Munich Personal RePEc Archive

Convergence Heterogeneity at the Local Level in Sub-Saharan Africa

Charpe, Matthieu

International Labour Organization

8 June 2022

Online at https://mpra.ub.uni-muenchen.de/114860/
MPRA Paper No. 114860, posted 11 Oct 2022 01:16 UTC

Convergence Heterogeneity at the Local Level in Sub-Saharan Africa

Matthieu Charpe *

International Labour Organization

June 8, 2022

Abstract

This paper tests for convergence in labor productivity at the local level in 10 Sub-Saharan countries, disaggregated into 1136 administrative entities. This work combines nighttime lights data and a unique set of population censuses to produce local measures of growth, employment and sectoral shares. We find evidence of unconditional convergence across sectors in the range of 2%. However, convergence is heterogeneous and conditional on both manufacturing and services employment shares. Convergence is also associated with proximity to the main city, moderate population density, low land suitability and relatively moderate temperature. Lastly, the within effect dominates the between effect.

Keywords: Local convergence, nighttime lights, heterogeneity, local labor market, structural transformation, census

JEL classifications: J23, J46, R11, R23, O14

*International Labour Organization, 4 route des Morillons, 1211 Geneva, Switzerland, E-mail: charpe@ilo.org. We would like to thank Slim Bridji for his contribution to data construction and preliminary analysis. We would like to thank the Institute of Statistics of Côte d'Ivoire in particular Désiré Aka Dore and Edmond Yao Koffi, the Ghana Statistical Service in particular David Kombat, Godwin Gyebi and Rosalind Quartey; the National Institute of Statistics of Rwanda in particular Dominique Habimana and Jean Marc Mukundabantu as well as Tania Smith from the Extended Public Works Program of the Republic of South Africa.

1 Introduction

The literature on structural transformation in Africa stresses the importance of sector specialization for productivity growth at national level (McMillan et al., 2014, 2017; Vries et al., 2015). However, implicit in the discussion on the emergence of a modern sector is the question of local specialization in production and employment, as well as labor movements from low productive areas to high productive areas. The aggregate data mobilized in this literature makes the exploration of the local dimension of structural transformation difficult.[1] Against this backdrop, the literature on convergence has recently explored the regional/subnational dimension but has left aside, in most cases, the issue of sector specialization partly due to the difficulty of measuring specialization locally (Gennaioli et al., 2014; Lessmann and Seidel, 2017; Adhikari and Dhital, 2021; Chanda and Kabiraj, 2020). A noticeable exception is Martin et al. (2018) for British cities.

The objective of this paper is to fill some of these gaps. This paper tests for convergence in nighttime lights per employment at the local level for 1136 administrative entities and 10 countries that are representative of sub-Saharan African. In addition to the speed of convergence, this paper explores convergence heterogeneity and its determinants. The contribution of sector specialization is the focus of this work together with the contribution of geographic and natural characteristics. A discussion of convergence regime is also provided.

This analysis combines nighttime light data and local employment data based on a unique set of population and housing censuses. Based on this, we can produce a proxy for labor productivity at the local level as well as a measure of local sector specialization based on local employment shares.

This paper therefore makes a twofold contribution: i) to the literature on structural transformation in Africa by providing a geographic disaggregation ii) to the literature on local convergence by including the dimension of sector specialization.

The literature on structural transformation in Africa focuses very much on the contribution of sector specialization and manufacturing, in particular for productivity growth. McMillan et al. (2014) and McMillan et al. (2017) bring evidence that the employment share of manufacturing is too small, and sometimes declining, to sustain aggregate productivity growth in Africa. Vries et al. (2015) point both to the increasing weight of services in Africa as well as to their declining productivity growth.

There is a limited number of papers that look at local convergence with either regional GDP data (Gennaioli et al., 2014), disaggregated national account in Great Britain (Martin et al., 2018) or nighttime lights (Adhikari and Dhital, 2021; Chanda and Kabiraj, 2020).[2] Regarding the importance of sector specialization for convergence, Martin et al. (2018) point to its central role for British cities.

[1] The literature on structural transformation proposes productivity decomposition based on sectoral national account and input-output tables at country level.

[2] A related paper is Lessmann and Seidel (2017) that look at regional inequality level and dynamic using nighttime lights data.

Rodrik (2012) tests explicitly for the importance of manufacturing for convergence but at the national level rather than subnational level.

A first result is that we find evidence of unconditional convergence across sectors around 2% in line with the "iron law of convergence" (Barro, 2012). The evidence of global convergence confirms the findings of Gennaioli et al. (2014) with regional GDP and the absence of a faster speed of convergence at subnational level. It is also in line with Adhikari and Dhital (2021) and individual country studies as Chanda and Kabiraj (2020) for India using nighttime lights. This result holds across different specifications and is robust to the possibility of measurement errors related to low lights and to different measures of the sum of lights.

However, it appears that convergence is heterogeneous and that some areas are left behind the convergence process. We therefore try to identify the characteristics of the administrative areas that explain convergence. When looking at the contribution of sector specialization for convergence locally, we find that convergence is conditional on the initial share of employment in manufacturing as well as in services, while the initial share of agriculture employment affects lights per employment negatively. There is a clear ranking as the magnitude of the effect is twice as large for manufacturing as it is for services. This confirms the central role of manufacturing for lights per employment growth. It is an important result as manufacturing employment has stagnated and sometimes declined at the aggregate level in Africa over the period considered. In addition, the result that services matter for convergence is new to the best of our knowledge.

Further disaggregation of manufacturing and services into subsectors confirms that subsector productivity impact the speed of convergence. Within services, the contribution of subsectors is heterogeneous with relatively high productivity sectors such as transport and relatively low productivity sectors such as retail. These results shed a new light on the contribution of manufacturing and services to structural transformation in Africa.

Looking at the determinants of convergence beyond sector specialization, we expand our set of explanatory variables with geographic characteristics and natural characteristics as in Henderson et al. (2017). We also add a measure of conflicts as armed conflicts persist in sub-Saharan Africa. We are able to show that administrative entities that converge the fastest are not only administrative entities with a specialization away from agriculture into manufacturing and services but also administrative entities with a proximity to the main city, relatively low population density, relatively low land suitability and more moderate temperatures. We then go a step further and test whether growth regimes are associated with these characteristics by estimating a threshold regression. Threshold regressions go a step further than the dummy interactions commonly found. We show that sectoral shares are associated with linear effects while geographic and natural characteristics are associated with convergence regimes. The impact on convergence speed is substantial as it can be 2 to 3 times larger in high convergence regimes.

Lastly, to echo the productivity decomposition that is central to the work on structural transformation in Africa, we perform a decomposition to enquire whether our proxy of productivity growth is explained by within administrative entities growth or whether it is explained by labor relocation between administrative entities. It also constitutes a first attempt to measure the importance of internal migration that is implicit when discussing structural transformation from rural/agricultural economies to urban/manufacturing economies as in Gollin et al. (2016). We find that in most cases, within administrative entities growth explains overall lights per employment. We also find that labor movements between administrative entities contribute positively to growth except in Côte d'Ivoire, Ghana and South Africa where labor movement went from high lights per employment areas to relatively low lights per employment areas. We discuss these three cases and show that they are consistent with conflicts in Côte d'Ivoire, existing evidences in Ghana and post-apartheid industrial reorganisation in South Africa.

Indirectly, this paper contributes to the flourishing literature that relies on nighttime light data or census data to explore a particular dimension of economic development such as transportation coast in Africa (Jedwab and Moradi, 2016), urbanization (Gollin et al., 2016), mining (Fafchamps et al., 2016) or the quality of institutions (Iddawela et al., 2021) amongst others.

The paper is organized as follow. Section 2 presents the data and a graphic representation of convergence. Section 3 estimates the speed of (un)conditional convergence under different specifications. Section 4 discusses the importance of sector specialization. Section 5 discusses geographic factors explaining convergence. Section 6 performs productivity decomposition across administrative entities. Section 7 concludes. Robustness check is presented in the appendix.

2 Data

The analysis of the convergence of labor productivity that we perform in this paper relies on a combination of two main databases. We use nighttime lights as the proxy for economic activity (Henderson et al., 2012). Nighttime lights over the period 1992-2013 is taken from the Earth Observation Group (EOG) from the Colorado School of Mines (Elvidge et al., 1997; Baugh et al., 2010).[3] Nighttime lights are recorded by DMSP[4] satellites and the EOG provides cloud-free composites for each calendar year. The data is cleaned to account for sunlight, glare, moonlight, observations with clouds and lightning from the aurora. There are different sets of satellite collecting data over time. We construct the annual observation by taking the mean light across available satellites for each year.[5] The data values range from 0 to 63 with zero cloud free observation hard coded to 255. Given the continent wide coverage of our study, the radiance is adjusted for the latitude of the pixel. However, we show that

[3] https://eogdata.mines.edu/dmsp/downloadV4composites.html

[4] Defense Meteorological Satellite Program

[5] In the appendix we show how the results are impacted if we use a three-year average ($t-1$, t, $t+k$) of the mean light across available satellites.

our results hold with or without latitude adjustment. The radiometric properties of the satellite are such that DMSP has difficulty measuring radiance at the two extremes: low lights maybe missed and the satellite suffers from saturation with bright lights. Regarding low lights, to avoid measurement errors, we add as a control variable the share of unlit pixel for each administrative entity. Regarding saturation, Henderson et al. (2012) argue that the share of top coded lights is close to zero in middle-and low-income countries.[6]

The radiance of each pixel is summed at the second level of administrative entities.[7] Some administrative entities display zero sum of lights at a point in time. Henderson et al. (2017) consider that a sum of light of zero in areas with non-zero population is a censoring issue. They assign the lowest observed light to the grid cell in order to reduce the gap between areas with no light and areas with the smallest non-zero values. Here we propose two transformations, $(1 + y^s_{i,j,t})$ in the main body of the paper and $max(1, y^s_{i,j,t})$ in the appendix as robustness check, with $y^s_{i,j,t}$ the sum of lights in the administrative entity i of country j in year t with light definition s. Note that other studies exclude administrative entities with zero light at the beginning of the period. As our objective is to look at convergence, we chose to retain all administrative entities. However, in the appendix, we show how the inclusion or exclusion of certain administrative entities impacts the speed of convergence.

Local employment is measured based on housing and population censuses. Since censuses are not available for every country and every year, we restrict our analysis to those years and countries that are available. In particular, we make use of the employment module contained in the census; which gives information on employment and sector of activity. While labor force surveys provide a richer set of information on employment, labor force surveys are often non-consecutive over time and are not representative at the local level. As a key contribution of this study is the dynamic and local dimensions, we rely on census data for our measure of employment.

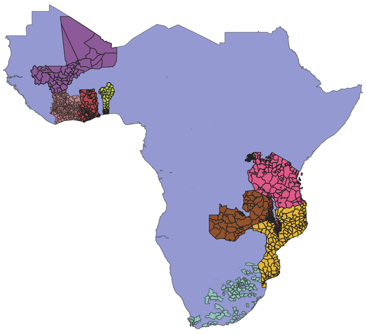
This analysis covers ten countries in sub-Saharan Africa representing Western, Eastern and Southern Africa: Mali, Côte d'Ivoire, Ghana, Bénin, Rwanda, Tanzania, Zambia, Malawi, Mozambique and South Africa. The set of countries is restricted to countries with two consecutive censuses containing an employment module and with geographic entities that are or can be made consistent over time. The country sample represents 35% of GDP and 26% of the population of sub-saharan Africa. The main employment ratios (employment population and employment share) in the 10 countries does not deviate from the ratios for sub-saharan Africa by more than 3 percentage points. In addition, the sum of light to population ratio is similar between the 10 countries 0.01 and sub-saharan Africa 0.007.

We use two consecutive censuses and match the census year with the same year for the nightlight satellite. The only exception is Côte d'Ivoire, whose census took place in 2014 and which is matched with the year 2013. For the ten countries, we have a census at the beginning of the 2000s and at the

[6] In sub-Saharan Africa, pixels with top-coding are 15 per 100000 overall and 3 per 100000 when excluding Nigeria and South Africa.

[7] See next paragraph for a discussion of the size of the administrative entities as defined in the census data.

Figure 1: Maps of 1136 administrative entities



end of the 2000s. For three countries, we have a third census in the early 1990s enabling us to run panel regressions. These ten countries are sub-divided into 1336 administrative entities corresponding to the second level of administrative entities. We worked with the Ghana statistical service to match the administrative entities of the census over time to generate 101 administrative entities. We re-aggregated the 416 administrative entities of Rwanda to 115 administrative entities to match the 'akare' administrative entities of 2001. We worked with the Institute of Statistics of Côte d'Ivoire and BNETD[8] to digitalize maps of administrative entities and reconcile maps over time as the number of 'sous-préfécture' increased from around 232 to 509 between the two census. We have also identified 'sous-préféctures' with a low response rate in the second census and aggregated these polygons with neighbouring polygons to produce 218 administrative entities over time. In South Africa, the numerous territorial reforms since the end of apartheid have made the reconciliation of administrative maps over time difficult. For censuses in 2001 and 2011, we matched the main 90 municipalities only. For South Africa, the analysis relies therefore on urban areas contrary to other countries. Cities and agglomerations have been identified using the Global Rural-Urban Mapping Project of the Socioeconomic Data and Application Center from Columbia university.[9] For the remaining countries, we use the IPUMS[10] data at the second level of administrative entities (IPUMS, 2019): 47 in Mali, 77 in Bénin, 113 in Tanzania, 55 in Zambia, 177 in Malawi and 143 in Mozambique. From the census, we aggregate employment at the local level and when a distinction is made per sector of activities we use the agriculture, manufacturing and services sectors (excluding the public sector). A map of administrative entities drawn upon can be found in figure 1.

We compute our proxy for labor productivity for the total economy, using the gross measure of nightlight luminosity as a proxy for output, as follows:

$$\omega^s_{i,j,t} = \frac{(1 + y^s_{i,j,t})}{n_{i,j,t}} \tag{1}$$

[8]Bureau National d'Etudes Techniques et de Développement

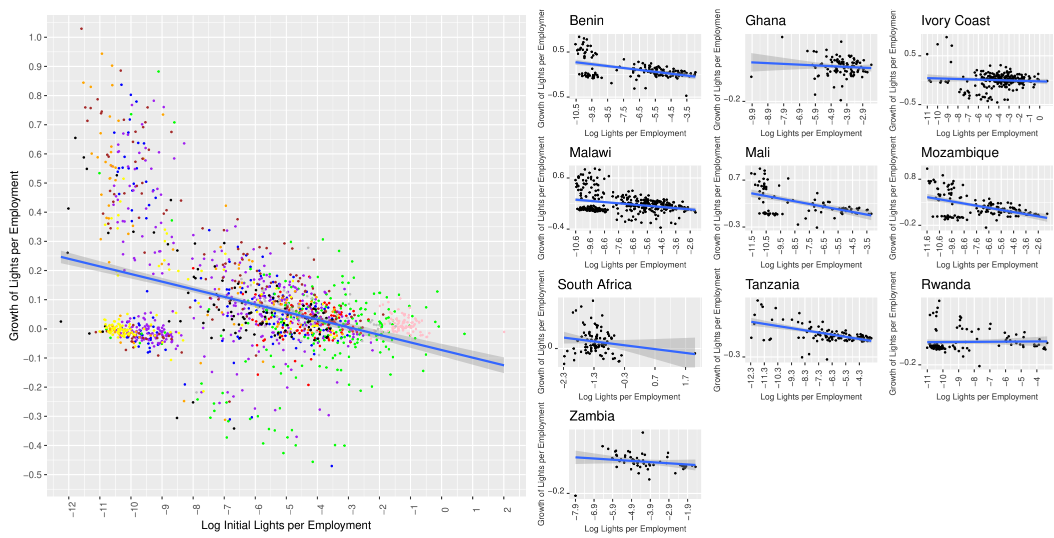
[9]http://sedac.ciesin.columbia.edu/data/set/grump-v1-settlement-points/data-download.

[10]Integrated Public Use Microdata Series

where $\omega^s_{i,j,t}$ denotes light per employment in the administrative entity i of country j in year t using the $s = gross$ definition of nightlight luminosity, $y^s_{i,j,t}$ represents the sum of lights, and $n_{i,j,t}$ is the total employment level. Then, we compute the annualised growth rate of light per employment for the total economy: $\hat{\omega}^s_{i,j,t,t-k} = \left(\frac{\omega^s_{i,j,t}}{\omega^s_{i,j,t-k}}\right)^{1/k} - 1$. There are alternative transformations in the literature such as $max(1, y^s_{i,j,t})$ (see appendix A for robustness checks).

Figure 2a displays the relationship between initial lights per employment and its annual growth rate in the subsequent decade. The colours represent countries. In particular, administrative entities with lower levels of lights per employment in base periods undergo more rapid growth in lights per employment in the subsequent decade, whatever the definition of nighttime light luminosity we use. Hence, the downward and significant slope represents a convergence result for local-level lights per employment.

Figure 2: Convergence in 1336 Administrative Entities and 10 African Countries



(a) All countries

(b) Individual Countries

Figure (a) illustrates unconditional convergence for 10 Sub-Saharan African countries and 1136 administrative entities. Figure (b) is country by country subfigures.

We can also plot the relationship between lights per employment in the base periods and its annual growth rate over the respective subsequent decade, country by country. Interestingly, Figure 2b shows that in most countries administrative entities with low initial lights per employment tend to be split into two groups: low and high growth rate. This is an indication that convergence is not homogenous across all administrative entities. This shows the importance of geographic disaggregation to study structural transformation. Administrative entities with high and low growth rate may cancel each other out in the aggregation process. See section 4 for further discussion of heterogeneity.

Table 1: Baseline and fixed effects specifications

	(1)	(2)	(3)	(4)	(5)	(6)
	lights per employment growth					
	Cross section				Panel	
log lights per employment t-k	-0.018***	-0.020***	-0.020***	-0.024***	-0.050***	-0.213***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.006)	(0.011)
unlit pixel share t-k		-0.052***				
		(0.010)				
R^2	0.11	0.11	0.14	0.32	0.15	0.82
N	1136	1136	1136	1136	602	602
Country FE	No	No	Yes	No	No	No
Regional FE	No	No	No	Yes	No	No
Admin FE	No	No	No	No	No	Yes
Time FE	No	No	No	No	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the results of the baseline specification both with and without country, regional and administrative entity fixed effects. This table also includes the control variable for the share of pixel with non zero population and zero sum of light.

3 Global (un)conditional convergence at local level

This section tests for unconditional and conditional convergence at the sector aggregate level and geographic disaggregate level. First, we estimate the baseline cross section specification, that is, we fit a linear regression model, with the annualised growth rate of lights per employment (over the 2000-2010 period) as the dependent variable and the log of initial lights per employment (in 2000) as the independent variable:

$$\hat{\omega}^s_{i,t,t-k} = \beta_0 + \beta_1 \log\left(\omega^s_{i,t-k}\right) + \varepsilon_{i,t,t-k} \tag{2}$$

where i denotes the administrative entity index, $t-k$ is the fixed initial time, and s represents the measure of nighttime light luminosity used. Note that in a first step, we test for unconditional convergence and therefore ignore any country fixed effects in this specification. We then discuss censoring issues linked to the radiometric properties of satellites, the inclusion of country and regional fixed effects on the estimated coefficient and the inclusion of second level administrative entities fixed effects in a panel of a smaller set of countries.

Table 1 column (1) reports the estimation results of equation (2). The significantly estimated coefficient in the baseline specification shows that a 1 percent higher lights per employment in the base period leads to a lower growth rate of lights per employment by 1.8 percent per year across the subsequent decade. The magnitude of the coefficient is in line with the *iron law of convergence* of 2% put forward in Barro (2012) using national GDP data. This magnitude is also in line with studies finding that convergence is surprisingly not faster at regional level despite smaller barriers (Gennaioli et al., 2014). These two papers have a global coverage with a weight for the number of African countries of 22% in the former and 7% in the latter. This result is also in line with studies using nighttime lights across country (Adhikari and Dhital, 2021) or for individual countries as Chanda and Kabiraj (2020) for India and Carrington and Jiménez-Ayora (2021) for Ecuador. Measuring convergence in labor productivity, Rodrik (2012) find evidence of unconditional convergence for manufacturing only in a magnitude of 2.9 percent a year. Discussing the case for African countries, Rodrik (2012) points to three reasons that could explain why convergence does not aggregate up. This includes nonconvergence within nonmanufacturing, the small size of manufacturing and the limited shift of labor towards manufacturing.

A question arises whether this first result is driven by the data to proxy production growth. The radiometric properties of satellites are such that it makes difficult to measure lights at both ends. Saturation may underestimate the growth of lights in bright areas. However, this effect is likely to be small in Sub-Saharan Africa as the share of top coded pixel is small.[11] The difficulty of DMSP satellites in measuring low lights could lead to biased estimates against convergence. This censoring issue tends to minimize the magnitude of convergence amongst low light areas. In order to control for this effect, we include the share of unlit pixel in the initial period in column (2).[12] The control variable modifies the convergence speed marginally from 1.8% to 2%. The share of unlit pixel is negative. The satellite difficulty in detecting low lights seems to underestimate the rate of convergence.

The baseline regression does not include fixed effects. The impact of fixed effects on the speed of convergence has been largely discussed in the literature. Growth model points that productivity growth depends on total factor productivity. However, given the difficulty to control for the factors impacting TFP in a regression exercise, the estimation is subject to the omitted variable problem. The omitted variable problem is likely to be more severe in cross-country studies with heterogenous institutions. The solution to include country fixed effects has the consequence of generating a Hurwicz bias overstating the rate of convergence. It follows that Barro (2012) does not include country fixed effects. Gennaioli et al. (2014) argue that the omitted variable problem is less severe at regional level given the homogeneity of institutions within countries. They therefore estimate the convergence equation with country fixed effects but not with regional fixed effects. The basic specification is a

[11] See discussion in Section2.

[12] The unlit pixel share is measured as the number of pixel with zero light but non zero population divided by the number of pixel with non zero population. The data for population at pixel level is from landscan.https://landscan.ornl.gov/

cross country regression in which the bias maybe strong. Including a country fixed effects raises the coefficient from -0.18 to -0.2 (column 3). Including a regional fixed, the coefficient increases further to -0.24.[13]

In order to control for administrative entity fixed effects, we can focus our analysis on the subsample of three countries (Bénin, Malawi, and Mali) that can be observed across two time units (1990-2000 and 2000-2010).[14] Table 1 column 6 reports the estimation results with both time and administrative entity (AE) fixed effects and has to be compared with column 5 with time fixed effects only. With both time and administrative entity fixed effects, the coefficient is 4 times the coefficient without fixed effects. This seems to suggest that disaggregate administrative entities fixed effects tend to bias the speed of convergence upward as discussed in the literature. It follows that we choose to include only country fixed effects in the following sections in line with Gennaioli et al. (2014).

These results are robust to various measures of sum of lights (appendix A), different measures of convergence (appendix B), excluding or including administrative entities with zero lights (appendix C).

4 Heterogenous convergence and sector specialization

Section 2 touches briefly upon the heterogeneity in the convergence process. Below we plot the normal frequency distribution for the growth rate in lights per employment for two categories: administrative entities above or below the median labor productivity level at $t-k$ (Figure 3). For the group with initial low lights per employment, lights per employment growth is on average higher as well as with higher standard deviation than for the group with initial high lights per employment level (av=8% and sd=0.2 in the low group and av=0.5% and sd=0.08 in the high group). In particular, the high dispersion in the group with initial low lights per employment level is an indication that convergence is heterogenous. Visual inspection shows that a non-negligible number of administrative entities with initial low lights per employment level experience small or negative growth.

This raises the question of the determinants of convergence. A central result in the existing literature on structural transformation is that the central role played by the manufacturing for productivity growth. This result is particularly well documented for African countries. The main explanation is that there is an important productivity gap between manufacturing and non-manufacturing sectors in low income countries. It follows that growth of non-manufacturing sectors does not contribute to productivity growth. In sub-Saharan Africa, the agriculture sector contains largely subsistence activities and services are dominated by low productive activities while modern services remain small. In addition, the size of manufacturing is limited in Africa and cannot absorb enough labor to generate

[13] The regression with fixed effects is : $\hat{\omega}^s_{i,t,t-k} = \beta^p_1 \log\left(\omega^s_{i,t-k}\right) + D_c + D_r + \varepsilon_{i,t,t-k}$ where D_c represents the country fixed effects and D_r represents the regional fixed effects.

[14] $\hat{\omega}^s_{i,t,t-k} = \beta^p_1 \log\left(\omega^s_{i,t-k}\right) + D_i + D_t + \varepsilon_{i,t,t-k}$, where i and j denote, respectively, the administrative entity and country indices, D_i represents the administrative entity fixed effects, and D_t is the time fixed effects.

Figure 3: Distribution - lights per employment growth

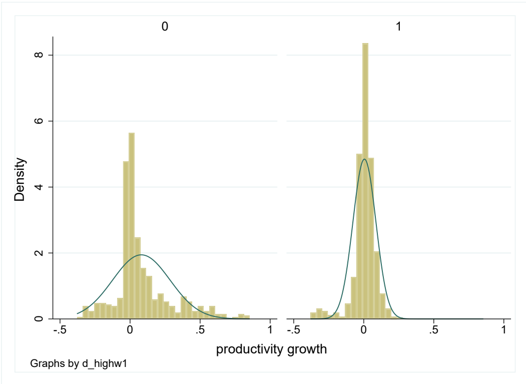


Figure (a) illustrates the density distribution of growth for low (0) and high (1) initial lights per employment level.

aggregate convergence. Rodrik (2012) shows that convergence is conditional on the initial employment share in manufacturing using country level national account data. In addition, McMillan et al. (2014) have documented negative structural change in Nigeria and Zambia, with manufacturing experiencing a declining employment share between 1990 and 2005. Using an annual panel of 11 African countries between 1960 and 2010, Vries et al. (2015) point to the fact that structural transformation has contributed positively to productivity growth. However, they underline that labor has been moving to sectors with above average productivity level but declining productivity growth.

In table 2, we investigate whether the initial employment allocation between sectors has impacted the growth in lights per employment in the subsequent decade. Thus, we consider four broad sectors: agriculture, mining, manufacturing, and the service sector. Note that we leave aside the public sector. We calculate the share of the workforce employed in each of these sectors $\phi^h_{i,t} = \frac{n^h_{i,t}}{\sum_{h \in \mathcal{H}} n^h_{i,t}}$ with $\mathcal{H} \equiv \{agriculture, mining, manufacturing, services\}$. Then, we estimate the relationship between lights per employment and its growth in the subsequent decade (2000-2010), controlling for sectoral employment shares at the beginning of the period $\hat{\omega}^s_{i,t,t-k} = \beta_0 + \beta_1 \log\left(\omega^s_{i,t-k}\right) + \beta_2 \phi^h_{i,t-k} + \varepsilon_{i,t,t-k}$, where h indicates the sector.

Columns (1), (2), and (3) in Table 2 display the estimation results when controlling for the employment share of agriculture, manufacturing, and services, respectively.[15] The coefficient for convergence is not strongly impacted by the control variables, although the magnitude is slightly larger. However, the sectoral shares enter with different signs.

The initial employment share of agriculture has a negative effect on the annual growth rate of lights per employment in the subsequent decade. Hence, administrative entities with initially higher

[15] We do not adjust for the employment share in the mining sector because this sector is not operating in many administrative entities.

Table 2: Economic structure and convergence

	(1)	(2)	(3)	(4)	(5)	(6)
			lights per employment growth			
log lights per employment t-k	-0.026***	-0.023***	-0.025***	-0.025***	-0.023***	-0.025***
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
Et share agri t-k	-0.104***			-0.126***		
	(0.019)			(0.024)		
Et share manufacturing t-k		0.247***			0.215***	
		(0.068)			(0.081)	
Et share services t-k			0.133***			0.149***
			(0.024)			(0.028)
unlit pixel share t-k				0.026*	-0.014	0.016
				(0.015)	(0.015)	(0.014)
R^2	0.15	0.14	0.15	0.15	0.14	0.15
N	1136	1136	1136	1136	1136	1136
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the results of the convergence regression controlling for initial sectoral employment shares.

employment share of agriculture are more likely to have lower annual growth in lights per employment in the subsequent decade. This result points that convergence does not take place primarily in rural area.[16]

In addition, the manufacturing employment share has a strong and positive impact on lights per employment growth. Existing results for Africa based on aggregated data show a negative contribution of employment share in the 1990s and a slightly positive contribution in the 2000s. This result confirms existing results in the literature. The result is valuable given the widespread stagnation of manufacturing employment share at the aggregate level. Further, while manufacturing employment may stagnate at the aggregate level, it does not exclude the possibility of manufacturing growth at the local level through a rellocation/concentration effect.

In turn, the initial employment share of local services shows positive effects on the annualized growth rate of lights per employment in the subsequent decade. Administrative entities with initially higher employment share of local services are more likely to experience higher growth in lights per employment in the subsequent decade. This result is also interesting as services are often presented as low productive sectors in Africa and are ruled out as an engine for productivity growth. The coefficient is twice smaller than the coefficient associated with the initial employment share in manufacturing. These results are robust to the inclusion of the share of unlit pixels as a control variable (see columns (4) to (6) of Table 2). The coefficient for unlit pixel is not significant in contrast with the baseline regression.[17]

We make use of the information contained in the census data to explore the relative importance of subsectors' employment shares (Figure3). Given the small share of manufacturing employment, this category is divided into 5 subcategories (food and beverages, textile, others manufacturing, electricity and water, and construction). We also report the results for 4 service categories (wholesale, hotel, transport, finance).[18] We find here that the coefficients reflect the productivity of the subsector, with for instance food and beverage associated with a coefficient of 0.37, textile with a coefficient of 0.67 and construction with a coefficient of 0.67. Interestingly, the service sector is composed of both high productive and low productive sectors such as transport for the former and wholesale for the latter. The difference of coefficient between manufacturing and high productivity services is also small with some high productivity services producing a larger coefficient than manufacturing. This is also sometimes labeled as "wrong" manufacturing as these manufacturing sectors can be linked to sectors with high informality. These sectors are also producing for the domestic market and are therefore not very different from nontradable sectors (Osei and Jedwab, 2017). These subsector coefficients

[16]This result differs from the result presented in Chanda and Kabiraj (2020) for India. Note that the paper on convergence in India uses a different method to differentiate between urban and rural areas based on low versus high luminosity, and are not using a direct measure of sectoral specialization as in this paper.

[17]These results are not altered significantly by a modification of the manufacturing versus services categories that would for instance be aligned with a tradable versus nontradable classification.

[18]The lack of sectoral diversity in low income countries prevents us from further disaggregation.

Table 3: Economic sub-structure and convergence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	lights per employment growth								
log lights per employment t-k	-0.023***	-0.023***	-0.023***	-0.023***	-0.024***	-0.025***	-0.023***	-0.024***	-0.023***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
food, beverage t-k	0.377								
	(0.264)								
textile t-k		0.670***							
		(0.179)							
others beverage t-k			-0.089						
			(0.124)						
electricity, water t-k				0.345					
				(0.411)					
construction t-k					0.676***				
					(0.230)				
wholesale t-k						0.269***			
						(0.055)			
hotel t-k							-0.005		
							(0.308)		
transport t-k								0.782***	
								(0.190)	
finance t-k									-0.796**
									(0.396)
unlit pixel share t-k	-0.040***	-0.027**	-0.048***	-0.041***	-0.016	0.015	-0.044***	-0.004	-0.053***
	(0.012)	(0.012)	(0.014)	(0.012)	(0.013)	(0.015)	(0.015)	(0.013)	(0.014)
R^2	0.14	0.14	0.14	0.14	0.14	0.15	0.14	0.14	0.14
N	1136	1136	1136	1136	1136	1136	1136	1136	1136
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the results of the convergence regression controlling for initial sub-sectoral employment shares.

are consistent with the aggregate coefficients for manufacturing and services to the extent that the aggregate coefficients also reflect the relative size of the subsectors.[19]

In this section we highlight the importance of manufacturing employment to explain conditional convergence in line with existing studies using national account data. In addition, we also show that services can contribute positively to convergence as it gathers sectors that are heterogenous in terms of productivity. In light of these results, the combined contribution of manufacturing and service sectors may explain the global convergence highlighted in the previous section.

5 Geographic - natural characteristics and convergence regimes

In this section, we discuss additional variables that could explain this heterogenous convergence. We first look at potential candidate factors to account for conditional convergence in a linear cross section estimation. We then test whether these variables are robust to a threshold estimation to capture the non-linearities that could explain the absence of convergence in certain geographic entities.

Below, we explore the importance of different factors in explaining the convergence across administrative entities. The impact on the convergence coefficient of the different set of variables is moderate but the coefficient is higher compared to the baseline regression. We include variables related to population and sectoral specialization, variables that capture geographic location and natural characteristics in the spirit of Henderson et al. (2017) and Chanda and Kabiraj (2020). A first set of variables include sectoral employment in the last period, population density in the last period, as well as the number of deaths from conflicts within the administrative entity.[20] The coefficients associated with either sectoral employment in agriculture, manufacturing or services is similar to the sign in Table 2, highlighting further the robustness of our results.

The inclusion of population density is motivated by the fast rate of urbanization in Africa. Population density is also important as a result of the concentration of the modern sector in a handful of large cities in Africa. We expect convergence to take place between secondary cities and the main cities as well as between rural areas and urban centers. The sign is negative and significant. This indicates that convergence has taken place either in rural areas or secondary cities. Given the negative sign of agriculture related areas, this could indicate that convergence occurred in secondary cities. We also include a variable for conflicts that measures deaths (in thousands) taking place within the administrative entity. The data on deaths related to armed conflicts is from the Uppsala Conflict Data Program - Uppsala Georeferenced Event Dataset (Sundberg and Melander, 2013; Pettersson and Öberg, 2020).[21] The database compiles every episode of armed conflicts in a given location. Armed

[19] The coefficient associated with finance is uninformative given the concentration of employment in a handful of administrative entities.

[20] Adhikari and Dhital (2021) tests for decentralization and convergence in 69 countries but this cannot be done for this set of countries.

[21] `https://ucdp.uu.se/downloads/index.html#ged_global`.

conflicts are defined as an event during which an organized group used forces against another armed group or against civilians generating at least 1 death at a given date and in a given place. We aggregate the number of death in each administrative entity for the period between the two census. The decade 2000s has witnessed a relative stability for the set of countries in this study. For instance, it is only after 2010 that conflicts arise in Mali or Bénin. However, Côte d'Ivoire went through ten years of civil war that led to significant casualties. The coefficient associated with conflicts related death is negative and significant. This result echoes the literature that find a negative impact of conflicts on economic prosperity (International Monetary Fund, 2019; Lopes and Baskaran, 2015).

A second set of explanatory variables intend to capture the geographic location of the administrative entities including distance to the main city and distance to the coast.[22] These variables capture accessibility of an administrative entity to the domestic market/market size as well as international market. We find that the further away from the main city the lower the rate of convergence.[23] This result is in line with Storeygard (2016) in regard to the impact of transport costs on the income of cities in Africa. The literature also points to the positive impact of access to the coast as in Rappaport and Sachs (2003). In the regression, the coefficient is not significant. Note however that if the metric is replace by a dummy for administrative entities located on the coast, the coefficient is negative and significant. In addition, the threshold regression further below points to the positive impact of proximity to the coast.

A third category of control variables captures the natural characteristics of the administrative entity using terrain's ruggedness, land suitability, precipitation, temperature and malaria.[24] Land suitability is taken from Ramankutty et al. (2002). It is an index combining measures of growing season length, moisture availability to crops, soil carbon density and soil pH. This variable captures the probability that a cell is cultivated. Two other agriculture related variables include average precipitation and average temperature. All three agriculture variables appear with a negative sign. However, only land suitability is significant across the different specifications. In the existing literature, the sign differs in relation with the scope of the study and the geographic coverage. In global studies, a positive sign associated with these variables is sometimes interpreted as consumption amenities. However, this might be the case when the country coverage of the study includes temperate climate. Another common interpretation is the potential positive agriculture productivity. In sub-Saharan Africa relatively higher temperature or precipitation means either semi-arid areas or tropical forest areas. In addition, in sub-Saharan Africa, agriculture remains largely a subsistence activity and there is no consensus that a positive rural push is at work in this region. The negative sign of these

[22]These two variables are constructed by taking the distance from the centroid of the administrative entity in (100) km.

[23]The list of main cities are Bamako, Abidjan, Accra, Cotonou, Kigali, Daressalaam, Blantyre, Maputo, Lusaka and Johannesburg.

[24]Ruggedness is taken from Nunn and Puga (2012) and includes a latitude adjustment effect. Agricultural land suitability is taken from the Atlas of the Biosphere https://nelson.wisc.edu/sage. Rainfall and temperature are taken from https://www.worldclim.org/.

variables is consistent with the discussion of the importance of sector specialization in section 4 that highlighted the negative impact of agriculture specialization on convergence. The set of results for natural characteristics is similar to that of Chanda and Kabiraj (2020).

To control for natural characteristics, we also include ruggedness in the regression. The contemporaneous effect of ruggedness is expected to be negative as ruggedness increases trade costs, production costs in agriculture and building costs. Nunn and Puga (2012) show that ruggedness has a general negative impact on the level of income per capita. However, in Africa, ruggedness positively impacted on GDP level as historically the slave trade was more difficult to conduct in rugged terrain. In the regression below, ruggedness is negative but not significant. The difference with the result from Nunn and Puga (2012) is related in our analysis to the short term (10 years) growth dimension of our analysis that is impacted negatively by increased costs, while Nunn and Puga (2012) looked at the very long term impact of ruggedness on GDP level in interaction with the slave trade. Henderson et al. (2017) also find a negative coefficient associated with ruggedness but on the level of light (per area). The coefficient is not significant in our study may be because the size of the administrative entities is rather large and may reduce the specificity of the measure. For instance, Henderson et al. (2017) aggregate lights at the level of 900 pixels while we use the second level of administrative entities. We include a variable for susceptibility to tropical disease as measured by the prevalence of Malaria to measure the potential negative impact of tropical disease on lights per employment.[25] The coefficient is not significant and does not impact the estimated coefficient for convergence. The coefficient associated with geographic and natural characteristics is robust to the measure of sectoral employment specified. These results are also robustness to a measure of light with no latitude adjustment as shown in Appendix C.

From the regression above, it appears that the most important factors explaining conditional convergence are sectoral specialization, the density of the city, distance to the main city, and land suitability. In order to capture the heterogeneity in the convergence of administrative entities, we therefore want to examine whether the coefficient associated with convergence changes magnitude when interacting with some of the variables identified above. The objective is to gain a further indication that administrative entities which experienced convergence are administrative entities with relatively good access to market, with small population density and with a smaller prevalence of agriculture.

In order to do so, we perform threshold regression. A threshold regression is a spline regression where the cut-off point minimizes the sum of square errors (Hansen, 1999). The data are then divided into two groups and two regimes are estimated. The estimated equation is now :

$$\hat{\omega}^s_{i,t,t-k} = \beta_0 + \beta_1 \log\left(\omega^s_{i,t-k}\right)\Gamma\left(x_{i,t-k} \leq q\right) + \beta_2 \log\left(\omega^s_{i,t-k}\right)\Gamma\left(x_{i,t-k} > q\right) + \varepsilon_{i,t,t-k} \tag{3}$$

[25] The Malaria index is taken from Kiszewski et al. (2004).

Table 4: Conditional convergence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	lights per employment growth								
log lights per employment t-k	-0.025***	-0.020***	-0.022***	-0.027***	-0.028***	-0.025***	-0.026***	-0.027***	-0.028***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
unlit pixel share t-k	0.022*	-0.048***	-0.055***	0.020	0.007	-0.009	-0.022	0.007	0.000
	(0.013)	(0.011)	(0.012)	(0.016)	(0.017)	(0.016)	(0.017)	(0.015)	(0.015)
Et share agri t-k	-0.148***			-0.141***	-0.127***				
	(0.021)			(0.024)	(0.026)				
Et share manufacturing t-k						0.350***	0.244***		
						(0.082)	(0.084)		
Et share services t-k								0.177***	0.165***
								(0.030)	(0.030)
density t-k	-0.074***			-0.088***	-0.114***	-0.069***	-0.088***	-0.100***	-0.125***
	(0.017)			(0.017)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)
deaths t-k	-0.081*			-0.092*	-0.067	-0.067	-0.039	-0.100*	-0.070
	(0.046)			(0.050)	(0.052)	(0.042)	(0.044)	(0.053)	(0.054)
distance to main city		-0.002*		-0.003**	-0.006***	-0.004***	-0.006***	-0.003**	-0.006***
		(0.001)		(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
distance to coast		0.000		0.001	-0.002	0.000	-0.001	0.001	-0.001
		(0.002)		(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
land suitability			-0.066***	-0.065***	-0.050**	-0.071***	-0.053**	-0.068***	-0.051**
			(0.022)	(0.023)	(0.025)	(0.023)	(0.025)	(0.023)	(0.025)
ruggedness			-0.006	-0.009	-0.010	-0.010*	-0.011	-0.010	-0.011
			(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
precipitation			-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
temperature			-0.006**	-0.003	-0.006*	-0.005**	-0.006**	-0.003	-0.005*
			(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Malaria index			0.001	-0.000	-0.000	0.000	-0.000	-0.000	-0.000
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R^2	0.14	0.12	0.13	0.15	0.17	0.14	0.17	0.15	0.17
N	1136	1136	1136	1136	1136	1136	1136	1136	1136
County FE	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the results of the baseline specification controlling for variables measuring distances, natural characteristics, density and sector specialization.

Table 5: Convergence regimes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			lights per employment growth				
	Et share agri t-k	Et share manu t-k	Et share services t-k	density	Main city	Coast	Ruggedness
q	.952	.008	.026	0.612	.203	.625	.652
log lights per employment t-k$\leq$ q	-.020	-.024	-.042	-.024	-.087	-.034	-.024
.95 CI	[-.025,-.015]	[-.043,-.009]	[-.060,-.025]	[-.030,-.015]	[-.123,.0007]	[-.052,-.016]	[-.080,-.014]
log lights per employment t-k $>$ q	-.047	-.020	-.022	-.012	-.017	-.016	-.011
.95 CI	[-.061,-.034]	[-.027,-.014]	[-.026,-.017]	[-.020,-.008]	[-.091,.021]	[-.020,-.011]	[-.021,-.007]
.95 CI	[.952,.952]	[.005,.051]	[.026,.026]	[.356,4.612]	[.203,15.8]	[.428,2.106]	[.005,5.26]
N $\leq$ q	959	88	137	500	40	187	730
N	1136	1136	1136	1136	1136	1136	1136

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the results of the threshold specification for different thresholds q identified in the previous regression: distance to coast, ruggedness, density lagged and employment share in sectors lagged.

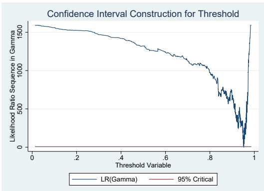
with β. the coefficient for each regime, $x_{i,t-k}$ the threshold variables and q the threshold parameter. This approach shares similarities with Henderson et al. (2017) in the sense that the emphasize is set on multiple regimes. Based on Durlauf and Johnson (1995), they split their sample into early and late developers, and include an interaction dummy between the group and the explanatory variables. We go a step further as threshold regressions estimate two separate coefficients for each of the regime.

A first result is the absence of a regime associated with the sectoral shares pointing to a linear effect. The likelihood ratio test reject the threshold for both agriculture and services. For manufacturing, the threshold is low with less than 100 data points below the threshold and the coefficients are similar under both groups. Note, however, that there could be another threshold around 5-6% (see Figure 4).

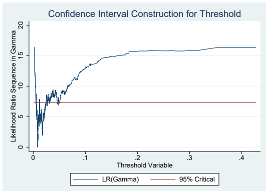
Contrastingly geographic and natural characteristics are associated with convergence regimes. Population density impacts the convergence rate with a threshold at 0.61. Both regimes have very similar size, with the low regime gathering around 500 observations. The rate of convergence is 2.4% and significant in areas with low population density and 1.2% in high population density. This result is intuitive as the more densely populated areas are also the more urbanized areas. Distance to the main city delivers a threshold at 0.2 (20km from the main city). In addition to the small distance implied by the threshold, none of the coefficients are significant and all the observations falls above the threshold with only 40 observations below the threshold. Looking at the likelihood ratio test, there seem to be potentially three additional thresholds at 400km, 1000km and 1400km. This indicates that a double threshold estimation or a smooth transition estimation might be more suited for the analysis. Interestingly, the variable distance becomes significant in the threshold regression. The threshold is at 60 km from the coast. The confidence interval at 95% comprises distance ranging from 40km to 200km. Administrative entities below this cut-off point converge at 3.4% while administrative entities beyond this cut-off point converge at 1.6%. The number of administrative entities below this cut-off point is mechanically smaller (16% of the sample). The threshold regression also points to two regimes associated with ruggedness. The two regimes are distinct by a value of ruggedness equal to 0.65.[26] The sample is split 65% below the threshold, with convergence being 2.4% for administrative entities relatively flat and 1.1% in more rugged terrains. The likelihood ratio indicates a threshold at 0.6 as well as possible alternative thresholds above 3.

From this exercise, it appears that the impact associated with good or bad geographic and or natural characteristics may imply a convergence rate that is 2 to 3 times more important than with poor characteristics. This analysis should be complemented with double threshold estimation or smooth transition estimation as well.

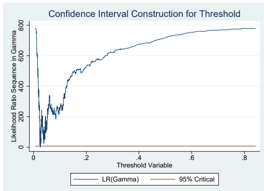
Figure 4: Likelihood ratio test for the thresholds



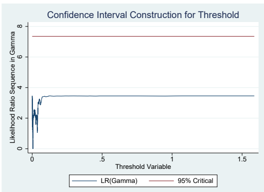
(a) Agriculture Et share



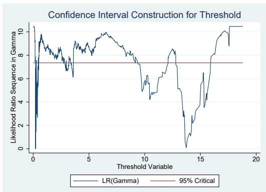
(b) Manufacturing Et share



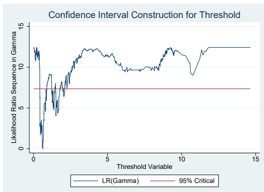
(c) Service Et share



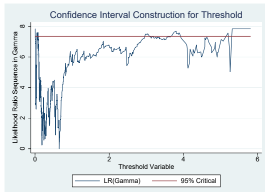
(d) Population density



(e) Distance main city



(f) Distance coast



(g) Ruggedness

This figure shows the likelihood ratio test comparing the sum of square errors of the model with the estimated threshold $\hat{q}$ and the model with alternative threshold q.

6 Decomposition across administrative entities

Convergence in labor productivity also raises the issue of the source of convergence and whether convergence is related to productivity growth within sectors or labor relocation between sectors. For instance, McMillan et al. (2014) find that structural change in Africa is growth reducing over the 1990s and growth enhancing in the first half of the 2000s, with both effects balancing each other overall. Similarly, Vries et al. (2015) note that labor reallocated towards sectors with relatively higher productivity but declining productivity growth. In this section, we ask a similar question

[26] Ruggedness is taken from Nunn and Puga (2012). Ruggedness is adjusted for latitude only. Sometimes latitude adjustment is multiplied by the undajusted cell areas as measured in the original raster file.

but from a geographic point of view. The decomposition determines the respective contribution of within administrative entities effect and the between administrative entities effect. In other words, we decompose our proxy of aggregate productivity growth into growth at the level of the administrative entity and labor movement between administrative entities.

$$\dot{\omega}^s_{t,t-k} = \sum_{i=1}^{n} \theta_{i,t-k}\dot{\omega}^s_{i,t,t-k} + \sum_{i=1}^{n} \dot{\theta}_{i,t,t-k}\omega^s_{i,t} \tag{4}$$

with ω aggregate lights per employment, ω_i local lights per employment, θ_i the share of local employment in aggregate employment and $\dot{x}$ denotes the change of variable x between t and $t-k$.

The decomposition performed here is not a direct observation of sectoral recomposition as made in the existing literature. However, when discussing the within and between effects the literature usually assumes labor movements between rural areas and urban areas. Rural push and urban pull theories of structural change identify whether transformation is initiated via a rise of productivity in agriculture in rural areas or a rise of productivity in manufacturing in urban areas. In the African context, these labor movements across administrative entities may be associated with negative push and pull effects, such as consumption of resources rent, biased urban policies, rural poverty and fast increase in labor supply (Gollin et al., 2016). The decomposition performed here complements existing sectoral decomposition by stressing the changing size of administrative entities and the movement of labor associated with these changes.

In Figure 5, we show that the within sector component is positive. The within sector also dominates the between sector component except in Malawi and Mali. This result indicates that lights per employment growth has taken place across both rural and urban administrative entities. This confirms that there are at play positive rural push and positive urban pull factors. Structural transformation is also positive indicating that labor movements across administrative entities have taken place from administrative entities with low lights per employment to administrative entities with high lights per employment. The magnitude of the effect seems however small. Despite the large movements of population between rural and urban areas and the fast growth of cities in the countries considered, the contribution of labor movement to lights per employment growth is small. This indicates that there are also negative push and pull factors at play.

Three countries display negative structural transformation effects: Côte d'Ivoire, Ghana and South Africa. Population has been moving from relatively high productive areas to low productive areas. This can be explained by the post-electoral crisis that covers most of the 2000s decade in Côte d'Ivoire. In South Africa, the post apartheid industrial reorganization may also explain this pattern as many industries have disappeared. The apartheid period is associated with high tariff and a pick in manufacturing employment.[27] The result for Ghana is consistent with existing results. Osei and

[27] See Diao et al. (2017).

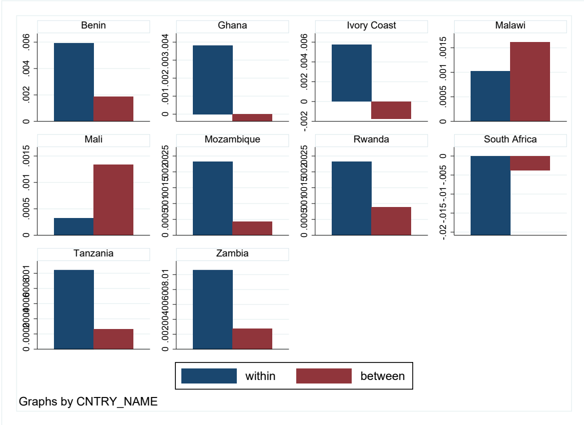


Figure 5: Decomposition, by Country.

Jedwab (2017) find, using national account data, that structural transformation is very volatile and in particular that structural transformation is negative over the period 2000-2006 and positive over the period 2006-2010. We find here a small negative effect.

Earlier, this paper stressed that contrary to studies using aggregate national account, a disaggregated approach finds evidence of convergence across sectors. For those three countries, movements of workers from areas with high productivity to areas with lower productivity could also explain the absence of convergence in studies using aggregate national accounts. The weight of the administrative entities that display convergence declines.

7 Conclusion

In this paper, we have tested for convergence at the local level for 1136 administrative entities representing 10 countries in sub-Saharan countries. We do so by combining a unique set of local measure of employment taken from population and housing census with nighttime lights data to produce a proxy for labor productivity. The period of analysis is centred on the period 2000-2010 but also covers the early 1990s for the countries for which there are three consecutive census. The objective is to give a new perspective to the literature on structural transformation in Africa that relied mostly on national account disaggregated at sectoral level and focused on the rise (or fall) of the manufacturing sector. Our paper is close to the handful of papers using nighttime light to study convergence. One important distinction, however, is that local sectoral employment shares constructed from the census data enable us to maintain the sectoral dimension central to the literature on structural transformation in Africa. This paper also distinguishes itself by performing productivity decomposition that touches upon the issue of rural urban migration. Another value added is to identify the economic, geographic and natural characteristics of the areas converging.

A first finding of this paper is the evidence of convergence across sectors around 2%. This result is in line with the (regional) growth literature but new to papers focusing on Africa and finding convergence in the manufacturing sector only. This result is robust to different measures of sum of lights as discussed in the appendix.

A second finding of this paper is that convergence is heterogeneous across administrative entities. Only a fraction of administrative entities with low initial lights per employment level experienced lights per employment growth. Discussing the relevance of sector composition for structural transformation, this paper highlights the importance of shifting away from agriculture to display fast convergence rate. The paper confirms the central role played by manufacturing activities as well as services for convergence. The manufacturing effect dominates the services effects, but some subsectors within services display relatively high productivity level.

Identifying additional factors explaining convergence, we bring evidence that convergence is conditional on certain characteristics such as sector specialization, proximity to the main city, relatively low population density, land suitability and moderate temperature. These variables emphasize the importance of trade cost, production cost and sector specialization for convergence. In this perspective, this analysis complements existing papers such as Henderson et al. (2017). However, we go one-step further and estimate convergence regimes using threshold regression. Evidence supporting convergence regimes is strong and suggests that in the high convergence regime, convergence rates can be 2 to 3 times larger than in the low convergence regime.

An important dimension of convergence is the rural-urban migration that is embedded in the push and pull theories of structural change. In this paper, we touch upon this issue by performing a decomposition of lights per employment into within administrative entities growth, and labor movement between administrative entities. We show that lights per employment growth is explained mostly by the within component, but that in most cases the between component is positive too. We leave aside the issue of sigma convergence but the heterogeneity in convergence implies that some areas are left behind the convergence process.

References

Adhikari, B. and S. Dhital (2021). Decentralization and regional convergence: Evidence from nighttime lights data. *Economic Inquiry 59*(3), 1066–1088.

Barro, R. J. (2012, August). Convergence and modernization revisited. Working Paper 18295, National Bureau of Economic Research.

Baugh, K., C. Elvidge, D. Ghosh, and D. Ziskin (2010). Development of a 2009 stable lights product using dmsp-ols data. *Proceedings of the Asia-Pacific Advanced Network 30*(0), 114.

Carrington, S. J. and P. Jiménez-Ayora (2021). Shedding light on the convergence debate: Using luminosity data to investigate economic convergence in ecuador. *Review of Development Economics 25*(1), 200–227.

Chanda, A. and S. Kabiraj (2020). Shedding light on regional growth and convergence in india. *World Development 133*.

Diao, X., M. McMillan, and S. Wangwe (2017, 12). Agricultural Labour Productivity and Industrialisation: Lessons for Africa. *Journal of African Economies 27*(1), 28–65.

Durlauf, S. N. and P. A. Johnson (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics 10*(4), 365–384.

Elvidge, C., K. Baugh, K. Kihn, E. Kroehl, and E. Davis (1997). Mapping city lights with nighttime data from the dmsp operational linescan system. *Photogrammetric Engineering and Remote Sensing 63*(6), 727–734.

Fafchamps, M., M. Koelle, and F. Shilpi (2016, 06). Gold mining and proto-urbanization: recent evidence from Ghana. *Journal of Economic Geography 17*(5), 975–1008.

Gennaioli, N., L. P. Rafael, L. D. S. Florencio, and S. Andrei (2014, 09). Growth in regions. *Journal of Economic Growth 19*(3), 259–309.

Gollin, D., V. Dietrich, and R. Jedwab (2016). Urbanization with and without industrialization. *Journal of Economic Growth 21(1)*, 35–70.

Hansen, B. E. (1999, December). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics 93*(2), 345–368.

Henderson, J. V., T. Squires, A. Storeygard, and D. Weil (2017, 09). The Global Distribution of Economic Activity: Nature, History, and the Role of Trade1. *The Quarterly Journal of Economics 133*(1), 357–406.

Henderson, J. V., A. Storeygard, and D. N. Weil (2012, April). Measuring economic growth from outer space. *American Economic Review 102*(2), 994–1028.

Iddawela, Y., N. Lee, and A. Rodríguez-Pose (2021). Quality of sub-national government and regional development in africa. *The Journal of Development Studies 57*(8), 1282–1302.

International Monetary Fund (2019). Regional economic outlook. sub-saharan africa : recovery amid elevated uncertainty. *Regional Economic Outlook: Sub-Saharan Africa*.

IPUMS (2019). Minnesota population center. integrated public use microdata series, international: Version 7.2 [dataset]. minneapolis, mn: Ipums, 2019.

Jedwab, R. and A. Moradi (2016). The permanent effects of transportation revolutions in poor countries: Evidence from africa. *The Review of Economics and Statistics 98(2)*, 268–284.

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs (2004). A global index representing the stability of malaria transmission. *The American Journal of Tropical Medicine and Hygiene 70*(5), 486–498.

Lessmann, C. and A. Seidel (2017). Regional inequality, convergence, and its determinants – a view from outer space. *European Economic Review 92*, 110–132.

Lopes, d. M. and T. Baskaran (2015). Re-evaluating the economic costs of conflicts. *CEGE Discussion Papers 246*.

Martin, R., P. Sunley, B. Gardiner, E. Evenhuis, and P. Tyler (2018). The city dimension of the productivity growth puzzle: the relative role of structural change and within-sector slowdown. *Journal of Economic Geography 18*(3), 539–570.

McMillan, M., D. Rodrik, and Íñigo Verduzco-Gallo (2014). Globalization, structural change, and productivity growth, with an update on africa. *World Development 63*, 11 – 32. Economic Transformation in Africa.

McMillan, M. S., D. Rodrik, and C. Sepulveda (2017). *Structural change, fundamentals, and growth: A framework and case studies.* International Food Policy Research Institute.

Nunn, N. and D. Puga (2012, February). Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics 94*(1), 20–36.

Osei, R. D. and R. Jedwab (2017). *Structural Change in a Poor African Country: New Historical Evidence from Ghana*, Chapter 4, pp. 184–219. International Food Policy Research Institute.

Pettersson, T. and M. Öberg (2020). Organized violence, 1989–2019. *Journal of Peace Research 57*(4), 597–613.

Ramankutty, N., J. A. Foley, J. Norman, and K. McSweeney (2002). The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography 11*(5), 377–392.

Rappaport, J. and J. D. Sachs (2003, March). The United States as a Coastal Nation. *Journal of Economic Growth 8*(1), 5–46.

Rodrik, D. (2012, 11). Unconditional Convergence in Manufacturing. *The Quarterly Journal of Economics 128*(1), 165–204.

Storeygard, A. (2016, 04). Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa. *The Review of Economic Studies 83*(3), 1263–1295.

Sundberg, R. and E. Melander (2013). Introducing the ucdp georeferenced event dataset. *Journal of Peace Research 50*(4), 523–532.

Vries, G. D., M. Timmer, and K. D. Vries (2015). Structural transformation in africa: Static gains, dynamic losses. *The Journal of Development Studies 51*(6), 674–688.

Table 6: Robustness check: Different measures of sum of lights

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	lights per employment growth Cross section								
	no latitude adjustment			alternative transformation $max(1, y^s_{i,j,t})$			average $y^s_{i,j,t} = (y^s_{i,j,t-1} + y^s_{i,j,t} + y^s_{i,j,t+1})/3$		
log lights per employment t-k	-0.017***	-0.018***	-0.017***	-0.018***	-0.020***	-0.020***	-0.021***	-0.024***	-0.022***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
unlit pixel share t-k		-0.030***			-0.052***			-0.057***	
		(0.009)			(0.010)			(0.010)	
R^2	0.10	0.10	0.13	0.11	0.11	0.14	0.15	0.16	0.21
N	1136	1136	1136	1136	1136	1136	1136	1136	1136
Country FE	No	No	Yes	No	No	Yes	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the results of the baseline specification for different measure of the sum of lights: sum of lights with no latitude adjustment (Columns 1 to 3), $max(1, y^s_{i,j,t})$ (Columns 4 to 6), $y^s_{i,j,t} = (y^s_{i,j,t-1} + y^s_{i,j,t} + y^s_{i,j,t+1})/3$ (Columns 7 to 9).

A Robustness check 1: alternative measure of sum of lights

In this appendix we perform some robustness check. A first robustness check is to measure the coefficient for convergence under the baseline regression for different measure of nighttime lights. In Table 6, we show the baseline regression for alternative sum of lights. For each measure we display three regression, without FE, with the unlit pixel share as a control variable and with fixed effects. The alternative measures include sum of lights with no latitude adjustment (column 1 to 3). Another alternative measure modifies the transformation $(1 + y^s_{i,j,t})$ with $max(1, y^s_{i,j,t})$ in columns 4 to 6. A last alternative measure is to take a three years average sum of lights centered around the year corresponding to the census year $y^s_{i,j,t} = (y^s_{i,j,t-1} + y^s_{i,j,t} + y^s_{i,j,t+1})/3$. The motivation for this last alternative is that DMSP covers 6 different satellites that overlap except for the last two satellites. The aggregate series are constructed by taking the average across satellite. In order to smooth the potential gap for the last two series we that a three years average. The baseline results are robust to the different measure of sum of lights. There are no difference only slightly higher coefficient for the three years average.

B Robustness check 2: different measures of convergence

In this appendix, we discuss whether the convergence changes magnitude whether it is based on night light per capita versus productivity (night light per employment) versus employment per capita (see Table 7). The motivation is that macroeconomic studies using sectoral data rely on productivity

Table 7: Robustness check: Unconditional convergence for different measures of convergence

	(1) change lights per Et	(2) change Et pc	(3) change lights pc
log lights per Et t-k	-0.018***		
	(0.002)		
log Et pc t-k		-0.049***	
		(0.003)	
log lights pc t-k			-0.016***
			(0.002)
R^2	0.11	0.35	0.08
N	1136	1136	1136
Country FE	No	No	No

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the results of the baseline specification for different measures of convergence, light per employment, employment per capita and light per capita.

measures whereas geographic studies rely on night light per capita. This paper relies on a proxy for labor productivity. In sub-saharan Africa employment is very likely to be highly correlated with population given subsistence activities and the lack of replacement incomes.

In the table below, we show that convergence using nightlight per capita is very similar to convergence using productivity. The rate of convergence is 1.6% against 1.8% in the baseline calibration. The slight difference might be explained by the growth of population being faster than the growth of employment. However, the rate of convergence is twice larger at 4.9% when measured by employment to population ratio. This tends to indicate that convergence as proxied by light minimize the rate of convergence. The shortcoming of using employment rate is that in countries with no social safety nets and large share of informality, employment is less related to business cycle and more related to population growth.

C Robustness check 3: administrative entities with a sum of light of zero

The sample is made of 1136 administrative entities with a data point at the beginning of the 2000s and at the end of the 2010s. Some administrative entities have a zero sum of light either at one point in time or over the entire period. A zero sum of light could be indicative of a measurement error for area with low light as satellites have difficulties measuring low lights. This could be an issue has poor administrative entities might be going through either an increase or a decline in growth that

Table 8: Robustness check: Administrative entities with zero sum of lights

2000s	2010s	SOL (t)	Average SOL $(t-1,\, t,\, t+1)$
$SOL = 0$	$SOL > 0$	6%	7%
$SOL > 0$	$SOL = 0$	6%	5%
$SOL > 0$	$SOL > 0$	72%	76%
$SOL = 0$	$SOL = 0$	16%	12%

is not recorded by the satellite images. On the other end, a zero sum of light maybe indicative of administrative entities that are poor and are remaining poor.

In the table below, we are displaying the proportion of administrative entities across 4 categories (with positive or zero sum of light at both dates, with zero sum of lights at the beginning or at the end of the sample) and across two measures (the sum of light at the date of the census, the average sum of light for the years before and after the census year ($t-1$, t and $t+1$)).

Between 28% and 24% of administrative entities display zero lights at a point in time. Henderson et al. (2017) considers that sum of light of zero in areas with non zero population is a censoring issue. They assign the lowest observed light to the grid cell in order to reduce the gap between area with no light and area with the smallest nonzero values. Other studies exclude administrative entities with zero light at the beginning of the period as in Chanda and Kabiraj (2020). It might make sense to exclude certain areas that have zero light such as desert areas and forest areas. However, if the objective is to look at convergence, it might be important to include all areas. In particular, areas that display no light at a point in time and a non zero light at another point in time may capture areas converging or diverging. These two categories are non negligible as they account for more than 10% of the sample. Areas with zero lights across the entire sample account for 15% of the sample. Note that none of the administrative entities have zero population. Zero light might therefore be a censoring issue.

The inclusion or exclusion of different subgroups of administrative entities with zero light at some point in time has an impact on the magnitude of the convergence (see Table 9).[28] Convergence is 1.8% for the sample including all observations. However, excluding administrative entities with a zero sum of light both at the beginning and at the end of the period shifts the convergence rate to 2.9%. This is intuitive as this subcase excludes the administrative entities with zero growth in lights (Panel A column 2). However, when the subsample is restricted to administrative entities with non zero light

[28] As mentioned in the data section, in order to include observation with no light we apply the following transformation to the data $\omega^s_{i,j,t} = \frac{1+y^s_{i,j,t}}{n_{i,j,t}}$.

Table 9: Robustness check: Composition effect of administrative entities with zero light

	(1)	(2)	(3)	(4)	(5)
		lights per Et growth			
			panel A		
log lights per Et t-k	-0.018***	-0.029***	-0.031***	-0.034***	-0.024***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
R^2	0.11	0.21	0.26	0.31	0.13
N	1136	1001	868	944	925
Sample	(00,++,0+,+0)	(++,0+,+0)	(++)	(++,0+)	(++,+0)
Country FE	No	No	No	No	No
County FE	No	No	No	No	No
Admin FE	No	No	No	No	No
			panel B		
log productivity t-k	-0.020***	-0.033***	-0.035***	-0.036***	-0.028***
	(0.002)	(0.003)	(0.004)	(0.003)	(0.004)
L.unlit	-0.052***	-0.077***	-0.058***	-0.058***	-0.068***
	(0.010)	(0.012)	(0.012)	(0.011)	(0.012)
R^2	0.11	0.22	0.27	0.32	0.15
N	1136	1001	868	944	925
Sample	(00,++,0+,+0)	(++,0+,+0)	(++)	(++,0+)	(++,+0)
Country FE	No	No	No	No	No
County FE	No	No	No	No	No
Admin FE	No	No	No	No	No

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the baseline regression including or excluding administrative entities with zero lights at both dates (00), at the beginning of the period (0,+) or at the end of the period (+,0). Panel B includes the control for the share of pixel with zero light but non zero population.

at both dates, the convergence rate drops is high at 3.1%. This means that convergence measured in this paper is not driven by administrative entities with zero lights. The convergence rate is logically even higher (3.4%) when are added the administrative entities with zero light at the beginning of the sample and positive light at the end. The convergence is 2.4% when the sample gather administrative entities with positive lights at both dates and administrative entities with negative lights at the end of the sample. This results point to the importance of composition effect for the overall convergence rate. Some administrative entities have been left behind and are not part of the convergence process. Some administrative entities are diverging and others are converging quickly. In addition, this result shows that the speed of convergence rate across different set reflects the weights of the different types of trajectories (convergence vs divergence).

Panel B displays the same set of regression controlling for measurement errors related to the difficulty of satellites to measure pixel with low light. This variable measures the share of pixel with zero light but non zero population over the number of pixel with non zero population. The convergence rate is slightly larger across all sample composition but by a small magnitude (point one or point two decimal point). The coefficient is negative and significant. The larger the share with zero light but non zero population the smaller the convergence rate. The size of the coefficient across the different subsample reflects the weight of administrative entities with zero lights at the beginning of the period.

D Robustness check 4: regression geographic characteristics - no latitude adjustment

Table 10: Robustness check: Conditional convergence - no latitude adjustment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					lights per Et growth				
log lights per Et t-k	-0.021***	-0.018***	-0.019***	-0.022***	-0.021***	-0.021***	-0.021***	-0.022***	-0.021***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
unlit pixel share	0.019	-0.024**	-0.026**	0.024	0.002	0.002	-0.011	0.017	0.005
	(0.016)	(0.010)	(0.011)	(0.018)	(0.019)	(0.019)	(0.019)	(0.016)	(0.016)
Et share agri t-k	-0.093***			-0.083***	-0.032				
	(0.021)			(0.023)	(0.027)				
Et share manufacturing t-k						0.185**	0.028		
						(0.083)	(0.088)		
Et share services t-k								0.104***	0.055*
								(0.029)	(0.030)
density t-k	-0.050***			-0.070***	-0.091***	-0.060***	-0.083***	-0.077***	-0.097***
	(0.017)			(0.017)	(0.019)	(0.018)	(0.018)	(0.018)	(0.019)
deaths	-0.102***			-0.101***	-0.041	-0.083**	-0.031	-0.105***	-0.045
	(0.033)			(0.037)	(0.039)	(0.035)	(0.038)	(0.037)	(0.038)
distance to main city		-0.003**		-0.004***	-0.007***	-0.004***	-0.007***	-0.004***	-0.007***
		(0.001)		(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
distance to coast		-0.001		-0.002	-0.006*	-0.002	-0.006*	-0.002	-0.005*
		(0.002)		(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
land suitability			-0.069***	-0.073***	-0.055**	-0.077***	-0.056**	-0.074***	-0.055**
			(0.022)	(0.023)	(0.025)	(0.023)	(0.025)	(0.023)	(0.025)
ruggedness			-0.002	-0.003	-0.009	-0.003	-0.009	-0.003	-0.009
			(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
Precipitation			-0.000	-0.000	-0.000*	-0.000*	-0.000*	-0.000	-0.000*
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Temperature			-0.002	-0.001	-0.002	-0.002	-0.002	-0.001	-0.002
			(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Malaria index			-0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R^2	0.11	0.10	0.11	0.13	0.16	0.13	0.16	0.13	0.16
N	1136	1136	1136	1136	1136	1136	1136	1136	1136
County FE	No	No	No	No	Yes	No	Yes	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the results of the baseline specification controlling for variables measuring distances, natural characteristics, density and sector specialization. The sum of light is measured as (1+sol), no average, no latitude adjusted.