Rural labour allocation and urbanisation in Sub-Saharan Africa

VERY INCOMPLETE DRAFT, PLEASE DO NOT CITE

Hanne Van Cappellen
Institute of Development Policy (IOB), University of Antwerp, Belgium
hanne.vancappellen@uantwerpen.be – P +32 3 265 54 22

Joachim De Weerdt
Institute of Development Policy (IOB), University of Antwerp, Belgium
LICOS, Catholic University of Leuven, Belgium
joachim.deweerdt@uantwerpen.be – P +32 3 265 57 69

Abstract

Over the last few decades, Sub-Saharan Africa has been urbanizing at an unprecedented rate. While there is evidence that this has led to rural-to-urban migration, real structural transformation has not taken place: the majority of Africa's poor people still live in rural areas and are primarily engaged in low productivity agriculture. This paper addresses the link between urbanisation and the rural labour market. It is hypothesised that urbanisation stimulates both the demand for and supply of more working hours outside of agriculture. The proximity of an urban agglomeration induces a demand for diversified employment, and the highly seasonal agricultural calendar offers space for off-farm employment. By combining panel data on employment and night light data as a proxy for urbanisation, this paper explores the spatial and temporal link between rural labour supply and the proximity of agglomerations in three sub-Saharan African countries for the period 2008-2016. Not only does it evaluates the effect of urbanisation on the number of hours supplied, but it also provides insight in how these hours are allocated sector wise. Using LSMS-ISA employment data on a panel of more than 15 000 individuals, allows us to shift the focus from sector productivity to individual level productivity, as well as account for individual fixed effects. Nightlight data have shown to be a good proxy for urbanisation and are particularly interesting for Sub-Saharan Africa, where urbanisation statistics lag behind reality, are only sporadically available and lack international comparability. They provide us with fine-grained urbanisation information with which we can investigate the role of the emerging small towns (which are mushrooming up all over Africa) on the rural population. Our analysis finds that urbanisation has a significant positive effect in rural areas on hours supplied in wage labour as well as on the share of households engaging in non-farm enterprises. The fact that hours worked are not affected negatively confirms the premise that the seasonal agricultural schedule offers space for the supply of more working hours. These findings have the potential to inform the formulation of labour policies as well as urban planning that can maximize the positive effects of African urbanisation on the rural poor.

1. Introduction

Structural transformation in sub-Saharan Africa has been the subject of much debate. Macro-economic analysis show that the potential for economic growth stemming from reallocation of labour out of agriculture into services and manufacturing is substantial, given the large productivity gap between agriculture and the non-agricultural sector (Gollin, Lagakos, & Waugh, 2014; Lewis, 1955; McMillan, Rodrik, & Verduzco-Gallo, 2014). Puzzling is the fact that despite the large productivity gains that would come from such a move, structural transformation is not taking place.

Research into the explanations of why labour is not moving out of agriculture on a significant scale, broadly focus on two explanations: whether the agricultural productivity gap arises due to measurement error, or whether large movements out of agriculture are inhibited by significant exit barriers or selection (Adamopoulos & Restuccia, 2014; Beegle, De Weerdt, & Dercon, 2011; Bryan, Chowdhury, & Mobarak, 2014; Hall & Jones, 1999; Herrendorf & Schoellman, 2015; McCullough, 2017; Miguel & Hamory, 2009; Sen, 2006). Of particular interest are the studies nuancing findings on the large agricultural productivity gap as uncovered with macro-data. As such do Miguel and Hamory (2009) show that selection into migration due to differences in ability can explain a great deal of labour movements in Kenya. McCullough (2017) uses the LSMS-ISA micro data on individual labour supply to provide evidence that the measured agricultural productivity gap of Gollin et al. (2014) shrinks in half when taking into account differences in individual hours supplied.

If agriculture is not as unproductive as macro-economists claim it to be, economic growth stemming from structural transformation through merely a move out of agriculture, as explained by the theory of Lewis (1955), is limited. How then can economic growth in sub-Saharan Africa be materialised? Scholars such as Barrett, Christiaensen, Sheahan, and Shimeles (2017) Barrett, Reardon, and Webb (2001) and Steel and van Lindert (2017) have shown the potential of income diversification and nonfarm activities for economic development of rural Africa. McCullough (2017) and other scholars such as Nagler and Naudé (2013), Wiggins (2000) and Yumkella, Kormawa, Roepstorff, and Hawkins (2011) provide evidence of a strong growth linkage between agriculture and nonfarm activities, which can also stimulate rural development in sub-Saharan Africa.

This paper provides evidence that one such stimulator of rural economic growth through its positive effects on income diversification, nonfarm employment and working hours supplied, is growth in nearby towns. Urbanisation in sub-Saharan Africa is increasing in an unprecedented manner, yet much of it is still poorly understood and happening off the radar. One of the reasons why investigations on urbanisation dynamics in sub-Saharan Africa have remained scarce is the lack of reliable and regularly available statistical data. International comparison between urbanisation statistics is often made difficult by substantial definitional differences concerning urban status (Satterthwaite, 2010).

The rapid urbanisation in sub-Saharan Africa has also informed analyses on the process of structural transformation, or lack thereof (de Brauw et al., 2014; Henderson & Kriticos, 2017; McMillan, Rodrik, & Verduzco-Gallo, 2014). Many of these scholars touch upon the 'urbanisation without growth' phenomenon observed particularly in sub-Saharan African, while others try to dismiss these claims (Fay & Opal, 2000; Fox, 2012; Gollin et al., 2015; Jedwab, 2013; Onjala & K'Akumu, 2016). Certain is that the unfolding urbanisation patterns in sub-Saharan Africa and its links with development are puzzling from a classical economic point of view: urbanisation in sub-Saharan Africa has seemingly not been preceded by economic growth or industrialisation.

This paper uses the novel combination of the LSMS-ISA micro data on labour and night light data as a proxy for urbanisation into a panel dataset that investigates the impact of urbanisation on labour patterns in rural areas in three[1] sub-Sahara African countries. It finds evidence for a response of hours worked in rural areas and income diversification to an increase in urbanisation. Moreover, we find that it is primarily nearby urbanisation that has a significant effect on rural labour patterns. This adds evidence to the growing literature on the importance of small towns for rural development and poverty alleviation (Christiaensen, De Weerdt, & Kanbur, 2017; Gibson, Datt, Murgai, & Ravallion, 2017; Ingelaere, Christiaensen, De Weerdt, & Kanbur, 2018). This indicates that it is not the low productivity in agriculture that is the primary barrier to structural transformation in sub-Saharan Africa, but rather the rural underemployment. This also provides evidence that growth linkages between the agricultural and the non-agricultural sector are present, which shows that the Mellor-Johnson thesis is still relevant today (Dercon & Gollin, 2014).

This paper taps into two of the most challenging policy questions of today for sub-Saharan Africa. The majority of Africa's poor still live in rural areas and are predominantly employed in agriculture. Understanding livelihood strategies of the rural population is key in identifying proper policies to battle poverty. Focusing on the rural and thus predominantly agriculture, is still the most effective in reducing poverty (Christiaensen, 2018). At the same time does the steady urbanisation of sub-Saharan Africa require immediate policy attention. When urbanisation is poorly understood, it risks to be poorly managed and give rise to significant congestion costs, climate effects and rise in inequality (McGranahan & Satterthwaite, 2014). Mapping and understanding urbanisation is thus key in building effective institutional and policy frameworks to guarantee that urbanisation is advantageous for the economy and the society (Bloom, Canning, Fink, Khanna, & Salyer, 2010). Further doe the evidence that nearby urbanisation can have substantial positive effects on rural areas, warrants the need for focusing urban policymaking on the development of small towns, rather than focusing primarily on megacities.

2. Background

Household surveys have shown that rural labour markets in developing countries are characterized by two pervasive facts. The first is that agricultural workers work surprisingly few hours per year. For example, McCullough (2017) looks at LSMS-ISA data from four African countries and finds 700 hours worked per agricultural worker, per year, which would be equivalent to 88 working days of 8 hours work.

[1] Ethiopia, Malawi and Tanzania

The second is that agriculture requires labour intermittently throughout the agricultural season. For example, Arthi, Beegle, De Weerdt, and Palacios-López (2018) show that both at the extensive margin (who works on the farm) and at the intensive margin (how much they work), there is much irregularity.

These two stylized facts have a number of important implications. First, agriculture is not an intrinsically unproductive sector; at least not once we consider agricultural productivity as output per hour worked. Second, despite high per-hour-worked productivity, the total number of hours worked reveal significant levels of underemployment in agriculture. In short, the agricultural sector suffers much less from low productivity than it does from underemployment.

Off-farm activities present one opportunity to supply more hours of labour per year, but the supply of labour is restricted by the irregularity of the agricultural schedule. The challenge is for the labour demand to fit that schedule. As such do Nagler and Naudé (2014) show with a time series analysis on the LSMS-ISA data for six African countries that a significant part of non-farm entrepreneurship serves to complement seasonal agricultural labour. One source of demand for labour from outside agriculture could be in nearby urban centres. The pull factor of nearby urbanisation for nonfarm wage employment has already been shown by Fafchamps and Shilpi (2003) for the case of Nepal, and urbanisation as a stimulator for income diversification and off-farm employment has been mentioned by (Calì & Menon, 2013; Christiaensen, 2013; Nagler & Naudé, 2014). By linking urban growth data to rural household survey data, we will investigate empirically how urban growth is affecting local labour markets. The basic premise is that growth in a nearby town provides opportunities to supplement farm activities with off-farm work – opportunities that would not exist without the existence of the town.

3. Data description

3.1 LSMS-ISA

This paper draws on the World Bank Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) dataset to look at labour supplied in three different sectors: non-farm enterprises, wage (and other) labour, and farm labour. A panel dataset is constructed for three countries consisting of three waves: the first (2011-2012), second (2013-2014) and third wave (2015-2016) of the (Rural) Socioeconomic Survey of Ethiopia, the first (2010-2011), the second (2013) and the third wave (2016-2017) of the Third Integrated Household Survey of Malawi, and the first (2008-2009), second (2010-2011) and the third wave (2012-2013) of the Tanzania National Panel Survey. The LSMS-ISA project is implemented by the World Bank Group and aims at, in collaboration with national statistics offices, developing a database with novel and detailed statistics on household variables, with a particular focus on agriculture. Currently, it runs in eight sub-Saharan countries: Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania and Uganda. The LSMS-ISA data have been used in prior studies that look at labour productivity and labour patterns in sub-Saharan Africa (Allen, 2018; McCullough, 2017; Nagler & Naudé, 2013, 2014).

The datasets are nationally representative and have both rural and urban Enumeration Areas (EAs) in their sample. These EAs consist of both farming as well as non-farming households. The dataset is unique in its kind as it provides us with highly detailed information on household farming and labour involved, labour market participation and household non-farm enterprises. It generally consists of three questionnaires: a household questionnaire, an agricultural questionnaire and a community questionnaire. Most importantly are the survey data georeferenced at the EA level, which allows us to exploit spatial characteristics of the data.

The rural/urban status of the EAs in the LSMS-ISA dataset is derived from national definitions, often those applied in national censuses. Depending on the country, this definition applies a population size, settlement type or other criteria. As the criteria of these definitions often differ significantly, international comparability is limited (ILO, 2018; Satterthwaite, 2010). For this reason, rural/urban classification will be made based on a continent-wide definition that combines information from Night Time Lights datasets and the Africapolis dataset[2]. This will be explained in the next section.

The characteristics of the sample used in the analysis are described in Table 1. The sample of study consists of those individuals that are aged 15 or above, and reside in rural areas as defined by our definition based on Night Time Lights and Africapolis data. Only the individuals that did not move between waves are retained.

Table 1.
Sample characteristics at baseline

	Ethiopia	**Malawi**	**Tanzania**
Total number of individuals	17807	6382	14577
Individuals aged≥15	9506	3317	8370
% Male	48.9	46.3	47.6
% Female	51.1	53.7	52.4
Number of households	3639	1510	3088
Average household size	6.01	5.26	6.34
Average household size aged≤5	1.06	0.58	1.17
Average household size 5<aged≤15	2.10	1.84	1.93
Average household size15< aged≤65	2.68	2.72	3.07
Average household size >65y	0.16	0.13	0.18

The LSMS-ISA dataset is used to construct variables on individual labour supply in three main sectors: agriculture, non-farm household enterprises and wage (and other) labour. The constructed variables, their type and the linked survey question can be found in Table 2. Summary statistics can be found in Table 3.

[2] http://www.africapolis.org/home

Table 2.
Constructed labour variables

Last 12 months INDIVIDUAL LEVEL	**Ethiopia**	**Malawi**	**Tanzania**
Agricultural labour	Planting labour (land preparation, planting, ridging, weeding and fertilizing) + harvesting labour (harvesting, ridging) on household plots *Per plot: up to 8members/hh* *Hours (Weeks/year * days/week *hours/day)*	Agricultural labour in rainy and in dry season in three categories (land preparation and planting, weeding fertilizing and other non-harvest activities, and harvesting) on household plots *Per plot: up to 4 members/hh* *Hours(weeks/year*days/week*hours/day)*	Agricultural labour in long and in short rainy season in three categories (land preparation and planting, weeding, harvesting) on household plots *Per plot: up to 6 members/hh* *Days (days/year)*
Non-farm enterprise	Worked in any household non-farm enterprise *Per individual* *Dummy* Owns a non-farm enterprise *Per individual* *Dummy*	Worked in any household non-farm enterprise *Per individual* *Dummy* Owns a non-farm enterprise *Per individual* *Dummy*	Worked in any household non-farm enterprise *Per individual* *Dummy* Owns a non-farm enterprise *Per individual* *Dummy*
Wage (and other) labour	Main paid job (including casual/part-time labour)+ secondary paid job (including casual/part-time labour) + PSNP labour + other temporary/casual labour *Per individual* *Hours/year* *Main and secondary paid job: months/year*weeks/month*h/week* *PSNP and other temporary/casual labour: days/year**	Main paid job + secondary paid job + ganyu labour + unpaid apprenticeships + other unpaid labour *Per individual* *Hours/year* *Main and secondary paid job, unpaid apprenticeship and ganyu labour: Months/year*weeks/month*h/week* *Other unpaid labour: days/year**	Main paid job + secondary paid job + unpaid apprenticeships *Per individual* *Hours/year* *All categories: months/year*weeks/month*h/week*
HOUSEHOLD LEVEL			
Non-farm enterprise	Has any non-farm enterprise *Per household* *Dummy*	Has any non-farm enterprise *Per household* *Dummy*	Has any non-farm enterprise *Per household* *Dummy*

* For the analysis we assume 1 day consists of approximately 8 working hours

Table 3.
Summary statistics on labour variables

	Ethiopia			Malawi			Tanzania		
	W1	W2	W3	W1	W2	W3	W1	W2	W3
Share of households with at least one non-farm enterprise	28.33	33.83	35.52	25.44	34.43	35.66	45.10	49.52	47.90
Share of individuals aged≥15 that performed any wage labour	6.97	7.44	6.03	36.77	43.04	54.81	n.a	25.56	27.32
Av. hours worked in wage (and other non-farm) labour over the last 12 months	93.92	94.19	79.08	316.77	359.94	408.32	n.a.	322.58	396.08
Share of individuals aged≥15 that performed any household farm labour	66.35	68.10	61.94	71.39	72.81	72.81	57.24	56.49	54.70
Av. hours worked on the household farm over the last 12 months	273.14	260.57	232.72	168.03	182.18	210.94	314.30	314.97	308.02

3.2 Night Time Lights

3.2.1 Background

The use of Night Time Lights (NTL) to study socio-economic development is not new. Since the digitalization of Nighttime Lights datasets of the Defense Meteorological Satellite Program Operational Line Scanner (DMSP/OLS) in 1992, studies in a broad range of disciplines using these datasets have skyrocketed (Doll, 2008; Huang, Yang, Gao, Yang, & Zhao, 2014).

A significant part of these studies use NTL to study socioeconomic development or urbanization at the global, regional or national level (Bennett & Smith, 2017; Doll, 2008; Huang et al., 2014). Some investigate the statistical relationship between NTL emissions and socio-economic variables such as GDP, population density or built-up area, and it has been shown that in general NTL correlates well with socio-economic activity (Bennett & Smith, 2017; Briggs, Gulliver, Fecht, & Vienneau, 2007; Doll, Muller, & Morley, 2006; Sutton, Roberts, Elvidge, & Baugh, 2001; Zhang & Seto, 2013). The first studies proving the correlation between NTL and economic activity, population, electric power consumption and urban extent, date from already two decades ago (Elvidge, Baugh, Kihn, Kroehl, & Davis, 1997; C. D. Elvidge et al., 1997; L.Imhoff, Lawrence, Stutzer, & Elvidge, 1997). More recent studies have used NTL time series to investigate its power in explaining temporal changes in these variables (Bennett & Smith, 2017; Henderson, Storeygard, & Weil, 2012; Small & Elvidge, 2013; Yi et al., 2014). Other studies use NTL as a proxy for socio-economic variables of which reliable statistical data is lacking (Donaldson & Storeygard, 2016; Huang et al., 2014). These studies have been instrumental in uncovering distributional and temporal patterns of a range of variables identifying urban dynamics, such as urban boundaries, intercity dynamics, built-up area and population dynamics, both at one point in time as over time (Bennett & Smith, 2017; Doll, 2008; Huang et al., 2014; Ma et al., 2015; Zhang & Seto, 2011).

The bulk of the studies using NTL data to investigate urbanization dynamics are regionally skewed towards Asia and the US, with China leading the list (Bennett & Smith, 2017). In the last few years, several studies on Latin America have emerged (Álvarez-Berríos, Parés-Ramos, & Aide, 2013; Parés-Ramos, Álvarez-Berríos, & Aide, 2013; Rodriguez Lopez, Heider, & Scheffran, 2017; Soto Diaz, Vargas, & Berdegué, 2018). Studies using NTL in the context of sub-Saharan Africa are rather scarce. The region is featured in studies investigation the link between economic growth or urbanization and NTL on a global scale (Henderson et al., 2012; Zhang & Seto, 2013), and a few recent studies use NTL as a proxy for economic activity for one or more countries in Africa (Michalopoulos & Papaioannou, 2013, 2014; Rohner, Thoenig, & Zilibotti, 2013; Storeygard, 2016). To our knowledge, only three studies focusing specifically on urbanization and Nighttime Light in the African context (Binswanger-Mkhize & Savastano, 2017; Chen & Nordhaus, 2015; Savory et al., 2017).

The Nighttime Lights datasets are made available by the National Oceanic and Atmospheric Administration/National Geophysical Data Center (NOAA/NGDC). The NTL products come from two main sources. The oldest and most popular is the Version 4 Nighttime Lights Time Series Dataset from the Defense Meteorological Satellite Program Operational Line Scanner (DMSP-OLS). It was launched in the 1960s to detect cloud coverage, but since its digitalization in 1992, the nighttime images from the DMSP/OLS have become widely popular as an instrument to investigate anthropogenic activity (Doll, 2008). The latest and last Version 4 spans the period 1992-2013 and contains three different datasets collected from nine satellites. These three composites are a cloud free coverage, nighttime stable lights and average visible data. They differ in the extent of what they filter out and what they measure. The newest set of data comes from the Visible Infrared Imager Radiometer Suite (VIIRS) Day/Night Band(DNB) of the National Polar-Orbiting Operational Environmental Satellite System (NPOESS). It was launched in 2011 by NASA and NOAA and contains multiple significant improvements on the DMSP/OLS. Three cloud free products have been produced for 2015; a raw cloud-free composite, an outlier-removed cloud-free composite and a Night-Time Lights dataset (Elvidge, Baugh, Zhizhin, Hsu, & Ghosh, 2017). Chen and Nordhaus (2015) test the ability of the VIIRS data to serve as a proxy for population and economic output in Africa and find that it has the potential to improve the predictions of the DMSP/OLS datasets on socio-economic dynamics in Sub-Saharan Africa. The VIIRS data also have the potential to mitigate possible biases from DMSP/OLS data when using NTL as a proxy for urbanization in developing countries (Zhang & Seto, 2013).

Currently the DMSP/OLS datasets are still the most used due to its availability of long time series data and the limited availability of VIIRS processed composites. However, the improvements in the quality of the data compared with DMSP/OLS, the increase in products available in the near future and the development of intercalibration methods between the two sets of data (Li, Li, Xu, & Wu, 2017), have the potential to signal a new era in research using nighttime lights. More information on the datasets and the technicalities of the satellites can be found on NOAA/NGDCs website[3] and in (Doll, 2008; Elvidge et al., 2017; Huang et al., 2014).

Using DMSP/OLS datasets to study socio-economic patterns however come with widely recognized limitations. The most serious setback for time series analyses is the lack of inter- and intracalibration between different satellites. This is a potential serious problem, as it has been shown that the amount of NTL picked up by two different satellites in the same year can differ by as much as 10%. This is problematic when studying temporal changes over a period in which night time light is recorded by multiple satellites. Different intercalibration methods have been

[3] https://ngdc.noaa.gov/eog/index.html

developed to solve this issue (Huang et al., 2014; Small & Elvidge, 2013). Other common setbacks of NTL data are the presence of blooming, overglow, and the oversaturation of pixels. Blooming and overglow are the phenomena that nighttime light is recorded where there is actually no light, due to reflection of for example lakes (blooming), or the extension of nighttime lights from lighted pixels into the periphery (overglow). Oversaturation of pixels stems from the fact that the maximum DN value for light intensity is set at 63 which causes the DN values of very bright pixels to be truncated on top, with the consequence of possible loss of information (Huang et al., 2014; Zhang & Seto, 2011)

3.2.2 Used NTL dataset

Until recently, no paper documenting or dealing with these issues in the case of sub-Saharan Africa was available. Savory et al. (2017) however, have developed and made freely available an intercalibrated dataset for Africa for the period 2000-2013 based on the Stable Lights composites of the DMSP/OLS dataset[4]. We chose to use this dataset for our analysis for multiple reasons. First of all because with the use of the invariant region and quadratic regression method (IRQR) and Gaussian Process Modelling, it was able to develop a database that smoothed out the discrepancies between different satellite signals. This discrepancy was particularly present between satellite F16 and F18, covering the period 2004-2009 and 2010-2013 respectively, which is the time period of our analysis. Secondly because next to intercalibration, it also removes NTL stemming from gas flares and corrects for blooming by masking the datasets with Water Bodies datasets. Thirdly was not only the dataset especially developed for studying temporal urbanization dynamics in Africa, its capacity to do so was also tested by calculating correlations of the new dataset with GDP and urban population. It is shown that for our sample, the dataset performs particularly well with correlations at the country level with urban population all being 0.9 or above.

This dataset does not correct for overglow or saturation of pixels. Savory et al. decided to forgo correcting for overglow "due to its uncertain efficacy" (Savory et al., 2017, p. 6). As we are looking at changes over time, the (systematic) occurrence of overglow is less of a concern. Saturation of pixels is not a limitation in the case of Africa, as less than 1% of the dataset recorded a light intensity with DN value 63 for 2013. In our sample, no pixel with the maximum DN value of 63 was recorded.

Without neglecting the potential issues with NTL data, they nevertheless provide a huge potential for studying urbanization in sub-Saharan Africa that despite the large interest from scholars, remains relatively poorly understood. As mentioned by Álvarez-Berríos et al. (2013), Bennett and Smith (2017) and Donaldson and Storeygard (2016), NTL can prove especially useful for studying urbanization in countries for which official reliable statistics are lacking. The potential is fourfold. First of all does it eliminate the reliability on national statistics that are often only sporadically available, of questionable reliability and lag behind reality. Next is international comparability possible due to the universal coverage of the satellite data. This is often not the case for urban/rural definitions that are decided upon nationally and are hard to interpret and compare, even in one national context (Allen, 2018). Thirdly does the availability of NTL data for a long time period

[4] Freely available at https://geodata.globalhealthapp.net/

create the opportunity to look at temporal changes in urbanization dynamics, instead of the snapshot pictures that are derived from statistical data gathered at one point in time. Finally does the high spatial granularity (30 arc-second image grid) allow us to investigate urbanization dynamics on a subnational level.

3.2.3 NTL as identification of urban areas

Despite the large amount of studies using NTL to study urbanization dynamics, Bennett and Smith (2017) mention that the relationship between NTL and urbanization is context dependent, which limits the applicability of the methodologies used in the many studies on urbanization dynamics in the US, China and India to the African context. For example the extent to which agricultural area is lighted differs significantly between countries, which shows that no consensus on threshold exist between contexts (Ma et al., 2015; Small & Elvidge, 2013).

So while the potential added value of NTL for studying urbanization in SSA is huge due to the aforementioned reasons, the challenge lies in identifying a proper identification strategy that is best tailored to the African context. One of the issues with NTL is that it is not always clear what it measures. Although the Savory et al. (2017) database has been shown to be highly correlated with urbanisation, we use the Africapolis[5] database as an additional check to identify agglomerations. The Africapolis database is an effort of the OECD to provide novel historical data on agglomerations in (West) Africa. The most recent update includes detailed information on continent wide African agglomerations in 2015. The database is constructed using the combination of satellite data on built-up surface with population statistics.

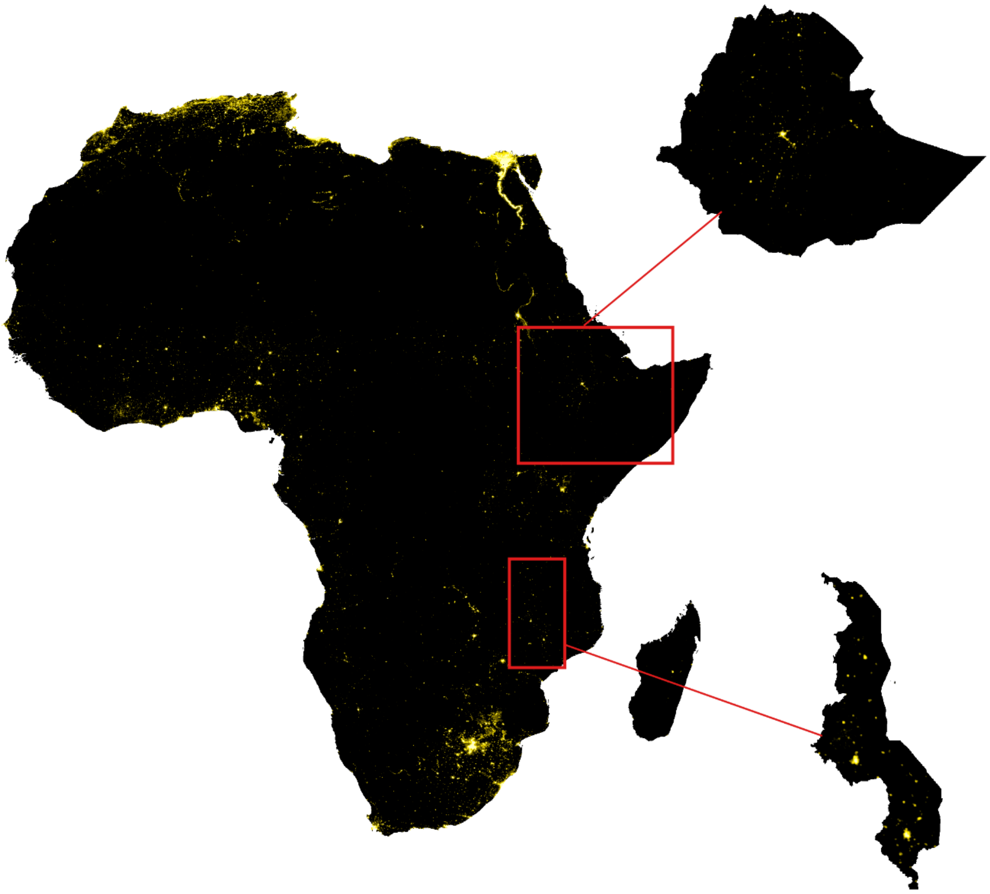
Over the course of our NTL sample (2005-2013), only the clusters of NTL that can be geographically linked with a 2015 Africapolis agglomeration are retained. By doing this, we increase the efficiency of NTL to identify agglomerations, and rule out NTL stemming from non-urban anthropogenic activities such as industry or mining. Further does this warrant against including NTL stemming from rural areas that have benefited from improved electrification due to rural electrification programs that have been rolled out across the continent over the last decade.

Consequently, we only retain EAs in our sample that have no NTL and are not linked to any Africapolis agglomeration.

Figure 1 shows NTL on the African continent as of 2013. We see that, rather than consisting of continuously lit surface, Africa mainly exists of a dark surface with some distinct light clusters. These light hubs are proven to be highly correlated with urbanisation (Savory et al., 2017).

[5] http://www.africapolis.org/

Figure 1
Night Time Lights in Africa in 2013, with a focus on Ethiopia and Malawi



The size of the NTL clusters in our sample range from only one lit pixel, which is approximately 1 km², to very large clusters of more than 1000 km².To make the analysis more intuitive, the remainder of this paper will refer to these light clusters as 'agglomerations'. Table 3 provides descriptive statistics on the agglomerations in the sample for each country and each year. Taking a look at the descriptive statistics of Table 4, we can see that both on the extensive and intensive margin, NTL increased between survey waves. As such did both the number of night light clusters and the total lit area, as well as their total sum of lights increased.

Table 4.
Descriptive statistics on Night Light Clusters

	Ethiopia			Malawi			Tanzania		
	2008	2010	2012	2007	2010	2013	2005	2007	2009
Number of urban clusters	181	191	193	44	45	46	115	124	128
Total sum of lights	76224	85285	92685	40654	42251	43625	70006	72745	78435
Largest sum of lights/cluster	28130	31164	34500	12949	12956	13002	21502	22641	24299
Smallest sum of lights/cluster	1	1	1	6	16	11	1	1	4
Max DN value in sample*	61	60	59	61	60	59	62	62	61
Av. max DN value/cluster	7.92	8.19	8.54	12.76	12.97	12.96	9.80	9.22	9.35
Total urban area	5763	6647	7018	3376	3594	3796	8556	9221	9961
As share of country area	0.51%	0.59%	0.62%	2.98%	3.18%	3.35%	0.89%	0.96%	1.04%

Each pixel in the used raster file is assigned a Digital Number (DN) that measures light intensity, ranging from 0 (no night light) to 63 (maximum light intensity that can be recorded). The sum of lights is the sum of the light intensity of each pixel that is part of a given light cluster.
*As in none of the country/years the maximum value of DN=63 is observed, there is no risk of saturation bias

4. Urban Access

Assuming the potential influence of surrounding urbanisation on a rural individual in a given EA depends on the size of as well as the distance to the surrounding agglomeration(s), we construct a variable that catches the 'Urban Access'[6] of an individual in a given EA in period t as follows:

$$\boldsymbol{UA_{EA,t}} = \sum_{\substack{j=1 \\ a<dist_{EA,j,t}\leq b}}^{n} \boldsymbol{sol_{j,t} * dist_{EA,j,t}^{-\beta}} \qquad (1)$$

With sol_j the sum of lights[7] of agglomeration *j* at time *t*, and $dist_{i,j}$ the distance between EA *i* and the border of agglomeration *j* at time *t*. β is a discount factor that weights this distance. The larger β, the less urban influence further away agglomerations are assumed to have on a rural area. The lower restriction provides the option to only include agglomerations that are location in a specifc concentric circle around the EA.

The choice of β is rather intuitive. Suppose we have three agglomerations with sol=50, sol=500 and sol=5000. This values are more or less representative in our sample for a small, a medium and large agglomeration. Setting β =-1 would mean that a big agglomeration 1000 km away has the same influence as a medium agglomeration 100 km away, and a small agglomeration just 10 km away. Intuitively, this clearly overestimates the importance of faraway agglomerations. Setting β =-2 would mean that a big agglomeration 316 km away has the same influence as a medium agglomeration 31 km away, and a small agglomeration just 10 km away.

[6] Inspiration was derived from the market access variable constructed by Blankespoor, Mesplé-Somps, and Spielvogel (2016)
[7] The sum of lights is the sum of the light intensity (expressed as a DN value ranging from 0 to 63) of each pixel that is part of a given light cluster.

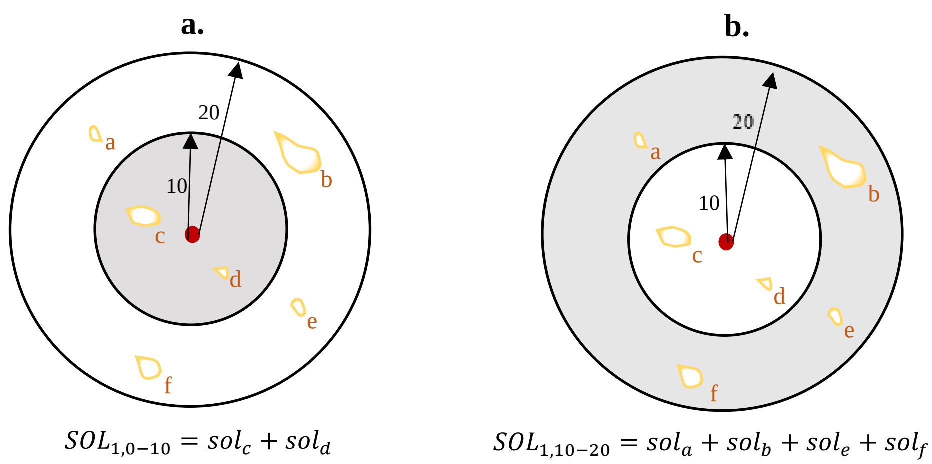
A more intuitive case of this Urban Access variable is the Sum Of Lights variable. This variable adds up the sum of lights of each agglomeration located in the area between two concentric circles around the EA *i*:

$$\boldsymbol{SOL_{EA,t,a,b}} = \sum_{\substack{\boldsymbol{j=1} \\ \boldsymbol{a<dist_{EA,j,t}\leq b}}}^{\boldsymbol{n}} \boldsymbol{sol_{j,t}} \qquad \textbf{(2)}$$

With a the radius of the lower limit concentric circle around *i*, b the radius of the upper limit concentric circle around *i*, $sol_{j,t}$ the sum of lights of agglomeration j at time *t*, and $dist_{i,j,t}$ the distance between EA *i* and the border of agglomeration j. The summation is conditional on $dist_{i,j}$ laying between the lower and upper radius of the concentric circles around *i*. Figure 2 provides an illustration of this SOL formula.

The Sum Of Lights is a popular measure in literature using NTL to study economic development, as it takes both the size of the lit area as well as the intensity into account (Ghosh et al., 2010; Gibson et al., 2017; Henderson, Squires, Storeygard, & Weil, 2018; Small & Elvidge, 2013). It can thus be thought of as balancing the importance of the intensive margin as well as the extensive margin of urbanisation.

Figure 2: Sum Of Lights urbanisation variable illustrated



Suppose the red dot is an EA with i=1 and the yellow spots are agglomerations surrounding EA_1 at time *t*, defined by night light clusters. The first concentric circle around $EA_{1,t}$ has a radius of 10 kilometres, while the second concentric circle around it has a radius of 20 kilometres. Suppose we want to calculate the SOL of the area up to 10 kilometres from the EA, and the area surrounding this area up to 20 kilometres from the EA, as indicated by the grey areas in a. and b. The SOL of the grey area in a. would then be $SOL_{1,0-10} = sol_c + sol_d$, and the SOL of the grey area in b. would be $SOL_{1,10-20} = sol_a + sol_b + sol_e + sol_f$.

It is useful to disentangle urbanisation stemming from different concentric circles around a rural area to be able to assess the relative importance of urbanisation happening in areas at different distances from rural areas. In our analysis, we will generally apply the 0-5 km, 5-10 km, 10-20 km, 20-30 km, 30-40 km, etc intervals. As such does the 0-10km interval captures more or less any location within walking distance, the 10-20km interval any location that can be reached by bike or public transport, the 20-50 km interval a distance that demands a longer travel of more or less maximum a day.

4.1 Descriptive analysis

A first exploration of the data clearly indicates that different labour moves over time in rural areas experience different spatial patterns of urbanisation. In Table 4 the demeaned change in SOL in four different concentric circles (doughnuts) around the rural survey areas are calculated for three main sectors of employment: non-farm enterprise, wage labour and agricultural labour. More specifically, it shows the average demeaned change in SOL for four distinct labour choices over time: the individuals who were not active in the particular sector in year 1 and were still not in year 2, those who were not active but became active in year 2, those who were working in the sector in year 1 but not anymore in year 2, and those who were already active in the sector in year 1 and remained active in year 2. By splitting up the change in urbanisation over time in these for different choice groups, we can get a clearer insight in the link between labour supply choices and urbanisation growth.

In general are the deviations from the average urbanisation trend diminishing with urban growth further away. If we zoom in on the individuals that selected into non-farm enterprise labour where they previously did not supply any hours of labour in this sector, we see that they experienced urban growth in the area within 10 km from their place of residence that is 4.3% higher than the average urbanisation trend for the 10km circle around rural areas for the sample as a whole. For wage labour, this is even 9.3%. However, this differential effect dissolves quickly with distance from the rural area. Figure 3 explores spatially the link between moving into non-farm enterprise jobs and wage labour respectively, and experienced urbanisation growth. Further do we see that above average urbanisation growth in the 10-20 concentric circle around a rural area, is linked with moving out of non-farm enterprises as well as out of agriculture. This might indicate that substantial urbanisation growth in not too close but not too far agglomerations might stimulate people to get a job in these towns. Lastly are differential urbanisation experiences becoming very small once we look at urbanisation happening at a distance of 50 kilometres and beyond. This could indicate that proximity of urbanisation is an important indicator for influencing labour choices.

Table 5 and Figure 3 provide a first indication that different kinds of urbanisation drive different labour decisions, and that proximity of agglomerations is possibly an important indicator for labour choice. The following section will explore empirically the link between urbanisation and rural labour supply.

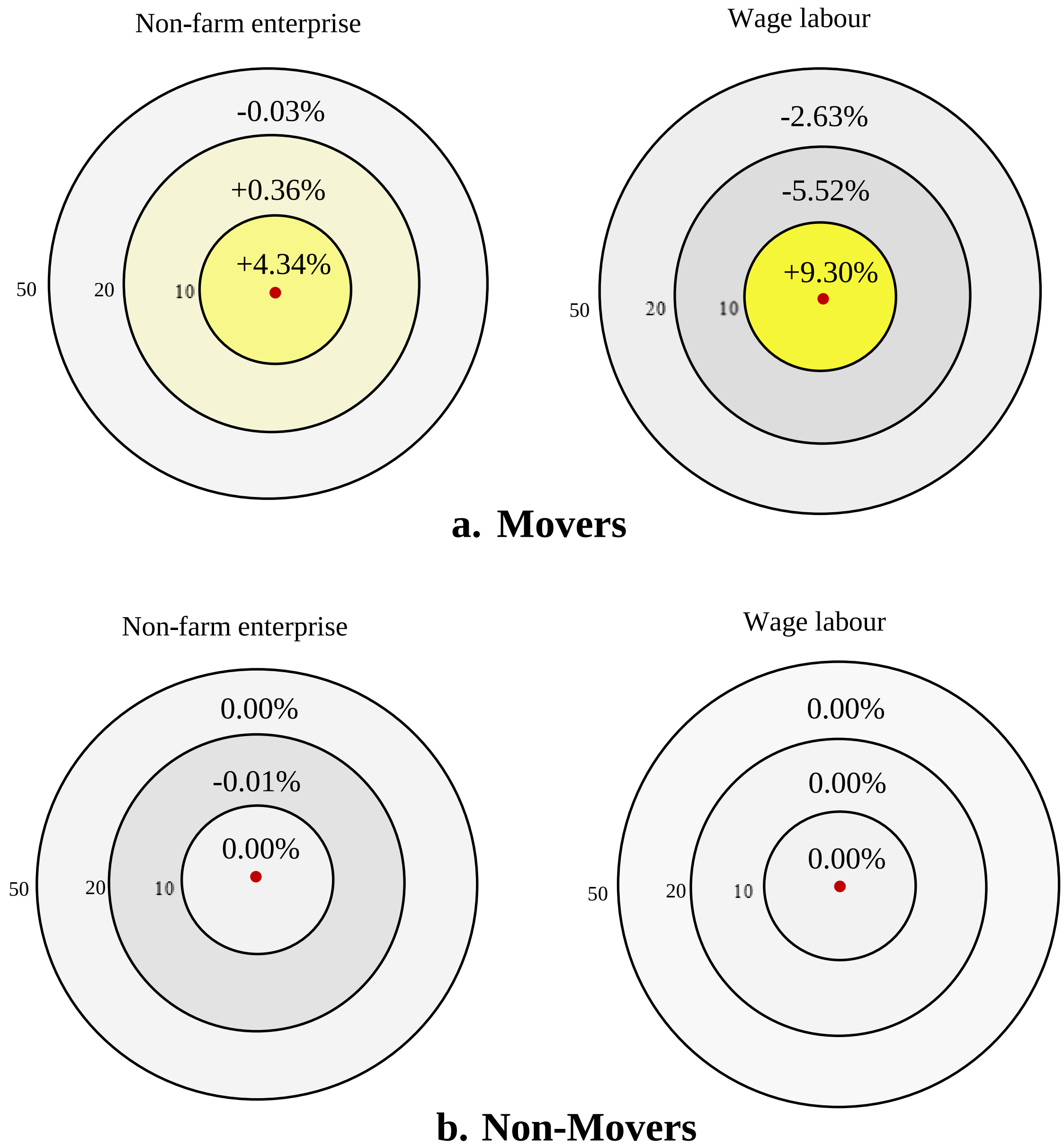
Table 5.
Average demeaned urbanisation growth in different concentric circles around different categories of rural workers

	0-10 km around EA*	10-20 km around EA*	20-50 km around EA*	50 km-country border*	N
		Employed in non-farm enterprise (Y1-Y2)**			
No-No	**-.0005304**	**-.0098712**	**-.000298**	-2.58e-06	10555
No-Yes	**.0434274**	**.0036803**	**-.0034757**	-.0032515	942
Yes-No	-.0370858	.0287074	.0440364	.0017455	790
Yes-Yes	-.0101034	.1303695	-.0473129	.0029662	603
		Employed in wage labour (Y1-Y2)**			
No-No	**-.0032051**	**.001723**	**-.0015904**	.0001446	11534
No-Yes	**.0930292**	**-.0552352**	**-.0263264**	-.0043161	363
Yes-No	-.0140562	-.0079301	.0034593	-.001394	405
Yes-Yes	.052105	-.0053796	.0611685	-.0002787	460
		Employed in agriculture (Y1-Y2)**			
No-No	.0270427	-.0775466	.0196506	.0099442	1990
No-Yes	-.0287958	-.0549578	-.0046742	.0069742	1747
Yes-No	-.0168436	.0231068	-.0171385	-.0022641	1006
Yes-Yes	.0016394	.027943	-.0017009	-.0036354	8147
Average change	0.1433	0.1629	0.1529	0.0913	

*The urbanisation is measured as the log of SOL formula, demeaned: $\Delta lnSOL_{i0,10} - \overline{lnSOL_{0,10}}$ to facilitate interpretation.
**These variables look at the extensive margin (working in a non-farm enterprise/wage labour/agriculture or not), but individuals are not necessarily working exclusively in these categories.
The percentages in bold are illustrated in Figure 3.

Figure 3

Differential urbanisation growth experienced by an average person moving into non-farm enterprise labour or wage labour respectively (a.) versus an average person that did not move into non-farm enterprise labour or wage labour (b.)



The top part of this figure (a.) shows what happens in terms of urban growth around an average individual that either moved into working in a non-farm enterprise or into performing wage labour over the period of analysis. The percentage in each concentric circle (or doughnut) indicates the change in urbanisation, trend demeaned. It shows that in both cases over the period between this move, both those who selected into non-farm enterprise labour as well as wage labour, experienced an above average increase of urbanisation (as measured by the sum of lights) in a radius of 10 kilometres around their place of residence, of respectively 4.34 and 9.30 percent. However, in the area further than 10 kilometres from the individuals place of residence, these individuals experienced almost no more than average change in urbanisation in the case of the non-farm enterprise mover, or even below average urbanisation growth in the case of the wage labour mover.

The bottom part of this figure (b.) shows what happens in terms of urban growth around an average individual that did not move into working in a non-farm enterprise or into performing wage labour over the period of analysis. It is clear that these individuals experienced a distinct pattern of urban growth compared with the movers: no above average or even slightly below average urban growth was found in a distance of up to 50 kilometres around their place of residence.

5. Estimation strategy

As we are interested in the within individual effect over time of urbanisation on rural labour supply, our baseline specification is the following:

$$\boldsymbol{\Delta L_{it} = \beta_0 + \beta_1 \Delta u_{it} + \beta_n \Delta c_{nit} + \beta_m c_{i1} + \Delta \varepsilon_{it}} \qquad \textbf{(3)}$$

With $\boldsymbol{\Delta L_{it}}$ the change in labour supply, $\boldsymbol{\Delta u_{it}}$ a measure of change in urbanisation, in our case specified by the UA (1) or SOL (2) constructed variable, $\boldsymbol{\Delta c_{nit}}$ a measure of change in control variable(s), $\boldsymbol{c_{i1}}$ baseline (individual) characteristics and $\boldsymbol{\varepsilon_{it}}$ a time variant error term.
By performing a within transformation on the panel data, we are able to control for all time-invariant unobserved individual characteristics. This is a powerful specification as it assures that our result are not driven by unobserved individual heterogeneity. One of the most important unobserved characteristics that can influence labour supply is undoubtedly ability (Miguel & Hamory, 2009). As we are dealing with geospatial data, spatial correlation stemming from an spatial autoregressive error term or spatial lagged dependence is a valid concern. To address these concerns, we use the Conley spatial HAC standard errors for models with fixed effects based on the code developed by Solomon Hsiang and building on the Spatial HAC errors as discussed in Conley (Conley, 1999; Conley, Hansen, McCulloch, & Rossi, 2008).[8]

We want to investigate the effect of urbanisation on labour supply in three distinct sectors: non-farm enterprises, wage labour and agricultural labour. By investigating the individual labour supply in each of this sectors on the individual level, we avoid assigning individuals to one main sector of employment, as is often done in macro-economic research on labour productivity. In this way, we are able to get a deeper insight on the extent to which individuals on the one hand seek income diversification and secondly seek fuller employment by increasing total working hours.
Further do we control for both time varying variables as well as for time invariant characteristics by including the (interaction with) baseline characteristics. As time varying control variables we include a household shock dummy and an amenities index at community level to control for (the change in) rural development.

[8] http://www.trfetzer.com/conley-spatial-hac-errors-with-fixed-effects/

6. Results

6.1 Main results

Table 6 reports the fixed effects estimation with Conley Standard Errors for the baseline specification. It looks at the effect of urbanisation in terms of SOL (expressed in SD) in different concentric circles around the EA on both hours supplied to wage (and other non-farm) labour, labour on own farm and the share of households engaged in non-farm enterprises. Specification (1) shows that a one standard deviation increase in the sum of light in the 0-5 km concentric circle around the EA translates into an increase in hours worked in wage labour of 139.8 hours a year. Similarly, this increases the probability of a household to have any non-farm enterprise with 7.1 percentage points. Given the mean baseline values of hours worked in wage labour of 219.08 hours and the percentage of households engaging in non-farm enterprises of 34.50%, this effect is quite substantial. However, it is not straightforward to interpret a one SD increase in SOL. So to place this effect in a more relatable setting, it is informative to look at the relationship between the size of an agglomeration and its SOL value.

In our sample, the average SOL value of an agglomeration of an area size between 40-80 km² is 330. An agglomeration of a size between 100 – 150 km² has an average SOL value of 927. This implies that the growth of a small agglomeration into a medium agglomeration means an average increase in around 600 SOL. If this happens in the 0-5km concentric circle, the hours worked in wage (and other non-farm) labour increase with 61 hours a year. Similarly will the probability of a household to have a non-farm enterprise increase with 3 percentage points. For the 5-10 concentric circle, this effect would be an increase of 34 hours and 4.5 percentage points respectively.

Table 6 also shows that the effect of urban growth on labour allocation outside of the farm dies out quickly: after a distance of 10 kilometres radius around the EA, no significant effect on wage (and other non-farm) labour and on non-farm enterprises is observed anymore. This indicates that it is especially nearby urbanisation that can positively affect hours supplied of rural individuals.

Specification (3) of Table 6 estimates the effect of urbanisation on hours worked on the household farm, and finds a positive effect on hours worked in the household farm in four of the six concentric circles. This provides a first indication that urbanisation is not only providing opportunities to fill in gaps in the labour calendar with off-farm activities, but that it is also positively affecting the agricultural sector.

Table 6.

	(1) Wage and other (non-farm) labour	(2) % of hh with non-farm enterprise	(3) (own) farm labour
Baseline average	219.08	34.50	272.92
1 year lagged sol 0-5 km in SD	139.8***	0.0714*	2.472
(1372.66)	(33.09)	(0.0409)	(23.21)
1 year lagged sol 5-10 km in SD	103.6***	0.137***	129.0**
(1811.75)	(35.78)	(0.0411)	(64.88)
1 year lagged sol 10-20 km in SD	41.17	-0.0402	107.3**
(2711.19)	(47.85)	(0.0414)	(51.59)
1 year lagged sol 20-30 km in SD	39.84	-0.0383	92.38**
(2150.103)	(41.97)	(0.0331)	(39.83)
1 year lagged sol 30-40 km in SD	10.95	0.00890*	-6.956
(3060.24)	(6.702)	(0.00519)	(6.709)
1 year lagged sol 40-50 km in SD	-2.439	-0.000144	28.49***
(2527.25)	(2.102)	(0.00291)	(3.575)
Observations	26,975	35,954	36,119
R-squared	0.001	0.002	0.004

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.2 Agglomeration effects

The baseline specification estimated the effect on an increase of urbanisation in terms of an increase of sum of lights, regardless of whether this increase came from growth in one particular city, or from growth divided among different (smaller) cities. However, it is plausible these two cases have a different impact on labour allocations due to agglomeration effects. (tbc – Table 7)

Table 7.

	(1) Wage and other (non-farm) labour	(2) % of hh with non-farm enterprise	(3) (own) farm labour
Baseline average	219.08	34.50	272.92
1 year lagged sol 0-5 km in SD	319.8***	0.357**	15.25
(1372.66)	(103.8)	(0.154)	(150.4)
1 year lagged sol 5-10 km in SD	-2,055***	-1.559*	-711.2
	(566.0)	(0.934)	(1,601)
1 year lagged sol 5-10 km in SD	-652.3**	2.733***	-7,641***
	(265.4)	(0.316)	(609.0)
Aggl 0-5 km * sol 0-5 km (1Y lag)	-0.138*	-0.000226**	0.00171
	(0.0726)	(0.000115)	(0.109)
1 year lagged # aggl 5-10 km	44.36	0.128***	-145.5***
	(28.78)	(0.0424)	(42.64)
Aggl 5-10 km * sol 5-10 km (1Y lag)	1.181***	0.000941*	0.457
	(0.311)	(0.000514)	(0.881)
1 year lagged # aggl 5-10 km	19.55	0.0350	-79.39
	(18.62)	(0.0314)	(55.90)
Aggl 10-20 km * sol 10-20 km (1Y lag)	0.240**	-0.00101***	2.819***
	(0.0980)	(0.000116)	(0.224)
1 year lagged # aggl 10-20 km	-12.45***	-0.0477***	-24.65**
	(4.364)	(0.00584)	(11.34)
Observations	26,975	35,954	36,119
R-squared	0.001	0.003	0.002

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.3 The effect of differential in baseline surrounding urbanisation level

Tbc

6.4 Socio-economic differential effects

Interactions with sex, age : tbc

7. Robustness checks

In this section we will show that our findings are robust to alternative specifications. As such we will control for existing road network in rural areas.
(tbc)

8. Conclusion

So far, true structural transformation in sub-Saharan Africa has not taken place. Studies questioning the size of the agricultural productivity gap in sub-Saharan Africa cast doubts on the potential gains to be materialised by labour moving out of agriculture.

McCullough (2017) showed that agricultural labourers have an excess of labour hours to be absorbed by a demand inside and/or outside of agriculture. Our findings resonate with this premise as they provide evidence that (nearby) urbanisation offers opportunities in rural areas to fill employment gaps in the agricultural labour schedule by supplying working hours outside of agriculture. As a consequence, (nearby) urbanisation might positively affect rural poverty. This is also in line with the findings of Nagler and Naudé (2014) that a significant part of non-farm entrepreneurship in sub-Saharan Africa serves to complement seasonal agricultural labour. Further does it add to papers that find evidence for growth linkages between the agricultural and non-agricultural sector, as well as the finding from Calì and Menon (2013) that urbanisation has a poverty reducing effect largely due to spill overs from the urban economy, and not necessarily due to rural-urban migration.

Secondly, did this paper find evidence that it is especially nearby urbanisation that is affecting rural labour supply choices. This adds to the continuously expanding literature that shows that growth in small towns can provide an important alley into rural poverty reduction (Christiaensen, 2013; Christiaensen et al., 2017; Christiaensen & Kanbur, 2017; Gibson et al., 2017). This has also potential implications for policymaking; investing in thriving small towns is not only beneficial for the urban population, but also has positive spill-overs to the surrounding rural areas.

A further understanding of the tight connection between urbanisation and labour patterns is essential for designing policies that may stimulate both agricultural and off-farm economic activities. Until now, rural entrepreneurship has been largely neglected in policy strategies for rural development (Nagler & Naudé, 2013). As Africa’s poor still mainly live in rural areas, understanding the livelihood strategies of the rural population is key for informing policies that can strengthen the possibilities of the rural population for fuller employment, and thus initiate rural poverty reduction.

Appendix

Table A
Non-farm enterprises survey questions

Ethiopia	Malawi
Over the past 12 months, has anyone in this household ..	Over the past 12 months has anyone in your household…
1) … owned a non-agricultural business or provided a non-agricultural service from home or a household-owned shop, as a carwash owner, metal worker, mechanic, carpenter, tailor, barber, etc.?	1) … owned a non-agricultural business or provided a non-agricultural service from home or a household-owned shop, as a carwash owner, metal worker, mechanic, carpenter, tailor, barber, etc.?
2) … processed and sold any agricultural by-products, including flour, local beer (tella), 'areke", "enjera", seed, etc., but excluding livestock by-products, fresh/processed fish?	2) … processed and sold any agricultural by-products, including flour, starch, juice, beer, jam, oil, seed, bran, etc., but excluding livestock by-products, fresh/processed fish?
3) … owned a trading business on a street or in a market?	3) … owned a trading business on a street or in a market?
4) … offered any service or sold anything on a street or in a market, including firewood, home-made charcoal, construction timber, woodpoles, traditional medicine, mats, bricks, cane furniture, weave baskets, thatch grass etc.?	4) … offered any service or sold anything on a street or in a market, including firewood, home-made charcoal, curios, construction timber, woodpoles, traditional medicine, mats, bricks, cane furniture, weave baskets, thatch grass etc.?
5) … owned a professional office or offered professional services from home as a doctor, accountant, lawyer, translator, private tutor, midwife, mason, etc?	5) … owned a professional office or offered professional services from home as a doctor, accountant, lawyer, translator, private tutor, midwife, mason, etc?
6) … driven a household-owned taxi or pick-up truck to provide transportation or moving services?	6) … driven a household-owned taxi or pick-up truck to provide transportation or moving services?
7) … owned a bar or restaurant?	7) … owned a bar or restaurant?
8) … owned any other non-agricultural business, even if it is a small business run from home or on a street?	8) …owned any other non-agricultural business, even if it is a small business run from home or on a street?

Reference list

Adamopoulos, T., & Restuccia, D. (2014). The Size Distribution of Farms and International Productivity Differences. *American Economic Review, 104*(6), 1667-1697. doi:10.1257/aer.104.6.1667

Allen, J. E. (2018). Are agricultural markets more developed around cities? Testing for urban heterogeneity in separability in Tanzania. *Food Policy, 79*, 199-212. doi:https://doi.org/10.1016/j.foodpol.2018.07.005

Álvarez-Berríos, N. L., Parés-Ramos, I. K., & Aide, T. M. (2013). Contrasting Patterns of Urban Expansion in Colombia, Ecuador, Peru, and Bolivia Between 1992 and 2009. *AMBIO, 42*(1), 29-40. doi:10.1007/s13280-012-0344-8

Arthi, V., Beegle, K., De Weerdt, J., & Palacios-López, A. (2018). Not your average job: Measuring farm labor in Tanzania. *Journal of Development Economics, 130*, 160-172. doi:https://doi.org/10.1016/j.jdeveco.2017.10.005

Barrett, C. B., Christiaensen, L., Sheahan, M., & Shimeles, A. (2017). On the Structural Transformation of Rural Africa. *Journal of African Economies, 26*, 11-35. doi:10.1093/jae/ejx009

Barrett, C. B., Reardon, T., & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications. *Food Policy, 26*(4), 315-331. doi:https://doi.org/10.1016/S0306-9192(01)00014-8

Beegle, K., De Weerdt, J., & Dercon, S. (2011). Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey. *Review of Economics and Statistics, 93*(3), 1010-1033.

Bennett, M. M., & Smith, L. C. (2017). Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics. *Remote Sensing of Environment, 192*, 176-197. doi:https://doi.org/10.1016/j.rse.2017.01.005

Binswanger-Mkhize, H. P., & Savastano, S. (2017). Agricultural intensification: The status in six African countries. *Food Policy, 67*, 26-40. doi:https://doi.org/10.1016/j.foodpol.2016.09.021

Blankespoor, B., Mesplé-Somps, S. S., Harris, & Spielvogel, G. (2016). The dynamics of villages, towns and cities in sub-Saharan Africa: The case of Mali.

Bloom, D. E., Canning, D., Fink, G., Khanna, T., & Salyer, P. (2010). Urban Settlement: Data, Measures, and Trends. In J. Beall, B. Guha-Khasnobis, & R. Kanbur (Eds.), *Urbanization and development: multidisciplinary perspectives* (pp. 19-40). Oxford: Oxford University Press.

Briggs, D. J., Gulliver, J., Fecht, D., & Vienneau, D. M. (2007). Dasymetric modelling of small-area population distribution using land cover and light emissions data. *Remote Sensing of Environment, 108*(4), 451-466. doi:https://doi.org/10.1016/j.rse.2006.11.020

Bryan, G., Chowdhury, S., & Mobarak, A. M. (2014). Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica, 82*(5), 1671-1748. doi:10.3982/Ecta10489

Calì, M., & Menon, C. (2013). Does Urbanization Affect Rural Poverty? Evidence from Indian Districts. *The World Bank Economic Review, 27*(2), 171-201. doi:10.1093/wber/lhs019

Chen, X., & Nordhaus, W. (2015). A Test of the New VIIRS Lights Data Set: Population and Economic Output in Africa. *Remote Sensing, 7*(4), 4937.

Christiaensen, L. (2013). Introduction: rural diversification, secondary towns and poverty reduction: do not miss the middle. *Agricultural Economics, 44*(4-5), 433-434.

Christiaensen, L. (2018). Five new insights on how agriculture can help reduce poverty. Retrieved from http://blogs.worldbank.org/jobs/five-new-insights-how-agriculture-can-help-reduce-poverty

Christiaensen, L., De Weerdt, J., & Kanbur, R. (2017). *Where to create jobs to reduce poverty: cities or towns?* : World Bank Group.

Christiaensen, L., & Kanbur, R. (2017). Secondary Towns and Poverty Reduction: Refocusing the Urbanization Agenda. *Annual Review of Resource Economics, 9*(1), 405-419. doi:10.1146/annurev-resource-100516-053453

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

Conley, T. G., Hansen, C. B., McCulloch, R. E., & Rossi, P. E. (2008). A semi-parametric Bayesian approach to the instrumental variable problem. *Journal of Econometrics, 144*(1), 276-305.

Dercon, S., & Gollin, D. (2014). Agriculture in African Development: Theories and Strategies. *Annual Review of Resource Economics, 6*(1), 471-492. doi:10.1146/annurev-resource-100913-012706

Doll, C. N. (2008). CIESIN thematic guide to night-time light remote sensing and its applications. *Center for International Earth Science Information Network of Columbia University, Palisades, NY*.

Doll, C. N. H., Muller, J.-P., & Morley, J. G. (2006). Mapping regional economic activity from night-time light satellite imagery. *Ecological Economics, 57*(1), 75-92. doi:https://doi.org/10.1016/j.ecolecon.2005.03.007

Donaldson, D., & Storeygard, A. (2016). The View from Above: Applications of Satellite Data in Economics. *Journal of Economic Perspectives, 30*(4), 171-198. doi:doi: 10.1257/jep.30.4.171

Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. C., & Ghosh, T. (2017). VIIRS night-time lights. *International Journal of Remote Sensing, 38*(21), 5860-5879. doi:10.1080/01431161.2017.1342050

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., & Davis, E. R. (1997). Mapping city lights with nighttime data from the DMSP Operational Linescan System. *Photogrammetric Engineering and Remote Sensing, 63*(6), 727-734.

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R., & Davis, C. W. (1997). Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *International Journal of Remote Sensing, 18*(6), 1373-1379. doi:10.1080/014311697218485

Fafchamps, M., & Shilpi, F. (2003). The spatial division of labour in Nepal. *The Journal of Development Studies, 39*(6), 23-66. doi:10.1080/00220380312331293577

Ghosh, T., L Powell, R., D Elvidge, C., E Baugh, K., C Sutton, P., & Anderson, S. (2010). Shedding light on the global distribution of economic activity. *The Open Geography Journal, 3*(1).

Gibson, J., Datt, G., Murgai, R., & Ravallion, M. (2017). For India's Rural Poor, Growing Towns Matter More Than Growing Cities. *World Development, 98*, 413-429. doi:10.1016/j.worlddev.2017.05.014

Gollin, D., Lagakos, D., & Waugh, M. E. (2014). The Agricultural Productivity Gap. *The Quarterly Journal of Economics, 129*(2), 939-993. doi:10.1093/qje/qjt056

Hall, R. E., & Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics, 114*(1), 83-116. doi:Doi 10.1162/003355399555954

Henderson, J. V., Squires, T., Storeygard, A., & Weil, D. (2018). The Global Distribution of Economic Activity: Nature, History, and the Role of Trade1. *The Quarterly Journal of Economics, 133*(1), 357-406. doi:10.1093/qje/qjx030

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review, 102*(2), 994-1028. doi:doi: 10.1257/aer.102.2.994

Herrendorf, B., & Schoellman, T. (2015). Why is measured productivity so low in agriculture? *Review of Economic Dynamics, 18*(4), 1003-1022. doi:10.1016/j.red.2014.10.006

Huang, Q., Yang, X., Gao, B., Yang, Y., & Zhao, Y. (2014). Application of DMSP/OLS Nighttime Light Images: A Meta-Analysis and a Systematic Literature Review. *Remote Sensing, 6*(8), 6844.

ILO. (2018). *Rural-urban labour statistics*. Paper presented at the 20th International Conference of Labour Staticians Geneva

Ingelaere, B., Christiaensen, L., De Weerdt, J., & Kanbur, R. (2018). Why secondary towns can be important for poverty reduction–A migrant perspective. *World Development, 105*(C), 273-282.

L.Imhoff, M., Lawrence, W. T., Stutzer, D. C., & Elvidge, C. D. (1997). A technique for using composite DMSP/OLS "City Lights" satellite data to map urban area. *Remote Sensing of Environment, 61*(3), 361-370. doi:https://doi.org/10.1016/S0034-4257(97)00046-1

Lewis, W. A. (1955). *Theory of Economic Growth*: Routledge.

Li, X., Li, D., Xu, H., & Wu, C. (2017). Intercalibration between DMSP/OLS and VIIRS night-time light images to evaluate city light dynamics of Syria's major human settlement during Syrian Civil War. *International Journal of Remote Sensing, 38*(21), 5934-5951. doi:10.1080/01431161.2017.1331476

Ma, T., Zhou, Y., Zhou, C., Haynie, S., Pei, T., & Xu, T. (2015). Night-time light derived estimation of spatio-temporal characteristics of urbanization dynamics using DMSP/OLS satellite data. *Remote Sensing of Environment, 158*, 453-464. doi:https://doi.org/10.1016/j.rse.2014.11.022

McCullough, E. B. (2017). Labor productivity and employment gaps in Sub-Saharan Africa. *Food Policy, 67*, 133-152. doi:10.1016/j.foodpol.2016.09.013

McGranahan, G., & Satterthwaite, D. (2014). *Working paper: Urbanisation concepts and trends*. London: IIED.

McMillan, M., Rodrik, D., & Verduzco-Gallo, Í. (2014). Globalization, Structural Change, and Productivity Growth, with an Update on Africa. *World Development, 63*, 11-32. doi:10.1016/j.worlddev.2013.10.012

Michalopoulos, S., & Papaioannou, E. (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica, 81*(1), 113-152. doi:doi:10.3982/ECTA9613

Michalopoulos, S., & Papaioannou, E. (2014). National Institutions and Subnational Development in Africa *. *The Quarterly Journal of Economics, 129*(1), 151-213. doi:10.1093/qje/qjt029

Miguel, E., & Hamory, J. (2009). *Individual Ability and Selection into Migration in Kenya*. Retrieved from

Nagler, P., & Naudé, W. (2013). Non-farm entrepreneurship in rural Africa: Patterns and determinants.

Nagler, P., & Naudé, W. (2014). Labor productivity in rural African enterprises: empirical evidence from the LSMS-ISA.

Parés-Ramos, I., Álvarez-Berríos, N., & Aide, T. (2013). Mapping Urbanization Dynamics in Major Cities of Colombia, Ecuador, Perú, and Bolivia Using Night-Time Satellite Imagery. *Land, 2*(1), 37.

Rodriguez Lopez, J. M., Heider, K., & Scheffran, J. (2017). Frontiers of urbanization: Identifying and explaining urbanization hot spots in the south of Mexico City using human and remote sensing. *Applied Geography, 79*, 1-10. doi:https://doi.org/10.1016/j.apgeog.2016.12.001

Rohner, D., Thoenig, M., & Zilibotti, F. (2013). Seeds of distrust: conflict in Uganda. *Journal of Economic Growth, 18*(3), 217-252. doi:10.1007/s10887-013-9093-1

Satterthwaite, D. (2010). Urban myths and the mis-use of data that underpin them: working paper No. 2010/28. In: World Institute for Development Economics Research.

Savory, D. J., Andrade-Pacheco, R., Gething, P. W., Midekisa, A., Bennett, A., & Sturrock, H. J. W. (2017). Intercalibration and Gaussian Process Modeling of Nighttime Lights Imagery for Measuring Urbanization Trends in Africa 2000–2013. *Remote Sensing, 9*(7), 713.

Sen, K. H. (2006). The kin system as a poverty trap? *Poverty traps*, 95.

Small, C., & Elvidge, C. D. (2013). Night on Earth: Mapping decadal changes of anthropogenic night light in Asia. *International Journal of Applied Earth Observation and Geoinformation, 22*, 40-52. doi:https://doi.org/10.1016/j.jag.2012.02.009

Soto Diaz, J. D., Vargas, O., & Berdegué, J. A. (2018). How Large Are the Contributions of Cities to the Development of Rural Communities? A Market Access Approach for a Quarter Century of Evidence from Chile. *A Market Access Approach for a Quarter Century of Evidence from Chile (December 1, 2018)*.

Steel, G., & van Lindert, P. (2017). *Rural livelihood transformations and local development in Cameroon, Ghana and Tanzania*. International Institute for Environment and Development.
Storeygard, A. (2016). Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa. *The Review of Economic Studies, 83*(3), 1263-1295. doi:10.1093/restud/rdw020
Sutton, P., Roberts, D., Elvidge, C., & Baugh, K. (2001). Census from Heaven: An estimate of the global human population using night-time satellite imagery. *International Journal of Remote Sensing, 22*(16), 3061-3076. doi:10.1080/01431160010007015
Wiggins, S. (2000). Interpreting Changes from the 1970s to the 1990s in African Agriculture Through Village Studies. *World Development, 28*(4), 631-662. doi:https://doi.org/10.1016/S0305-750X(99)00153-9
Yi, K., Tani, H., Li, Q., Zhang, J., Guo, M., Bao, Y., . . . Li, J. (2014). Mapping and Evaluating the Urbanization Process in Northeast China Using DMSP/OLS Nighttime Light Data. *Sensors, 14*(2), 3207.
Yumkella, K. K., Kormawa, P. M., Roepstorff, T. M., & Hawkins, A. M. (2011). Agribusiness for Africa's Prosperity.
Zhang, Q., & Seto, K. (2013). Can Night-Time Light Data Identify Typologies of Urbanization? A Global Assessment of Successes and Failures. *Remote Sensing, 5*(7), 3476.
Zhang, Q., & Seto, K. C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment, 115*(9), 2320-2329. doi:https://doi.org/10.1016/j.rse.2011.04.032