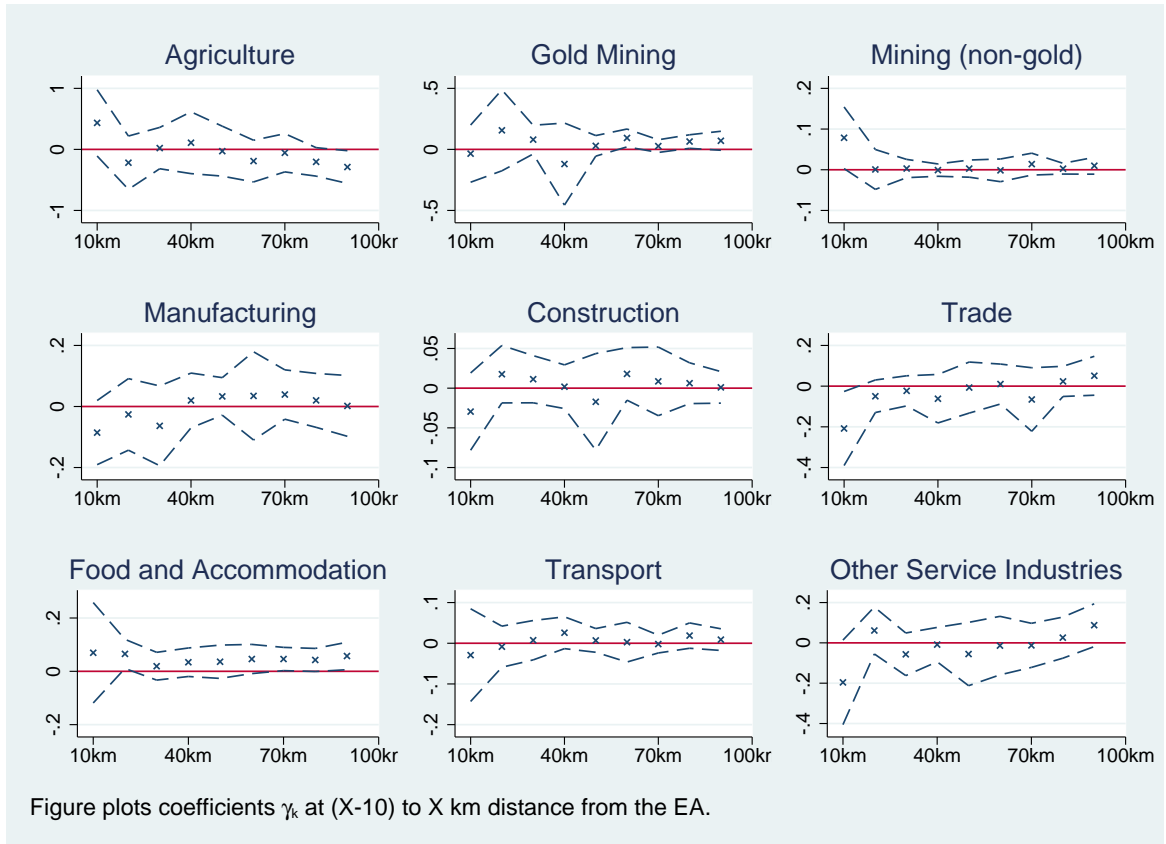


Figure 10: Gold Mining and Occupational Choice, First Differences

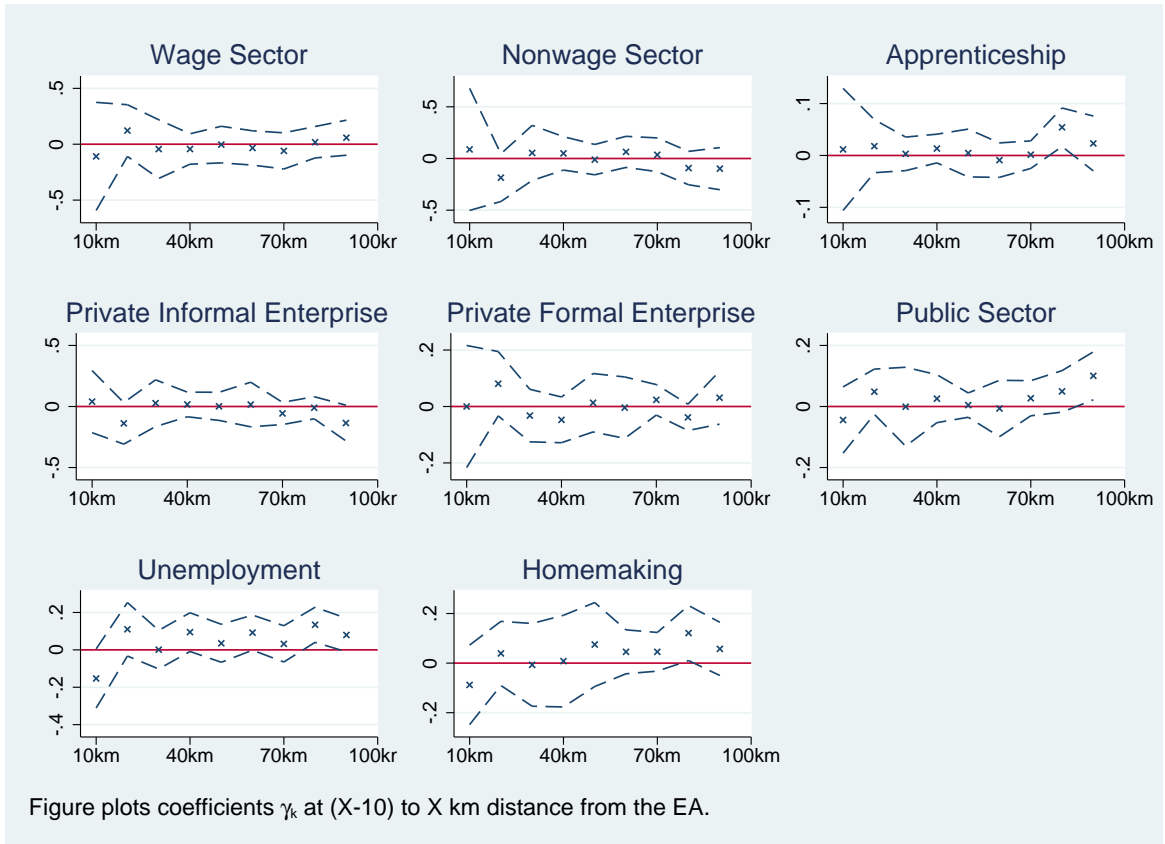


Figure 11: Gold Mining and Sector of Activity, Expanding Mine

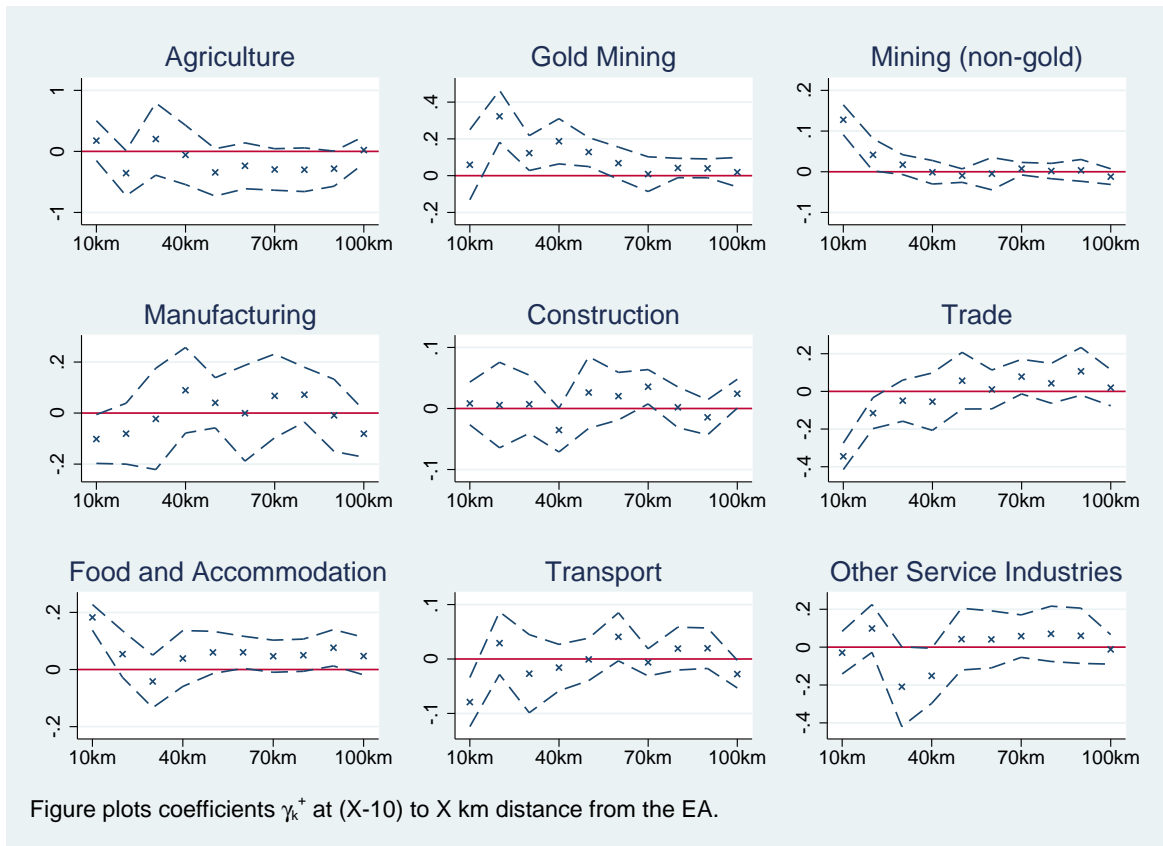


Figure 12: Gold Mining and Occupational Choice, Expanding Mine

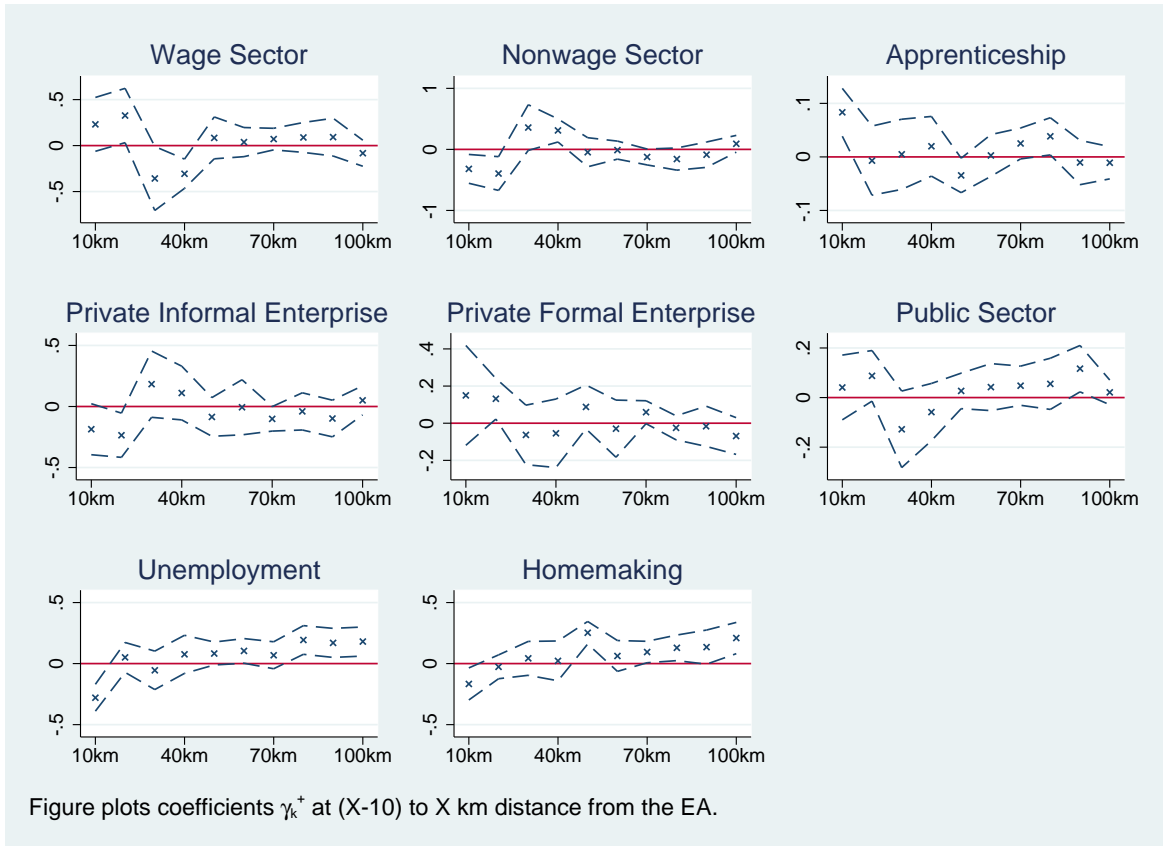


Figure 13: Gold Mining and Sector of Activity, Contracting Mine

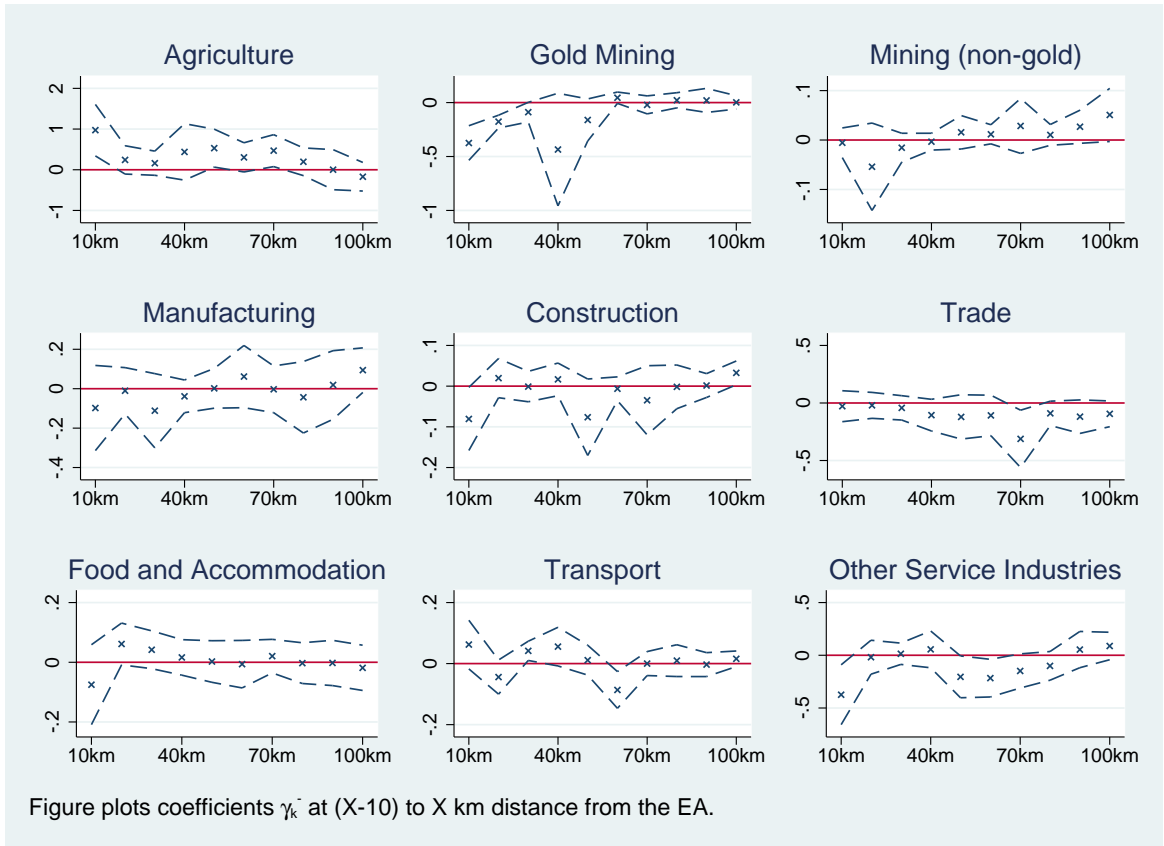


Figure 14: Gold Mining and Occupational Choice, Contracting Mine

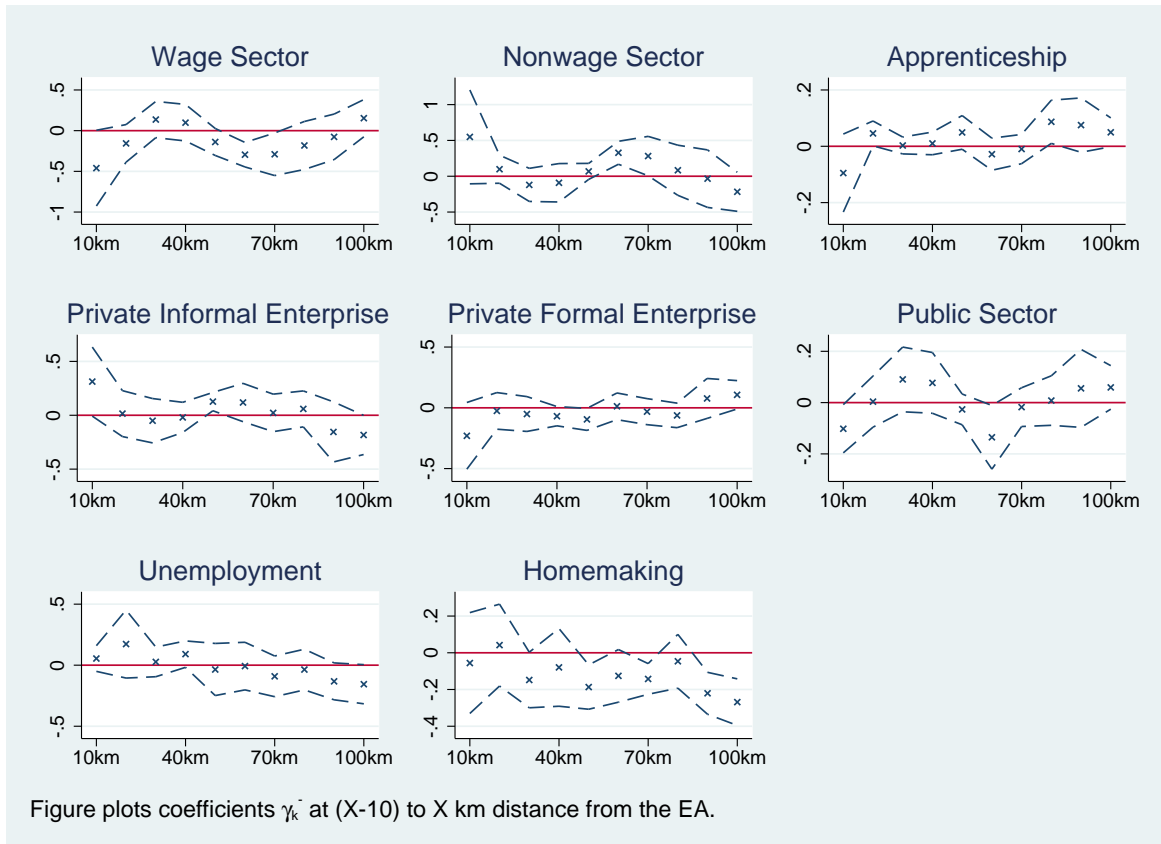


Figure 15: Gold Mining and Population & Employment Levels, Expanding Mine

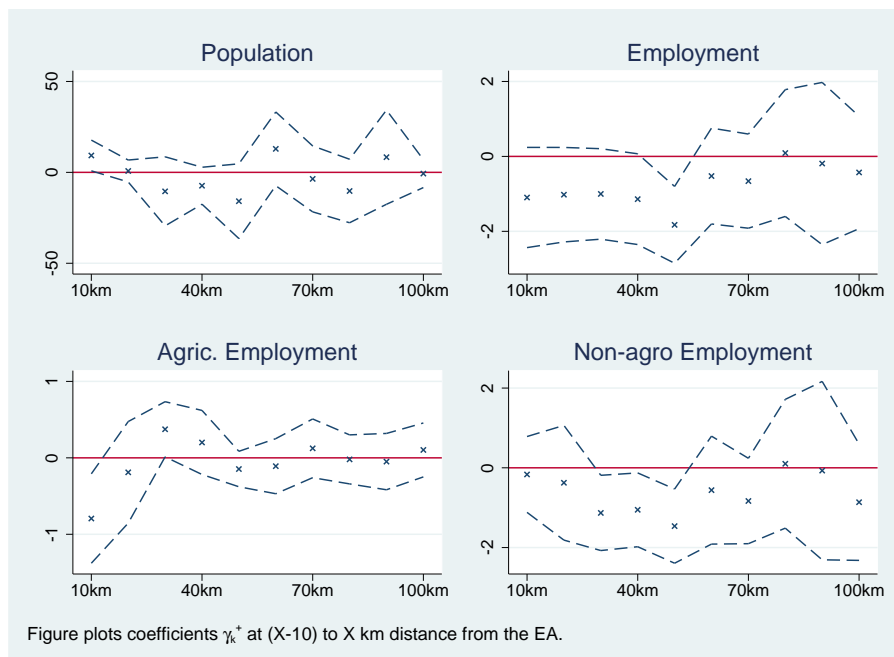


Figure 16: Gold Mining and Population & Employment Levels, Contracting Mine

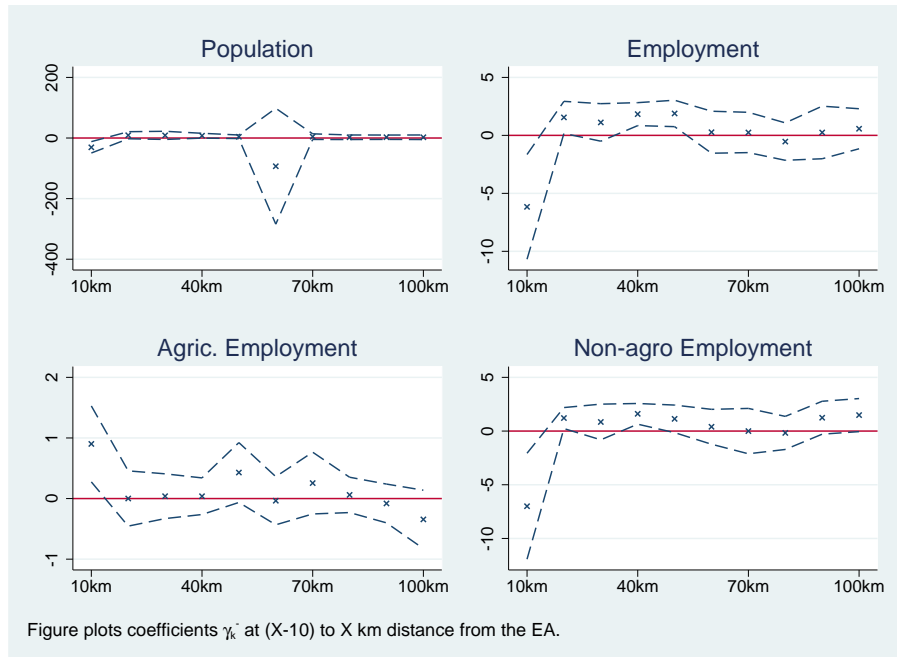


Figure 17: Gold Mining and Night Lights, 2000 and 2010 only

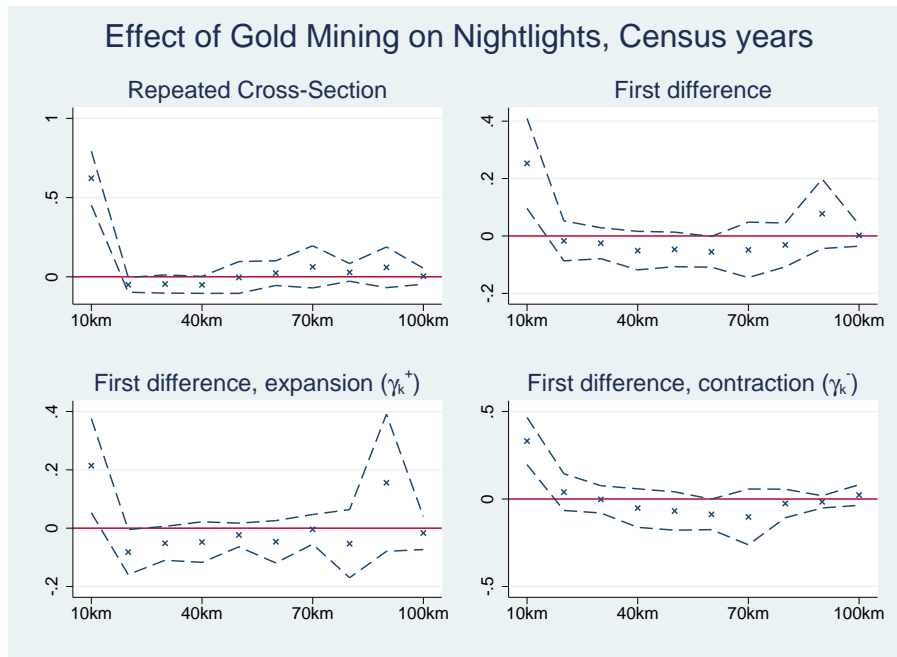


Table 1 - Gold Mines in Ghana, 1991-2010

Mine Name	District (in 2000 System)	Region	Average Yearly Production 1991-	Average Yearly Production 2001-	Opening	Closure
			2000	2010		
Obuasi	Adansi West	Ashanti	24,442	12,921		
Obotan	Amansie West	Ashanti	1,599	657	1997	2003
Bonte (alluvial)	Amasie West	Ashanti	759	482		
Obenemasi / Konongo	Asante Akim North	Ashanti	188	0	1998	
Ahafo	Asutifi	Brong Ahafo	0	4,970	2008	
Ayanfuri	Upper Denkyira	Central	936	36		2002
Dunkwa	Upper Denkyira	Central	97	0		1999
Asikam (alluvial)	East Akim	Eastern	27	0		2000
Goldenrae (alluvial)	East Akim	Eastern	26	0	1993	
Bibiani	Bibiani Anwhiaso-Bekwai	Western	1,938	3,330		
Chirano	Bibiani Anwhiaso-Bekwai	Western	0	2,564	2005	
Wassa	Mpohor Wassa	Western	598	2,787	1999*	
Damang	Wassa West	Western	2,775	7,841		
Iduapriem	Wassa West	Western	3,473	6,268		
Prestea	Wassa West	Western	945	0		
Bogoso	Wassa West	Western	3,200	4,386		
Tarkwa	Wassa West	Western	3,154	19,383		
Teberebie	Wassa West	Western	5,621	0		2001

Source: USGS Minerals Industry in Ghana, 1992-2010; company annual reports. Notes: * Wassa was not operating from 2002 to 2004.

Table 2 - Labor Market in Ghana in 2000 and 2010

	2000						2010					
	All			Male			Female			All		
	Number	%		Number	%		Number	%		Number	%	
A. Sectoral Composition												
Agriculture	314,130	50.82		158,463	51.46		155,667	50.18		338,918	38.58	
GoldMining	4,435	0.72		3,334	1.08		1,101	0.35		8,180	0.93	
OtherMining	3,766	0.61		2,353	0.76		1,413	0.46		2,167	0.25	
Manufacturing	67,441	10.91		32,301	10.49		35,140	11.33		97,600	11.11	
Construction	19,588	3.17		16,765	5.44		2,823	0.91		29,048	3.31	
Trade	99,036	16.02		32,583	10.58		66,453	21.42		173,439	19.74	
FoodAccom	14,540	2.35		2,715	0.88		11,825	3.81		48,933	5.57	
Transport	20,548	3.32		17,900	5.81		2,648	0.85		33,832	3.85	
Services	74,662	12.08		41,539	13.49		33,123	10.68		146,292	16.85	
Total	618,146	100.00		307,953	100.00		310,193	100.00		878,409	100.00	
B. Employment sector												
Wage sector												
Wage-employed	107,065	17.32		74,910	24.33		32,155	10.37		187,862	21.39	
Employers	33,112	5.36		17,171	5.58		15,941	5.14		44,842	5.10	
Non-wage sector												
Self-employed	423,730	68.55		191,516	62.19		232,214	74.86		537,223	61.16	
Family Worker	34,343	5.56		12,928	4.20		21,415	6.90		80,441	9.16	
Apprentice	14,011	2.27		8,247	2.68		5,764	1.86		21,299	2.42	
Other	5,885	0.95		3,181	1.03		2,704	0.87		6,742	0.77	
Total	618,146	100.00		307,953	100.00		310,193	100.00		878,409	100.00	
C. Employment by type of firm / organization												
Public (including parastatal)												
Private formal	49,170	7.95		32,720	10.62		16,450	5.30		63,129	7.19	
Private informal	54,478	8.81		35,271	11.45		19,207	6.19		65,484	7.45	
Other	512,425	82.90		238,532	77.46		273,893	88.30		744,807	84.79	
Total	2,073	0.34		1,430	0.46		643	0.21		4,989	0.57	
Total	618,146	100.00		307,953	100.00		310,193	100		878,409	100.00	
D. Labor Supply												
Working	618,146	76.96		307,953	79.45		310,193	74.63		878,409	78.43	
Unemployed	61,465	7.65		29,379	7.58		32,086	7.72		51,503	4.60	
Homemakers	54,802	6.82		14,280	3.68		40,522	9.75		67,671	6.04	
Other (student, pensioner, etc.)	68,832	8.57		35,976	9.29		32,856	7.90		122,338	10.93	
Total	803,245	100.00		387,588	100.00		415,657	100.00		1,119,921	100.00	

Table 3 - Night lights data for several periods

	1992-2010	1990s	2000s	1992	2000	2010
DN 0 (Unlit)	92.90%	94.07%	91.82%	96.88%	91.34%	89.62%
DN 1-5	5.19%	4.22%	6.07%	1.91%	6.61%	5.79%
DN 6-10	1.04%	0.92%	1.15%	0.65%	1.14%	2.88%
DN 11-20	0.40%	0.36%	0.44%	0.26%	0.42%	0.87%
DN 20-62	0.46%	0.41%	0.50%	0.30%	0.48%	0.76%
DN 63	0.01%	0.01%	0.01%	0.01%	0.01%	0.07%
Average DN	0.5654	0.4633	0.5867	0.2983	0.5903	0.9728

Table 4: Distribution of gold production by distance bins

	0-10km	10-20km	20-30km	30-40km	40-50km	50-60km	60-70km	70-80km	80-90km	90-100km	>100km
2000											
Mean	0.13	0.29	0.45	0.67	0.95	0.97	1.17	1.22	1.35	1.48	41.09
SD	1.49	2.28	2.81	3.43	4.12	4.02	4.48	4.45	4.76	4.99	14.39
Max	24.44	26.04	27.07	27.22	32.02	34.48	36.42	34.89	37.35	40.83	49.78
% Zero	97.3%	93.7%	89.7%	84.0%	80.4%	76.6%	75.1%	74.6%	74.8%	72.7%	0.0%
2000											
Mean	0.15	0.31	0.47	0.69	1.02	1.28	1.34	1.51	1.64	1.80	55.42
SD	1.45	1.99	2.38	2.74	3.75	4.40	4.12	4.26	4.45	4.82	16.55
Max	25.65	34.39	36.28	30.04	36.28	38.31	43.48	40.63	40.14	42.99	65.62
% Zero	97.7%	94.9%	92.7%	88.9%	86.0%	82.8%	81.0%	0.784876	77.2%	76.1%	0.0%

Note: Gold production [in t] within distance intervals from EA, distribution across all EA. % Zero refers to percentage of EA for whom gold production in the respective distance bin is zero.

Table 5: Nightlights Dynamic Panel Estimations

	(1) OLS	(2) Fixed Effects	(3) First Difference	(4) Dynamic Panel	(5) Dynamic Panel IV	(6) Dynamic Panel IV
L.Nightlight				0.966*** (0.009)	0.413*** (0.133)	0.374*** (0.134)
Mining 0-10 km	0.578*** (0.062)	0.213*** (0.051)	0.166*** (0.035)			0.242* (0.138)
Mining 10-20 km	-0.035 (0.021)	-0.003 (0.014)	0.015* (0.008)			-0.000 (0.081)
Mining 20-30 km	-0.033 (0.026)	0.005 (0.010)	0.015 (0.013)			0.064 (0.105)
Mining 30-40 km	-0.051** (0.022)	-0.009 (0.010)	0.004 (0.010)			0.027 (0.045)
Mining 40-50 km	-0.012 (0.040)	-0.015 (0.014)	0.001 (0.008)			0.244** (0.107)
Mining 50-60 km	0.016 (0.034)	-0.006 (0.009)	0.008 (0.012)			-0.098* (0.057)
Mining 60-70 km	0.049 (0.052)	-0.003 (0.022)	0.005 (0.011)			0.051 (0.078)
Mining 70-80 km	0.018 (0.027)	0.003 (0.010)	0.019** (0.007)			-0.058 (0.046)
Mining 80-90 km	0.033 (0.042)	0.042*** (0.015)	0.027*** (0.010)			-0.070 (0.089)
Mining 90-100 km	0.001 (0.021)	0.025** (0.010)	0.024*** (0.009)			-0.111 (0.092)
Constant	0.927*** (0.303)	0.957*** (0.094)		0.352*** (0.052)		
N	235410	235410	223020	223020	210630	210630
AB AR(1) test				0.000	0.000	0.000
AB AR(2) test				0.012	0.208	0.298

*** p<0.01 ** p<0.05 * p<0.10. All regressions contain dummies for every time period. Yearly data for 12,390 EA from 1992-2010. Column(5) uses instruments the difference equation with the 4th lag of Nightlight. Column (5) additionally instruments the difference of gold production with its 2nd lag. AB AR(*p*) test reports p-values of the Arellano-Bond (1991) test for serial correlation of order *p*.

Appendix for online publication

Figure A1: Gold Mining and Sector of Activity, 2000

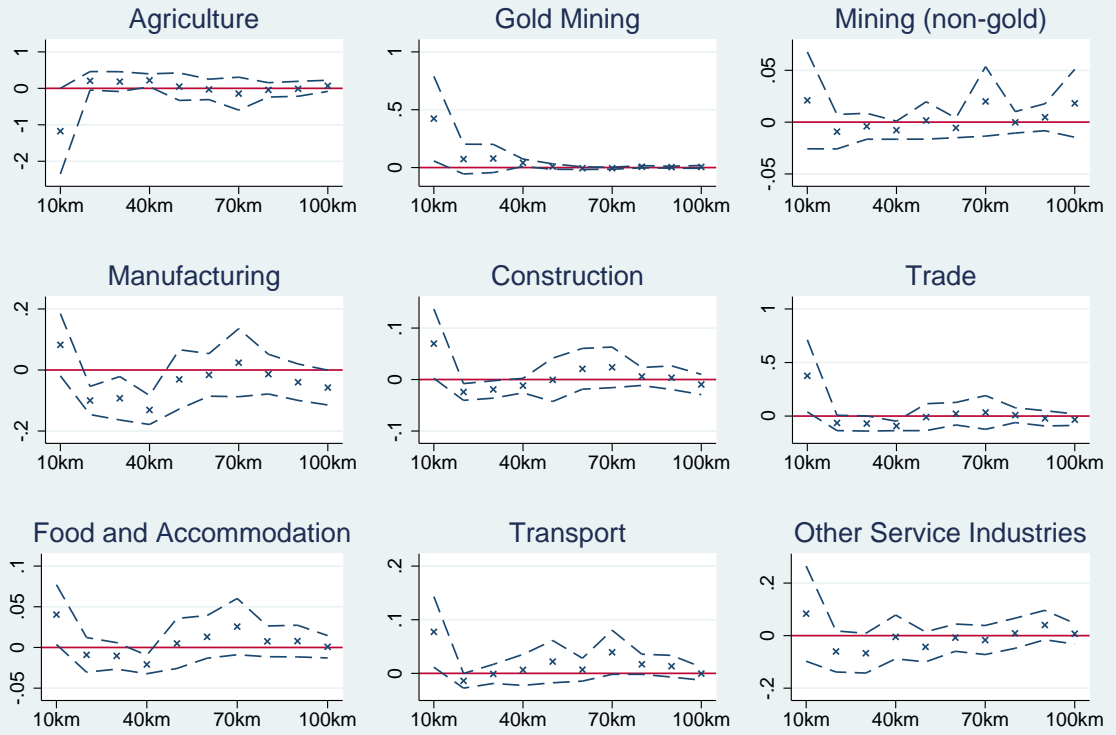


Figure plots coefficients γ_k at (X-10) to X km distance from the EA.

Figure A2: Gold Mining and Occupational Choice, 2000

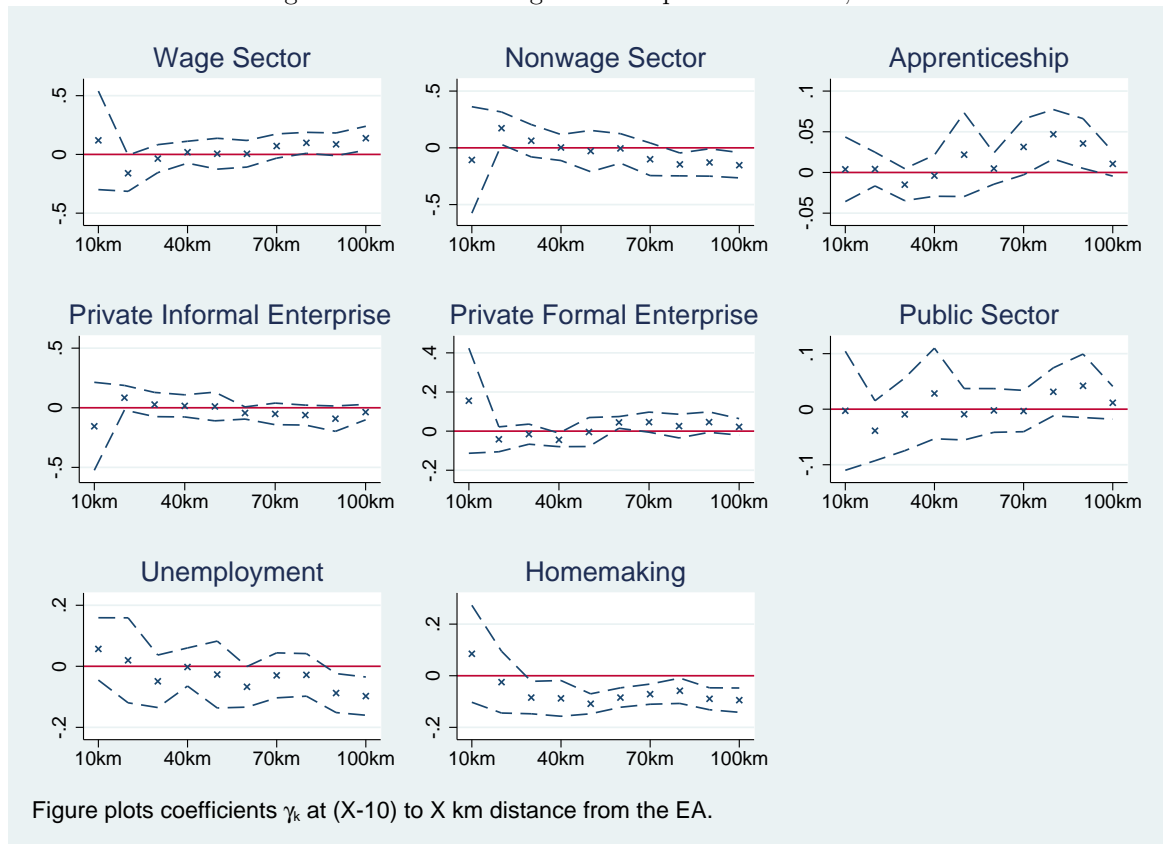


Figure A3: Gold Mining and Sector of Activity, 2010 - City controls

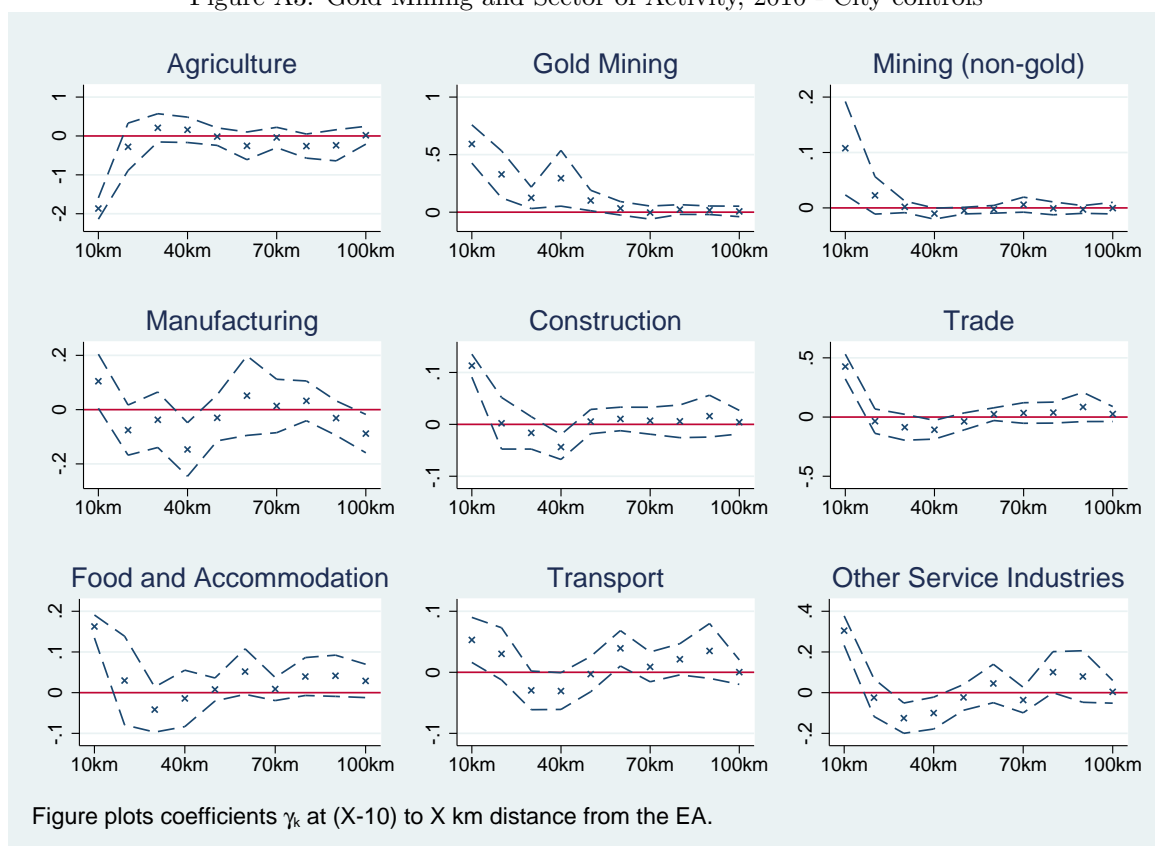


Figure A4: Gold Mining and Occupational Choice, 2010 - City controls

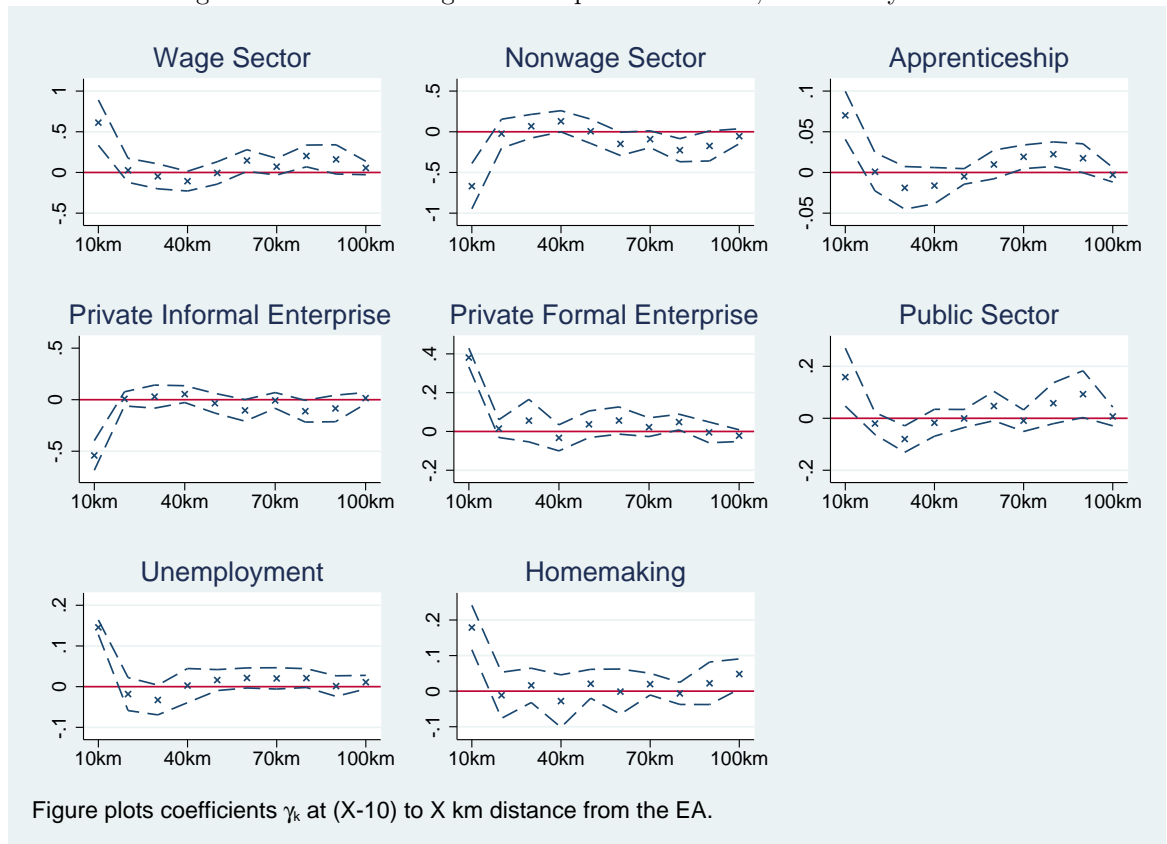


Figure A5: Gold Mining and Sector of Activity, 2000 - City controls

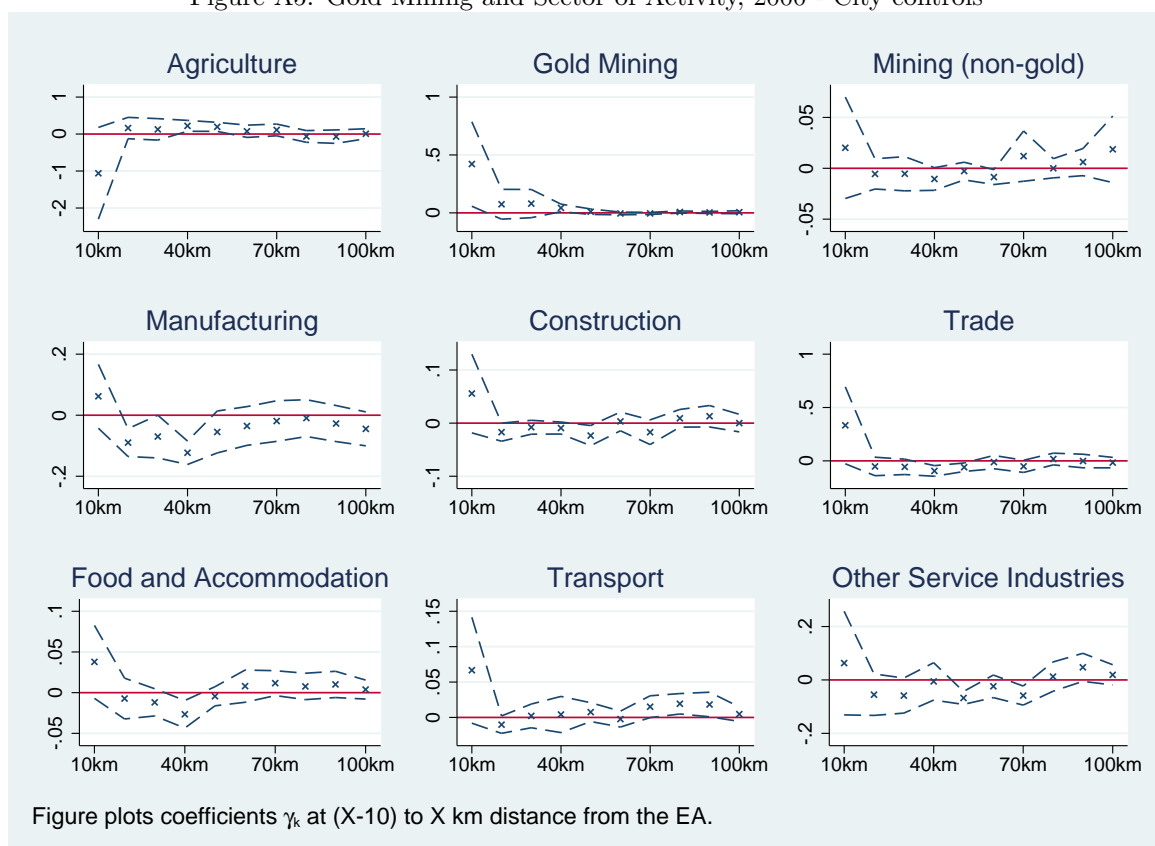


Figure A6: Gold Mining and Occupational Choice, 2000 - City controls

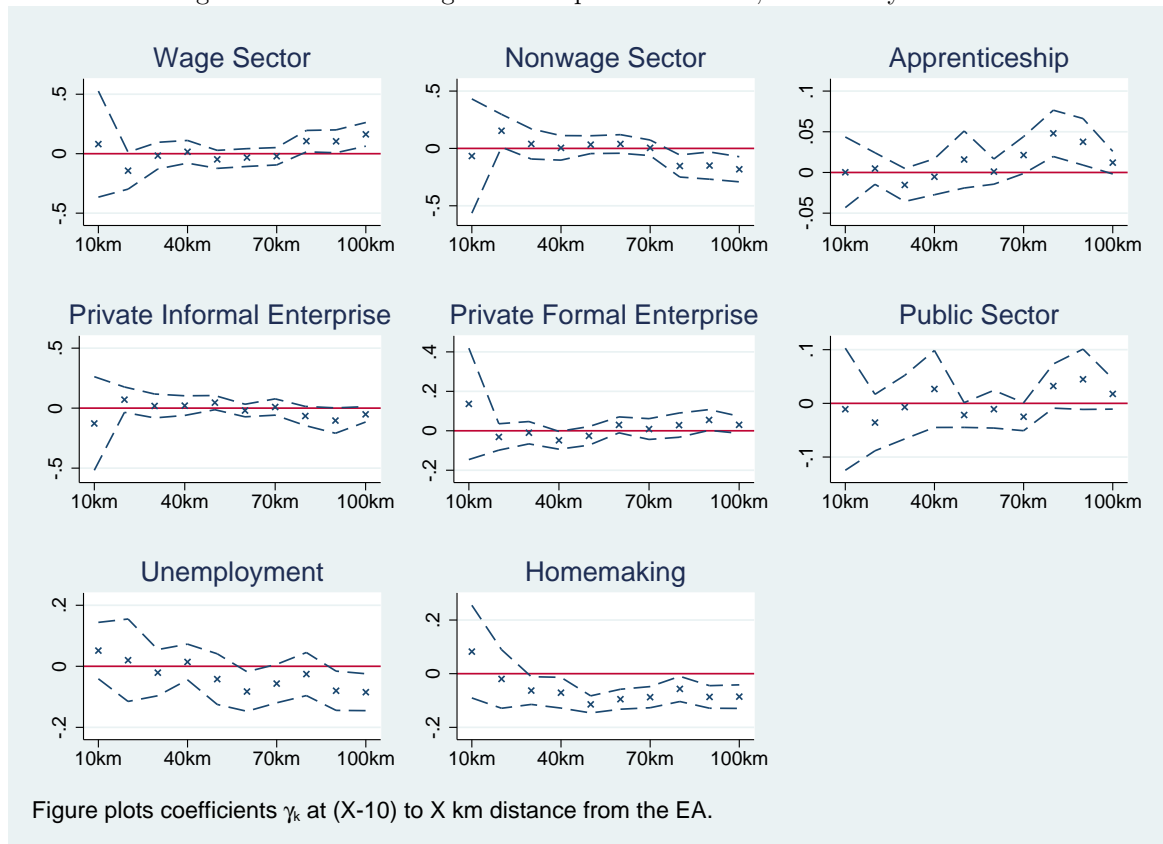


Figure A7: Gold Mining and Sector of Activity, 2010 - Slope, elevation and roads controls

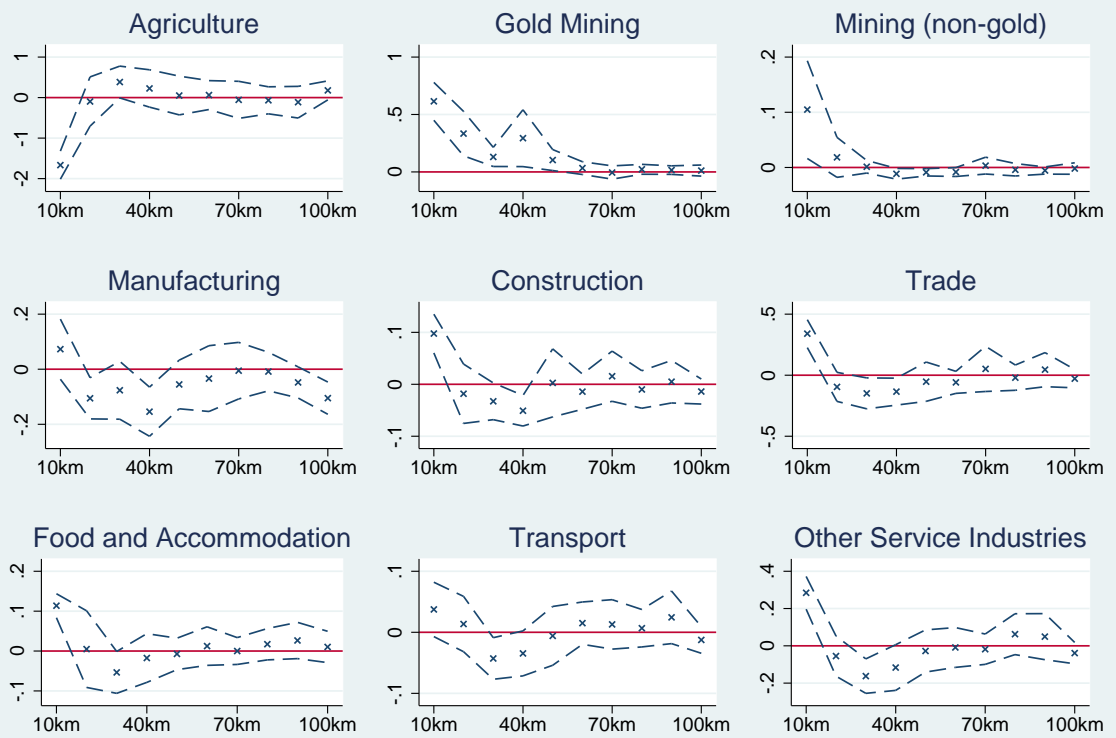


Figure plots coefficients γ_k at (X-10) to X km distance from the EA.

Figure A8: Gold Mining and Occupational Choice, 2010 - Slope, elevation and roads controls

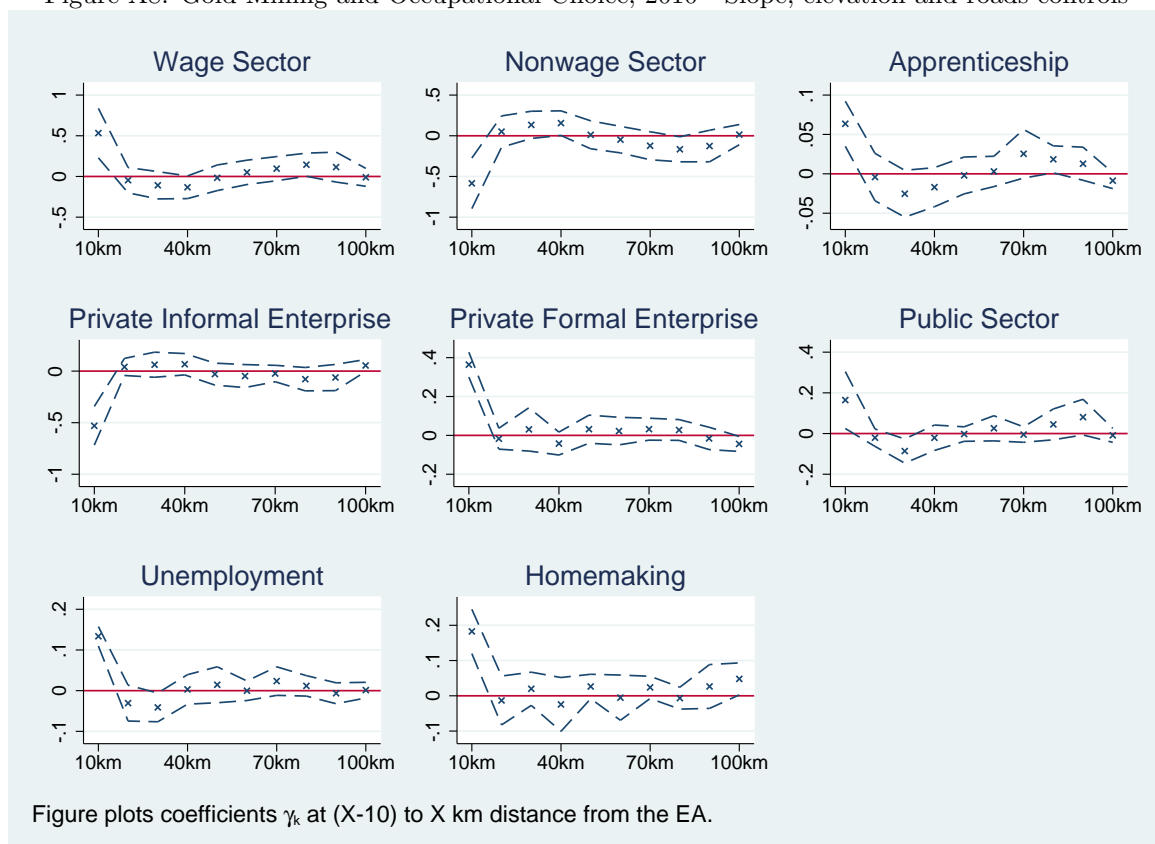


Figure A9: Gold Mining and Sector of Activity, Expanding mine - Slope, elevation and roads trend controls

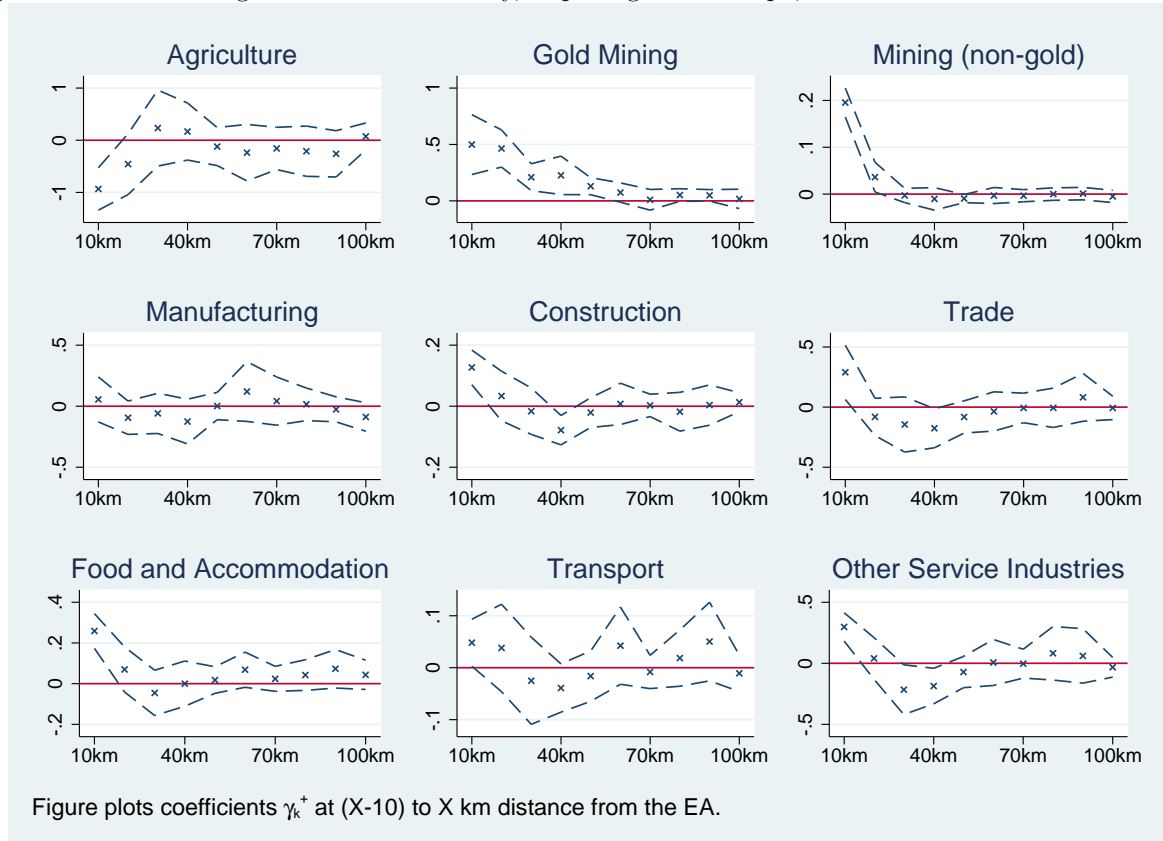


Figure A10: Gold Mining and Occupational Choice, Expanding mine - Slope, elevation and roads trend controls

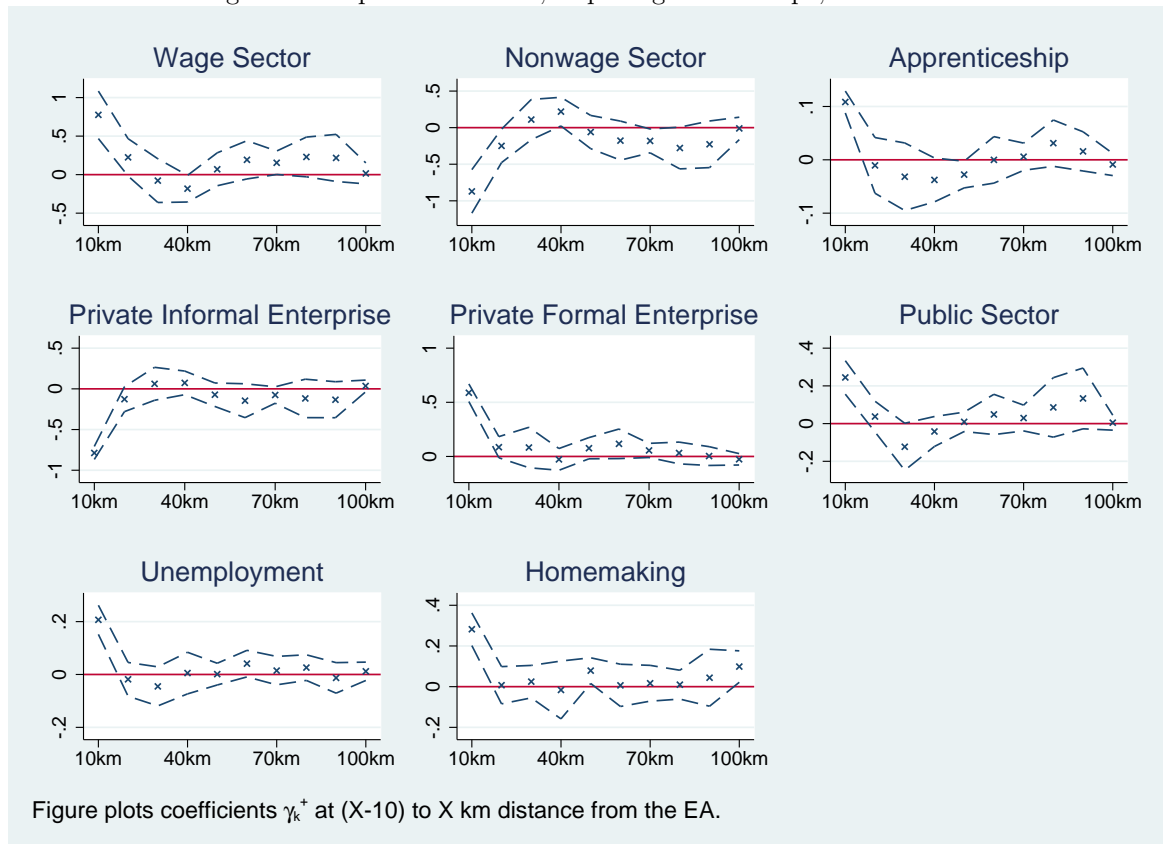


Figure A11: Gold Mining and Sector of Activity, Contracting mine - Slope, elevation and roads trend controls

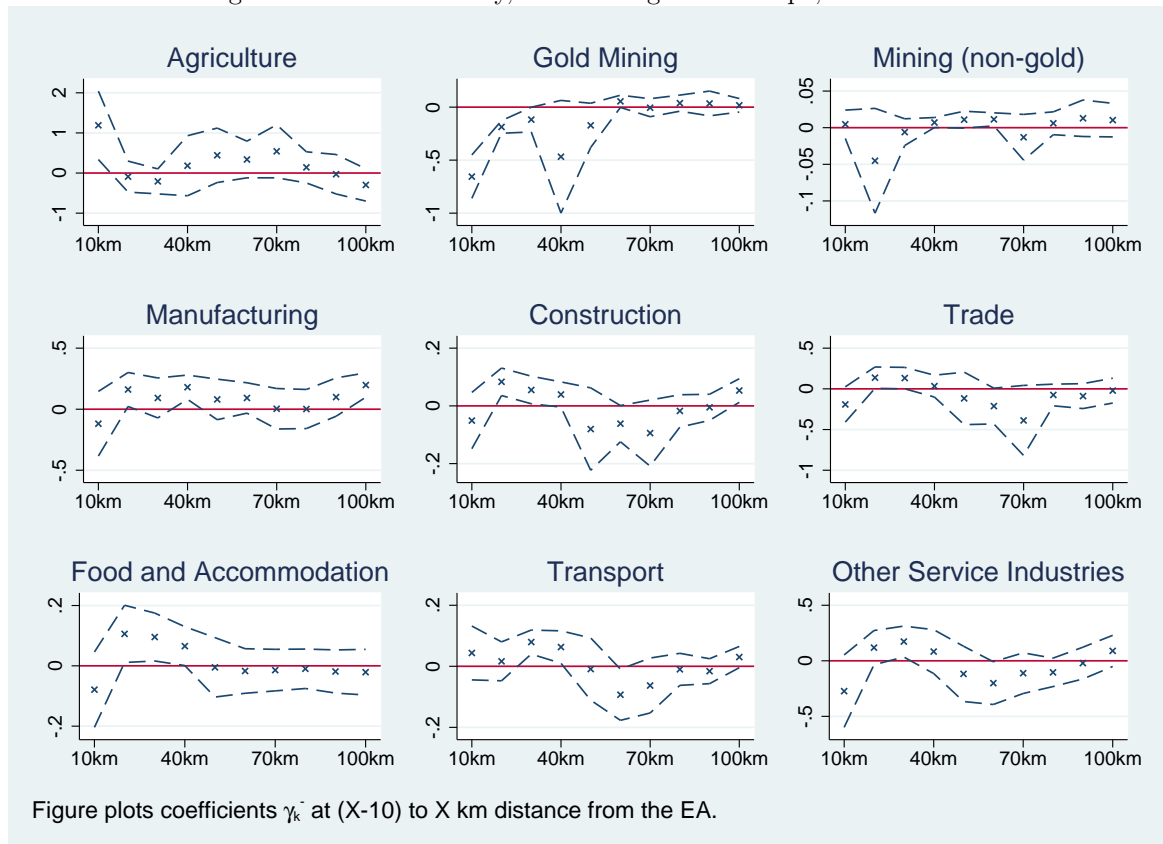


Figure A12: Gold Mining and Occupational Choice, Contracting mine - Slope, elevation and roads trend controls

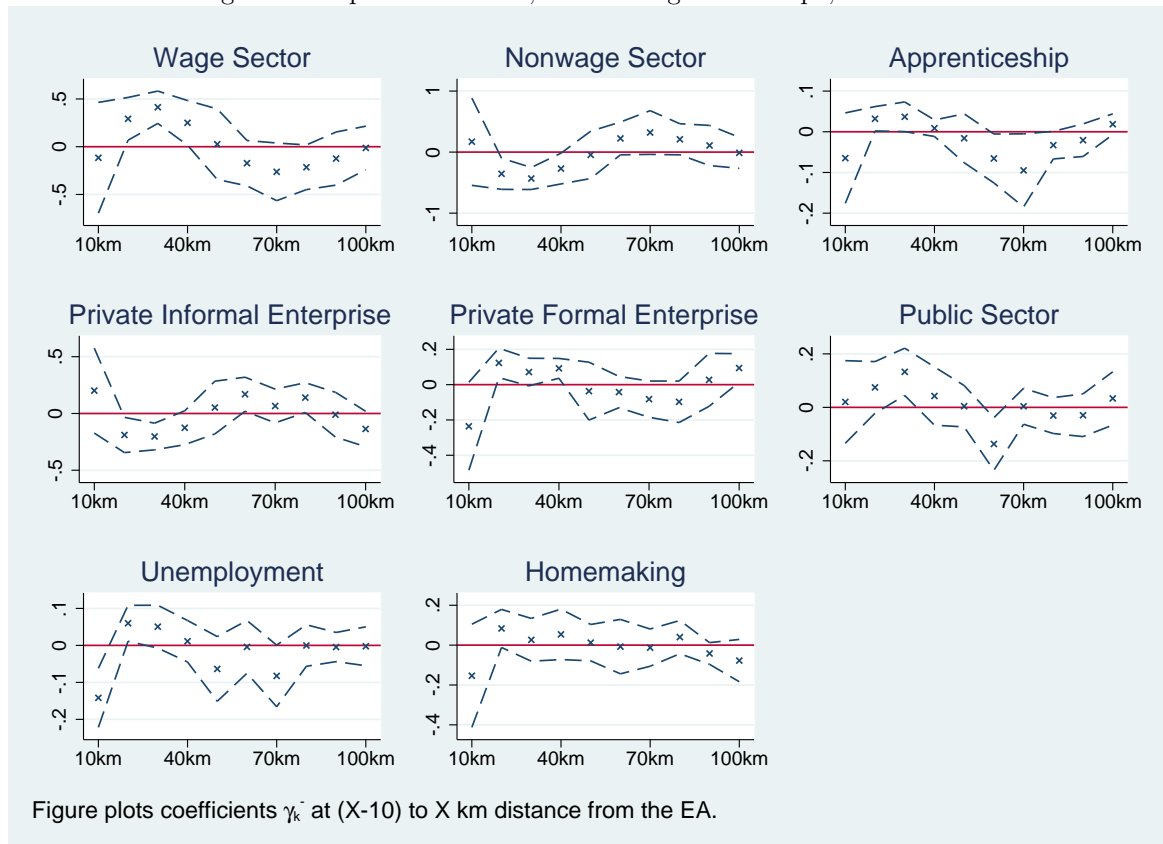


Figure A13: Employment Change in Gold Mining When Mine Expands

Split of Gold Mining employment

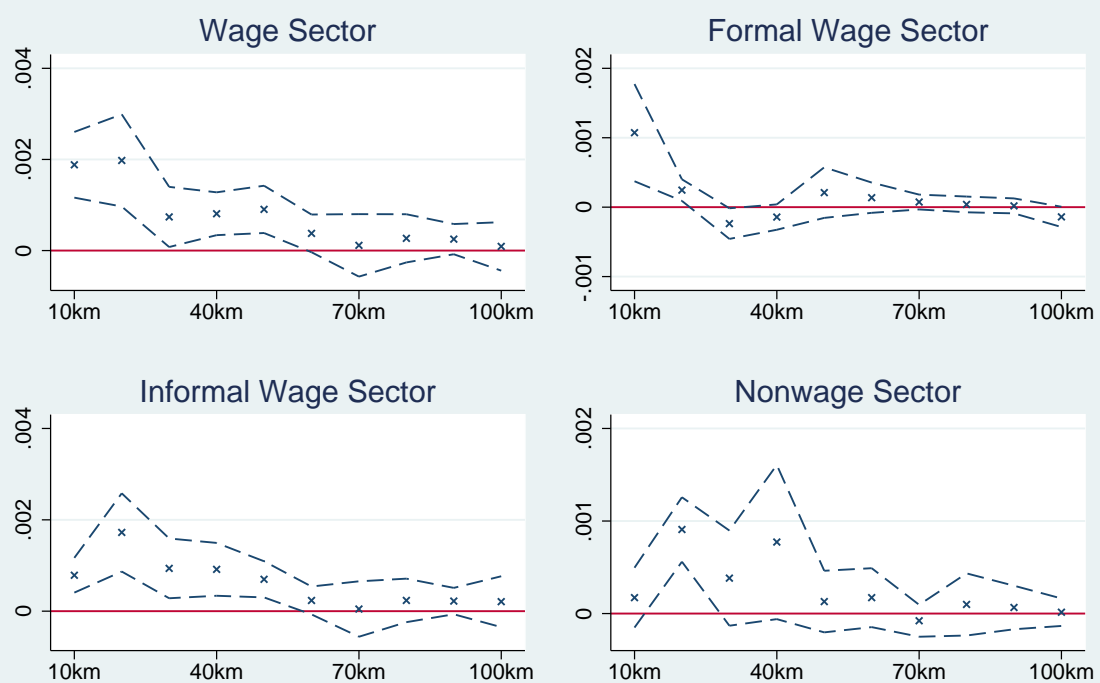


Figure plots coefficients γ_k^+ at (X-10) to X km distance from the EA.

Figure A14: Employment Change in Gold Mining When Mine Contracts

