MALAWI
IFPRI

STRATEGY SUPPORT PROGRAM | WORKING PAPER 43 SEPTEMBER 2023

Rural underemployment and urbanization

Insights from a nine-year household panel survey from Malawi

Hanne Van Cappellen
Joachim De Weerdt

CONTENTS

TABLES

FIGURES

ABSTRACT

Rural labor markets in Africa are frequently characterized by underemployment, with farmers unable to fully deploy throughout the year one of their most important assets—their labor. Using a nine-year panel data set on 1,407 working-age adults from rural Malawi, we document changes in rural underemployment over this period and how they are associated with urbanization. Nearby urban growth results in increased hours worked in casual labor (*ganyu*) and in non-agricultural sectors, at the expense of work on the household farm. Improved urban access is also associated with a small increase in wage labor and, at the intensive margin, with hours supplied in household enterprises. We draw lessons from these results for policies, investments, and interventions to leverage urban growth for rural development.

1 INTRODUCTION

Extreme poverty numbers are coming down worldwide, but the progress is slowest in sub-Saharan Africa (or Africa for short), which implies that Africa is hosting an ever-larger share of the world's extreme poor. Within a decade, 87 percent of the extreme poor will live in Africa (Beegle & Christiaensen, 2019). African poverty is concentrated in the rural areas. Currently, four out of every five Africans living under the extreme poverty line are found in rural areas, with most pursuing livelihoods that center around small-scale, labor-intensive farming with very low levels of productivity per worker. At the same time, Africa is rapidly urbanizing and will continue to do so both in number of urban agglomerations and in the share of the population living in urban areas. Africa's population is expected to double between now and 2050, and two-thirds of that growth is anticipated to be absorbed by urban areas, adding 950 million urban residents over the next 30 years (OECD, 2020; U.N., 2018). With rapid urbanization happening amid deep-rooted rural poverty, a key question should be whether and how urban growth can be leveraged for rural poverty reduction.

Urban growth can influence rural poverty through a variety of channels (Binswanger-Mkhize, Johnson, Samboko, & You, 2016; Diao, Magalhaes, & Silver, 2019; Soto, Vargas, & Berdegué, 2018).

- Rural areas can cater to growing urban demand for food and, in turn, rural areas can benefit from flows of capital, services, and information from nearby urban areas (Dorosh & Thurlow, 2014; Gibson, Datt, Murgai, & Ravallion, 2017; Vandercasteelen, Beyene, Minten, & Swinnen, 2018; Vandercasteelen, Minten, & Tamru, 2021).
- Urban connectedness has been shown to positively impact consumer welfare through prices and variety (Gunning, Krishnan, & Mengistu, 2018; Krishnan & Zhang, 2020).
- Urbanization can stimulate income diversification and off-farm employment through increasing the demand for activities auxiliary to the agricultural sector, which in turn can trigger rural development (Barrett, Christiaensen, Sheahan, & Shimeles, 2017; Barrett, Reardon, & Webb, 2001; Calì & Menon, 2013; Christiaensen, 2013; McCullough, 2017; Nagler & Naudé, 2014; Steel & van Lindert, 2017; Tacoli, 1998).
- Urbanization also goes hand in hand with changes in dietary patterns, increasing demand for processed foods and opening up income earning opportunities along the value chain (Cockx, Colen, & De Weerdt, 2018; Reardon et al., 2015; Tschirley, Reardon, Dolislager, & Snyder, 2015).
- Finally, urban growth can absorb rural labor through temporary or permanent migration, putting upward pressure on rural wages and potentially increasing urban-rural remittance flows (Beegle, De Weerdt, & Dercon, 2011; Bryan, Chowdhury, & Mobarak, 2014; de Brauw, Mueller, & Lee, 2014; Fafchamps & Shilpi, 2003; Lewis, 1955).

All of these linkages will depend not only on the extent of urban growth, but also on the distance between the rural and the urban area (De Weerdt, Christiaensen, & Kanbur, 2021; Lucas, 2001; Soto et al., 2018).

Building on this research literature, this paper studies the association between urban growth and rural labor markets. Two stylized facts motivate this focus. First, productivity levels of smallholder farmers are low, not only compared to farmers in high-income countries, but also compared to workers in the manufacturing or service sectors in the same country. In the poorest quartile of countries, value-added per worker in the agricultural sector is two to three times smaller than the value-added by a non-agricultural worker (Gollin, Lagakos, & Waugh, 2014). Secondly, rural labor markets tend to be characterized by pervasive underemployment. McCullough (2017) shows how across four African countries people in the agricultural sector work 700 hours per year, compared to 1,850 hours per year for non-agricultural workers. Whether one views agriculture as an unproductive sector then very much depends on whether one uses per capita productivity numbers or per hour productivity numbers. Per person, relatively little value is added. But per hour, value added is much higher.

Thus, a more nuanced picture emerges of an agricultural sector that is, per hour worked, roughly equally productive compared to the non-agricultural sector, but in which people are supplying much fewer hours of labor. Most of the farmers supplying these hours worked are among the poorest people in the world, whose marginal valuation of additional consumption should be very high. It therefore seems safe to assume that the low hours of labor supplied is not due to a preference for leisure, but is due to a lack of demand for their labor. And because labor is one of the most important assets of the poor, the possibility that it is underutilized is, both from an academic and policy perspective, an issue that requires attention.

But agricultural labor schedules are highly irregular. Arthi, Beegle, De Weerdt, and Palacios-López (2018), for example, show that both at the extensive margin (who works on the farm) and at the intensive margin (how much they work), there is a lot of temporal variation and irregularity in the agricultural calendar. The extent to which labor demand induced by urban growth can fill gaps in the rural labor calendar may depend on whether it can flexibly fill these irregular gaps, dictated by the agricultural season and the idiosyncrasies of each farm and farming household. On-farm activities that diversify agricultural production and, particularly, those that have different seasonal cycles could play a role here. The scope for doing this with rainfed crop production may be limited, but irrigation and greenhouse farming could reduce dependence on seasons, as could livestock rearing. Off-farm employment, to the extent that it is not pro-cyclical to agricultural production, also holds promise. Nagler and Naudé (2014) suggest with a time series analysis on household survey data for six African countries that a significant part of non-farm entrepreneurship serves to complement seasonal agricultural labor.

In this paper we study how urban growth is associated with rural labor markets in Malawi, a country characterized by high levels of rural poverty and rapid urbanization. We first establish baseline characteristics of rural labor markets for our sample. We find evidence of rural underemployment and labor seasonality and that job diversification off-farm offers opportunities to increase total hours worked. We examine 1,407 individuals aged 15 to 65 years living in 852 households in rural Malawi. These individuals were surveyed in 2010 and again in 2019 as part of the Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) of the World Bank. The LSMS-ISA data have been used in prior studies that look at labor productivity and labor patterns in sub-Saharan Africa, such as Allen (2018); McCullough (2017), and Nagler and Naudé (2013, 2014). We assess how the number of hours of labor individuals supply each year changed and how that has varied by differential exposure to urban areas. More specifically, we look into what employment categories are associated with urban growth and for which workers. We hypothesize that the higher nearby urban growth is, the more hours of work rural individuals will do. We split total hours worked annually into work hours on the household farm of the individual, in own non-farm household enterprises, in wage labor, and in casual labor (*ganyu*).

We find that people living in rural areas with expanded urban access supply more hours of work. A 10 percent increase in urban access is associated with 14 days of extra work annually. These effects mainly play out at the intensive margin, which means that people already involved in a category of activity supply more hours within it with improved urban access, but few people are entering or exiting categories of activities in response to nearby urbanization. We identify a tendency for urbanization to be linked to more casual work and away from the household farm, an association that is more pronounced for young and female rural dwellers. Finally, we find that mainly growth in older, well-established cities and urban agglomerations drive these effects.

To the best of our knowledge, we are the first to characterize the geographical distribution of hours worked in income earning activities by distance to and size of (nearby) urban centers in an African context and the first to study how the ongoing urbanization process is affecting that distribution. Our findings are important as they add to the understanding of how Africa's steady urbanization, one of the most important demographic trends on the continent, can be leveraged for rural poverty reduction.

The article is structured as follows. The following section provides background information on the Malawian setting. Section 3 discusses the data and the measures on rural labor and urbanization that we use. Section 4 looks at the rural labor market in Malawi through our labor data, discussing rural worker profiles, rural labor seasonality, and job diversification. Section 5 explains our estimation strategy to link changes in urban access to changes in rural labor allocation. Section 6 gives the main results and discusses heterogeneity around them. Section 7 concludes and elaborates on the policy implications of our results.

2 BACKGROUND ON MALAWI

Malawi is one of the poorest countries in the world, with 80 percent of households relying on agriculture for their income and 94 percent of the poor living in rural areas (Benson, 2021; Caruso & Cardona Sosa, 2022). The vast majority of farmers engage in rainfed agriculture with farming systems largely based on growing maize as a staple crop alongside a limited number of other food and cash crops. Rainfall determines the agricultural labor schedule. The busiest period on the farm is at the beginning of the rainy season in December and January when all crops need planting and are competing for labor. After planting, tending to the fields and harvesting requires a less concentrated schedule and is spread out more over time. For example, the harvests of tobacco (the main export crop) and maize (the main staple crop), do not coincide closely. Once the rains stop and all harvests are in, a long dry period follows with relatively little work on the farm until the fields need to be prepared for the next agricultural cycle.

With an official urbanization rate of only 16 percent and with just four official cities, Malawi is one of the least urbanized countries in Africa. At the same time, it is also one of the fastest urbanizing countries on the continent. Remotely sensed data show that the number of population agglomerations with at least 10,000 inhabitants increased exponentially from just one in 1950 to 77 in 2015 (Africapolis 2022). This suggests that currently about one third of Malawi's population lives in agglomerations of over 10,000 people, which could reasonably be considered urban. Lilongwe and Blantyre, both with populations of about one million inhabitants, make up 45 percent of the urban population, while several dozen smaller urban agglomerations with populations up to around 100,000 make up 41 percent. Squeezed in the middle of this bimodal population distribution lie the cities of Zomba and Mzuzu, with 284,000 and 236,000 inhabitants, respectively (Africapolis, 2022).[1]

This rapid population increase together with small average plot size and declining soil fertility increases the pressure on land across Malawi (Asfaw, Orecchia, Pallante, & Palma, 2018). With stagnating agricultural productivity, alternative labor opportunities off the land are necessary to reduce poverty (Benson & De Weerdt, 2023; Caruso & Cardona Sosa, 2022). We hypothesize that nearby urbanization might offer opportunities to increase hours worked through strengthened rural-urban linkages.

3 DATA AND ANALYTICAL MEASURES

3.1 Measuring rural labor

We use the 2010-11 and 2019-20 rounds of Malawi's Integrated Household Panel Survey (IHPS).[2] This nationally representative panel is part of the Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) program in which the National Statistics Office of Malawi collaborates with the World Bank to collect high-quality household panel survey data, with a focus on agriculture. We impose a number of restrictions on the sample for the purpose of our analysis. We include only rural[3] Enumeration Areas (EA) and individuals aged between 15 and 65 years who either did not move between the two survey rounds, or if they did, moved within a 10 km radius of their baseline location.[4] Permanent rural migration is a possible channel through which urbanization affects rural labor patterns, but not one we will document in this analysis. Our interest is to document changes in local labor markets and how they are associated with nearby urbanization, keeping location constant. We remove migrants who permanently moved further than 10 km away from their baseline location, resulting in a loss of 4% (165) individuals from our sample.[5] Temporary migration, another important channel through which urbanization might affect rural employment, does not affect inclusion in the sample, as such individuals remain listed on the household roster. We further restrict the sample to individuals who have worked at least one hour in any employment category, who were interviewed in both rounds, who were not in school in either of the two rounds, and have no labor data used in the analysis missing. Table 3.1 provides the construction of the final sample of 1,407 rural individuals, given these restrictions. Some sample characteristics at baseline are also presented.

[1] These estimates are of population numbers within the remotely sensed urban agglomeration boundaries, which differ from administrative boundaries.

[2] We use data from the 2016-17 IHPS round in our robustness checks section in Annex A.

[3] For our main results we follow the urban/rural EA classification of the IHPS baseline survey that is based on the official administrative definition of what constitutes as urban area in Malawi. However, 45 out of the 77 Africapolis agglomerations used in our sample are not officially recognized as an urban area, so it is possible that some so-called 'rural' EAs are located within one of these 45 unofficial agglomerations. In Annex A, we test the robustness of our results against different subsets of rural EAs.

[4] 3.6% of the individuals in our sample moved to a location less than 10 km from their baseline location.

[5] Of the 4228 rural individuals that were reinterviewed in 2019, 467 (11%) migrated more than 10 km from their baseline location. Of those 467, only 165 satisfied the sample restrictions of having worked at least one hour in any employment category, of being aged between 15 and 65 and of not being in school in either baseline or endline.

The initial 2010-2011 rural sample consists of 48.31 percent men and 51.69 women. Women had higher recontact rates and the sample of rural individuals for which both baseline and endline data is available consists of 46.35 percent men and 53.65 percent women. Our sample restrictions further exacerbate this gender imbalance: relatively more men were in school (50.6% of rural school-going individuals over both rounds were men) and proportionally more women were 15-65 years old, satisfying the age restriction. There is no additional effect when dropping migrants, or excluding people who did not work even one hour in the last year.

We study the link between urbanization and rural labor markets for the period 2010-2019. This is the longest period over which we have panel data and neither of these two survey rounds were conducted in years with extreme weather events, unlike 2016 which coincided with a drought (further details in Annex Figure 7 and Annex Figure 8 in Annex D).

There are two important drawbacks of the labor data that we use. The first is that we do not have accurate data on reproductive work and household chores. Two questions ask about time spent collecting firewood and water the day prior to the interview, which women report doing on average for 77 minutes in 2010 and 81 minutes in 2019. For men this is 8 and 9 minutes, respectively.[6] Excluding household work therefore has important implications when comparing levels of employment across men and women, but as the average stays constant over time it will be plausibly differenced out in the main regressions (see Section 5.2).

Secondly, recalling labor over a long time period can be cognitively taxing on respondents when schedules are highly variable during the season. Arthi et al. (2018) show how, in the absence of a regular work schedule, responses appear to be extrapolated from recent experiences. Furthermore, less salient work tends to be forgotten, for example if done by someone only intermittently involved in the activity. We expect the joint effect of these two potential reporting errors to bias the estimates of the levels, but it is unclear in which direction. Because respondents are interviewed during the same period in 2010 and 2019, we expect such biases to be attenuated in the differences.

[6] Wodon and Beegle (2006) use 2004 data, which, in addition to asking about collecting firewood and fetching water, also asks about time spent "cooking, doing laundry, cleaning your house, and the like" the day prior to the interview. They find, first, that all these tasks add 23 hours to a woman's working week and 4 hours to a man's and, second, that there is very little seasonal variation in hours worked on these tasks.

Table 3.1. Sample construction and characteristics

Derivation of sample		Sample characteristics at baseline (2010)	
Individuals in sample for 2010-11 IHPS	7,682	Average age, years	32.9
Households	1,619	Attended school, %	76.2
Enumeration Areas	102	Household size, average	5.55
Rural individuals	5,512	Members aged under 5 years	1.25
Died or left household permanently	190 (3.4%)	Members aged 5 to 15 years	1.63
From households not found or replaced in a previous round	1,092 (19.8%)	Members aged 15 to 65 years	2.57
Re-interviewed in 2019	4,228	Members aged over 65 years	0.10
Did not move or relocated within 10km of baseline location	3,761		
Aged between 15 and 65 years in both rounds	1,608		
Worked at least one hour in at least one round	1,591		
Not in school in both rounds	1,407		
Final sample for analysis			
Enumeration Areas	72		
Households	852		
Individuals	1,407		
Male, %	41.7		
Female, %	58.3		

Source: Authors' analysis.

We use the detailed data available on hours worked in the past 12 months across four employment categories: work on the household farm, wage labor, casual labor (*ganyu*[7]), and labor in household non-farm enterprises. Table 3.2 gives an overview of the various labor variables available and how they were constructed.

[7] "*Ganyu* labour is short-term labour hired on a daily or other short-term basis. Most commonly, piecework weeding or ridging on the fields of other smallholders or on agricultural estates. However, *ganyu* labour can also be used for non-agricultural tasks, such as construction and gardening" (National Statistical Office, 2012b, p. 48).

Table 3.2. Constructed labor variables

Variable	Construction
Household farm labor, hours/year	Agricultural labor in rainy season and in dry season on household plots (three categories—land preparation and planting; weeding, fertilizing, and other non-harvest activities; harvesting), with up to four household members working on a plot.
Labor in household non-farm enterprises (NFE), including self-employed fishing business, hours/year	Labor in any kind of non-farm household enterprise (NFE) + fishing labor in high and low seasons (four categories—full-time fishing, part-time fishing, fish trading, and fish processing), with up to five household members working in an NFE
Labor in household NFE in agriculture sector, hours/year	Labor in any kind of household NFE in agricultural industry + fishing labor
Labor in household NFE in non-agricultural sector, hours/year	Labor in any kind of household NFE in non-agricultural industry
Wage labor, hours/year	Main paid job + secondary paid job
Wage labor in agriculture sector, hours/year	Main paid job + secondary paid job in agricultural industry
Wage labor in non-agricultural sector, hours/year	Main paid job + secondary paid job in non-agricultural industry
Casual labor, hours/year*	Ganyu labor
Total hours worked, hours/year	Household farm labor + NFE labor + wage labor + ganyu labor

Source: Authors' compilation.
Note: * Casual labor was reported in days/year. For comparability with other labor measures, it was assumed that an average working day consists of eight working hours.

3.2 Measuring urbanization

To measure urbanization, we use two different sets of satellite-based data: Africapolis and WorldPop data.[8] Satellite data can prove especially useful for studying urbanization in countries for which official reliable statistics are lacking (Donaldson & Storeygard, 2016). Five advantages are typically cited.

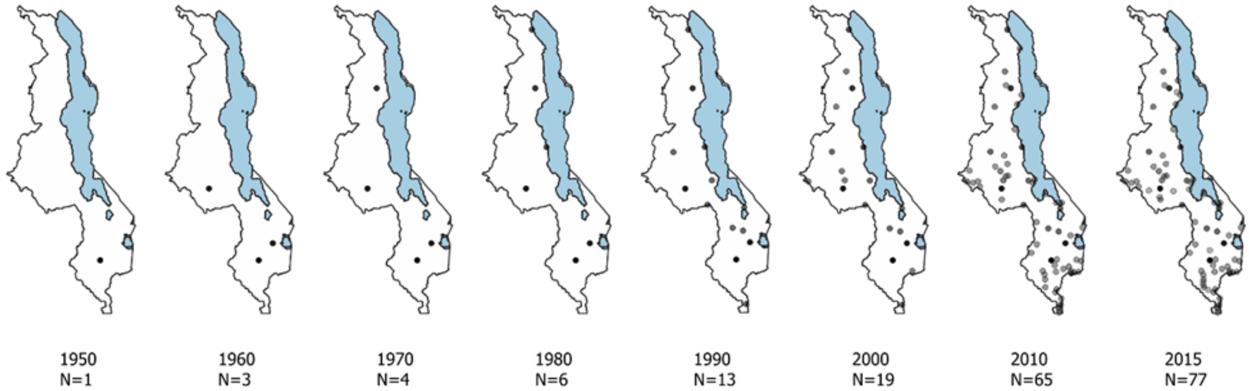
- Satellite-based measures overcome problems of comparability across and within countries of what constitutes an urban or a rural area (Allen, 2018).
- The availability of satellite data over a long time period creates the opportunity to look at temporal changes in urbanization dynamics.
- The high spatial granularity allows for investigating urbanization dynamics on a subnational level.
- Satellite data address concerns about the binariness of traditional urban measures, as it can be measured on a continuous scale. By modelling urbanization as a fundamentally spatial issue, the actual spatial connectedness between agglomerations and their surrounding rural areas can be exploited, instead of seeing rural and the urban as two opposing concepts (Beall, Guha-Khasnobis, & Kanbur, 2010; Brenner & Schmid, 2014).
- Remotely sensed data can be used for analyses in places where official statistics are unreliable or only sporadically available.

[8] A third one, the Visible Infrared Imaging Radiometer Suite (VIIRS) satellite Version 2 (V2) data, is used for a robustness check. See Annex A.

3.2.1 Africapolis

The Africapolis database is an effort of the OECD to provide historical data on urban agglomerations in Africa. The most recent update includes detailed information continent-wide on such agglomerations up to 2015. It uses a combination of satellite and census data to identify built-up area as well as to estimate the population in each agglomeration. A universal urban population threshold of 10,000 is applied. Population statistics are available at ten year intervals for the period 1950 to 2010 as well as for 2015. Figure 3.1 shows the growth in agglomerations in Malawi over the period 1950 to 2015—from just one in 1950 to 19 in 2000 to 77 in 2015. Much of this urbanization takes place under the official radar—45 of the 77 identified agglomerations are not officially recognized urban areas.

Figure 3.1. Spatial evolution of urban agglomerations in Malawi, 1950 to 2015



Source: Author's composition, based on Africapolis (2022).

The Africapolis database includes shapefiles of the built-up area of each agglomeration in 2015. We used these shapefiles as boundaries to demarcate each agglomeration in Malawi. To measure the level of 'urbanness' and changes in urbanization levels over time, we use yearly estimates of total population within these boundaries using the WorldPop database.

3.2.2 WorldPop

The second set of data used is the population count for each agglomeration which is derived from the WorldPop open-source database of gridded population estimates (Tatem, 2017).

There exists a multitude of gridded population data such as WorldPop, LandScan Global, the Gridded Population of the World (GPW), the Global Human Settlement Population Grid (GHS-POP) and the Global Rural-Urban Mapping Project (GRUMP). Yin et al. (2021) have shown for the case of Mainland South East Asia that WorldPop and Landscan Global provide more accurate estimates than GPW and GHS-POP. LandScan Global and WorldPop are the only databases that produce yearly data, but WorldPop provides the finest resolution, approximately 100m at the equator, versus one km for LandScan Global. Dorward and Fox (2022), in their study on population pressure, political institutions, and protests, favor the use of WorldPop over Landscan as it provides more consistent growth rates for the sample of African cities in their study. The high spatial granularity and relative superiority in performance leads us, also, to opt for WorldPop data. Other studies using WorldPop data to study urbanization or population dynamics in Africa are Chamberlain, Lazar, and Tatem (2022); Meredith et al. (2021); Pezzulo et al. (2017); Westlowski, Bengtsson, Buckee, Wetter, and Tatem (2014).

We use the unconstrained 100m resolution WorldPop product, which provides yearly estimates since 2000.[9] It uses a 'top-down' estimation modelling method to produce datasets that are based on census data projections but which are disaggregated to grid cell-based counts using the Random Forests estimation technique (Stevens et al., 2020). The disaggregation is done with the help of a range of detailed geospatial datasets, including land cover data, remotely sensed raster data on, for example, nighttime light, temperature, elevation, or slope, and spatial administrative data. The population counts are disaggregated over a 3 arc-second raster grid, which approximately corresponds to a 100-meter square at the equator.

This WorldPop dataset serves as our main source of measure of 'urbanness'. We use it to calculate population statistics for all 77 agglomerations in the Africapolis agglomeration shapefile, for baseline and endline. With GIS software (QGIS), the sum of the population of each 3 arc-second raster cell that falls inside the Africapolis agglomeration is calculated. We also use WorldPop data to calculate population within administrative boundaries, used in Annex B to test for local population growth. Table 3.3 provides WorldPop population statistics for the 77 urban agglomerations in Malawi. The 2019 population estimates for each of the 77 agglomerations are presented in Annex Table 8 in Annex D.

Table 3.3 reveals considerable discrepancy between Africapolis and WorldPop population estimates. For example, whereas Africapolis counts 77 agglomerations with a population above 10,000 people in 2015, WorldPop counts only 34. Annex C explains and illustrates these differences in more detail and motivates why we think these discrepancies are unlikely to bias our results.

Table 3.3. Population statistics for Africapolis-based urban agglomerations based on WorldPop spatial population data

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Population, average	30,150	31,133	32,257	33,384	34,296	35,425	36,725	38,204	39,674	41,300
Minimum	810	830	862	926	960	985	1,037	1,073	1,081	1,024
Maximum	682,103	715,118	753,048	795,436	835,145	883,352	933,927	990,710	1,054,810	1,124,959
Agglomerations, no.										
Population < 5,000	25	25	25	26	25	24	24	24	24	23
5,000 to 10,000	19	19	18	17	18	17	16	15	15	16
10,000 to 50,000	29	29	30	30	30	32	33	34	34	34
50,000 to 500,000	2	2	2	2	2	2	2	2	2	2
500,000 or more	2	2	2	2	2	2	2	2	2	2

Source: Authors' analysis of WorldPop unconstrained 100m resolution spatial population database product.
Note: Total agglomerations: 77.

[9] The unconstrained method divides the population estimates over all land grid squares globally, as opposed to the constrained method which projects estimates only within areas mapped as containing built settlements. https://www.worldpop.org/methods/top_down_constrained_vs_unconstrained

4 RURAL LABOR PATTERNS IN MALAWI

Before assessing the impact of nearby urbanization on hours worked annually by rural individuals, this section describes rural employment and underemployment patterns in our sample. We present statistics on annual hours worked in various categories of employment, generate socio-economic profiles of rural workers, and discuss labor market seasonality and job diversification.

4.1 Hours worked

Table 4.1 provides summary statistics at the baseline (2010) and endline (2019) of the rural individuals who work at least one hour in any category in at least one year, split into the four main categories: household farm labor, off-farm labor, casual labor (*ganyu*), and wage labor. Total labor is the sum of all four categories. These data suggest that the average individual works far less than their potential hours, especially on the family farm. Off-farm employment is able to offer a fuller work schedule, but participation in these jobs is significantly lower.

In Table 4.1 we see that at the baseline an average individual works 550 hours a year, which is equal to about 69 8-hour working days. 92 percent of individuals are involved in family farming. Conditional on working in an off-farm job, individuals work up to almost five times more hours in off-farm jobs than on their household farm: 270 hours/year on their household farm versus 371 hours/year in casual labor *(ganyu*), 710 hours/year in non-farm enterprises, and 1,265 hours/year in wage labor. However, while 92 percent of the individuals work on their household farm, the participation in off-farm jobs is much lower: 31 percent in *ganyu*, 12 percent in non-farm enterprises, and 8 percent in wage labor. The highest amount of conditional hours worked can be found in wage labor, where individuals at the baseline worked on average 1,265 hours/year, which is equal to 158 days of eight working hours or 32 40-hour work weeks.

Table 4.1. Annual labor allocation of rural workers at baseline and endline by employment category

	Baseline (2010)			Endline (2019)		
	Hours worked, average	Hours worked, conditional on working, avg.	Individuals that worked at least one hour, share	Hours worked, average	Hours worked, conditional on working, avg.	Individuals that worked at least one hour, share
Total Labor	549.6	573.7	0.96	756.4***	776.2***	0.97**
Household farm	247.3	269.6	0.92	142.4***	159.4***	0.89**
Household non-farm enterprise	85.8	710.1	0.12	216.3***	975.3***	0.22***
Agricultural	5.7	382.1	0.01	22.2***	538.1	0.04***
Non-agricultural	81.0	749.4	0.11	194.1***	1,054.3***	0.18***
Wage labor	100.7	1,265.4	0.08	99.0	1,151.6	0.09
Agricultural	25.4	1,116.1	0.02	26.0	1,354.7	0.02
Non-agricultural	72.3	1,303.7	0.06	73.0	1,093.2	0.07
Casual labor	115.7	370.9	0.31	298.6***	591.8***	0.50***

Source: Authors' analysis.
Note: Observations: 1,407.
Due to missing industry codes for 17 (22) individuals working in agricultural (non-agricultural) non-farm enterprises, the summary statistics for agricultural and non-agricultural non-farm labor might be biased.
T-test for difference from 2010 at significance level: *p<0.1; **p<0.05; ***p<0.01.

When we compare the baseline of 2010 with the endline of 2019, we see that the total hours worked increase significantly by 37 percent to 756 hours a year, or an increase of 26 8-hour working days. The increase in total hours worked stems from both *ganyu* and non-farm enterprise labor, while hours reported worked on the household farm decreased by 57 percent. Hours worked in wage labor remain roughly the same. *Ganyu* shows a particularly large increase between 2010 and 2019 with an increase in the participation rate of 19 percentage points and an increase in hours worked per year of 159 percent.

The group aged 25 to 45 years works more hours on average per year compared to the younger (<25 years) and older (>45 years) generations (Annex Table 4 in Annex D). The oldest generation spends more time on their household farm. Individuals aged under 45 years spend more time on and participate more in *ganyu*. Nevertheless, over the period of analysis, the hours worked in *ganyu* increased significantly for all age categories: they more than tripled for the youngest generation and increased 2.5 times for the oldest generation. The youngest generation spends the least time working in wage labor.

From Table 4.1 and Annex Table 4 in Annex D, we can conclude that, while we see differentials in hours worked between generations and over time, the relatively low hours worked overall per year indeed suggest that there is room for individuals to supply more hours. One important nuance, as explained in section 3.1, is that the ability of women to supply more hours in the four categories considered is significantly constrained by the large number of hours spent on household chores.

4.2 Socio-economic profiles of rural workers

Table 4.2 provides insights into the socio-economic profiles of the individuals who work in different employment categories. We regress a number of individual and household characteristics on a dummy indicating whether the individual is involved in an employment category in at least one of the two rounds, using a linear probability model (LPM):

$$P\left(Active_{i,y} = 1 \middle| X_{i,n}\right) = \beta_0 + \beta_n X_{i,n} + \varepsilon_i \quad (1)$$

with *Active*$_{i,y}$=1 if individual *i* worked at least one hour in employment category *y* in at least one round, 0 otherwise. $X_{i,n}$ is a set of socio-economic variables of interest consisting of sex, age, a dummy that indicates whether the individual went to school, a dummy that indicates whether the individual can read or write, a dummy that indicates whether the individual has higher education than primary school, and household size controls consisting of six categories based on age (0-5, 5-15, 15-65, >65 years of age) and sex. ε_i denotes the error term.

Table 4.2. Socio-economic profiles of individuals active in different categories

	Active on household farm	Active in household non-farm enterprise	Active in wage labor	Active in casual labor
Male	-0.0114	0.0295	0.1570***	0.1120***
	(0.0090)	(0.0262)	(0.0185)	(0.0273)
Age	0.0005	-0.0004	0.0010	-0.0085***
	(0.0004)	(0.0012)	(0.0008)	(0.0012)
Went to school	-0.0039	0.0430	0.0433*	-0.0432
	(0.0122)	(0.0352)	(0.0249)	(0.0367)
Can read or write	-0.0034	0.0500	0.0189	-0.0431
	(0.0107)	(0.0309)	(0.0219)	(0.0323)
Received education higher than primary school	-0.0493***	0.0937**	0.1630***	-0.2420***
	(0.0124)	(0.0359)	(0.0254)	(0.0375)
Constant	0.9990***	0.2510***	-0.0143	0.9930***
	(0.0191)	(0.0553)	(0.0391)	(0.0577)
Household size controls*	yes	yes	yes	yes
R-squared	0.034	0.033	0.128	0.086

Source: Authors' analysis.
Note: Coefficients come from LPM regression. Values in parentheses are standard errors.
Observations: 1,390. Seventeen individuals dropped from sample due to missing education variables.
*Eight household size controls based on four age categories (0-5, 5-15, 15-65, >65 years of age) for both sexes.
Significantly different from zero at level: *p<0.1; **p<0.05; ***p<0.01.

Men work significantly more in wage labor (16 percentage points) and casual labor (*ganyu*) (11 percentage points). In terms of education, people who work in non-farm enterprises and wage labor are significantly more educated on average than those who work on their household farm and in *ganyu*. In terms of age, we see that especially young people work in *ganyu*. Indeed, the generation aged less than 25 years spends about 29 percent of their working hours at baseline in *ganyu*, compared with 20 percent and 16 percent from the people aged 25 to 45 years and more than 45 years, respectively (Annex Table 4 in Annex D).

4.3 Seasonality

de Janvry, Duquennois, and Sadoulet (2022) document Malawian labor calendars in detail using 18,699 rural and 4,625 urban working-age adults. During the busiest two months, December and January, twice as many hours are worked compared to the least busy months in July and August. During the peak labor period at planting, a working-age adult will supply 24 hours of labor per week, while in the low-season that is only 12 hours. If rural dwellers in Malawi are not utilizing their most important asset, their labor, it is quite likely that they cannot find a market on which to sell their labor. Once the rains stop around April or May, there is a long dry spell until October or November during which there is very little scope for crop production, unless major investments in irrigation have been made.

Table 4.3 presents the labor supplied in different employment categories broken down in months and days for non-farm employment categories and in different agricultural activities and seasons for labor on the household farm. This table gives us insights into the intensity of labor supplied and the seasonality of the agricultural labor calendar. We see, for example, that, conditional on working, individuals working in wage labor work on average 8.7 out of the 12 months of the year, during which they work on average 3.9 weeks/month and 35.6 hours/week. This is very close to a full-time working schedule. When looking at casual labor (*ganyu*), we see that it is less intensive: individuals who worked at least one hour in *ganyu* over the past year did so on average over approximately 4.8 months, 2.8 weeks/month, and 4.4 days a week. Individuals working in non-farm enterprises did so on average over a period of almost 8 months, working approximately half of the days at 7.5 hours a day.

In contrast, when looking at household farm labor, which is recorded in the survey per plot, per activity, and per season (rainy and dry), we see that agricultural labor is indeed highly irregular. Especially for land preparation and planting tasks and for weeding, fertilizing and other non-harvest tasks during the rainy season, individuals supply a high number of working hours. An average household farm worker works, for example, 119 hours per year on land preparation and planting in the rainy season, which typically takes place in December and January. This is already more than half of the average conditional amount of total hours worked on the household farm for the entire sample, which is 215 hours per year. The hours supplied on the household farm in the dry season are only about half the amount of hours supplied in the rainy season for each of the three agricultural task categories.

Looking at hours worked per agricultural task category, we find additional evidence of seasonality. Significantly more individuals are employed for harvesting (82 percent) than for other tasks (67 percent), but conditional on working in the activity, preparing the fields and planting involves the highest number of hours worked. Furthermore, about 90 percent of all individuals is involved in farming during the rainy season, which is virtually everybody who worked at least one hour on the household farm all year round. In contrast, only 11 percent of individuals worked at least one hour in farming during the dry season.

These findings provide evidence that there is indeed irregularity as well as underemployment in the agricultural schedule, findings that are in line with de Janvry et al. (2022).

Table 4.3. Labor supplied in different employment categories

	Mean	Standard deviation	Individuals that worked	Conditional mean	Maximum
Household Farm					
Rainy season[1]					
Hours/year land preparation, and planting	78.38	110.23	1,846 (66%)	119.48	1,536
Hours/year weeding, fertilizing & other non-harvest tasks	61.15	82.40	1,851 (66%)	92.96	1,296
Hours/year harvesting	45.02	68.79	2,275 (81%)	55.68	852
Total hours worked HH farm, rainy season	184.55	206.26	2,545 (90%)	204.06	2,405
Dry season					
Hours/year land preparation, and planting	4.90	21.23	304 (11%)	45.33	406
Hours/year weeding, fertilizing & other non-harvest tasks	3.38	19.02	295 (10%)	32.21	504
Hours/year harvesting	2.24	16.20	209 (7%)	30.14	414
Total hours worked HH farm, dry season	10.51	46.97	318 (11%)	93.02	915
All year					
Hours/year land preparation, and planting	83.28	113.96	1,874 (67%)	125.05	1,536
Hours/year weeding, fertilizing & other non-harvest tasks	64.53	86.15	1,873 (67%)	96.94	1,296
Hours/year harvesting	47.26	70.52	2,302 (82%)	57.77	852
Total hours worked on household farm	194.89	214.75	2,548 (91%)	215.23	2,405
Household Non-farm Enterprise					
Months/year first non-farm enterprise (NFE)	1.21	3.30	423 (15%)	8.03	12
Days/last month first NFE	2.13	6.33	423 (15%)	14.15	31
Hours/day last month first NFE	1.12	2.92	423 (15%)	7.47	16
Total hours worked in first NFE	140.30	505.03	423 (15%)	933.33	3,650
Months/year NFE, average	2.26	6.52	443 (16%)	14.35	31
Days/last month NFE, average	1.17	2.95	443 (16%)	7.41	16
Hours/day last month NFE, average	1.25	3.33	443 (16%)	7.91	12
Total hours worked in NFE	145.46	508.98	443 (16%)	923.96	3,650
Wage Labor					
Months/year main wage job	0.72	2.65	233 (8%)	8.74	12
Weeks/month main wage job	0.32	1.08	233 (8%)	3.87	4
Hours/week main wage job	2.95	11.95	233 (8%)	35.60	84
Total hours worked in main wage job	100.20	443.91	233 (8%)	1,210.18	3,650
Casual Labor					
Months/year casual labor	1.97	3.30	1,149 (41%)	4.81	12
Weeks/month casual labor	1.14	1.54	1,149 (41%)	2.79	4
Days/week casual labor	1.78	2.36	1,149 (41%)	4.36	7
Total hours worked in casual labor	207.20	420.86	1,149 (41%)	507.44	2,688

Source: Authors' analysis.

Note: [1]Rainy agricultural season in Malawi varies spatially but in general refers to the period of November to May, while the dry season runs from June to October.
Observations (pooled): 2,814.

4.4 Job diversification

Table 4.4 and Table 4.5 give us a clear view of job diversification in our sample. Table 4.4 provides an overview of the number of different employment categories in which each individual worked. For the pooled sample, 43 percent held a job in one category, while 48 percent held jobs in two different categories in the previous year. If we compare both rounds however, we see that in 2019 significantly more people worked in two or more employment categories. We also see that across the whole sample people who hold more than one job work significantly more hours—people who worked in two employment categories work more than double the amount of hours per year than people only active in one. Rural workers active in three or four categories work another 50 percent more hours than those active in only two categories.

Table 4.4. Average total hours worked annually by number of employment categories in which individual worked

Employment categories in which worked, number out of four	Pooled sample		2010 (Baseline)		2019 (Endline)	
	Observations	Hours per year, avg.	Observations	Hours per year, avg.	Observations	Hours per year, avg.
None	95 (3%)	0.0	59 (4%)	0.0	36 (3%)**	0.0
One	1,206 (43%)	350.0	745 (53%)	353.4	461 (33%)***	344.6
Two	1,340 (48%)	883.0	546 (39%)	812.5	794 (56%)	931.5***
Three	166 (6%)	1,340.5	53 (4%)	1,141.1	113 (8%)***	1,434.0*
Four	7 (0.2%)	1,381.5	4 (0.3%)	1,482.7	3 (0.2%)	1,246.7
Observations	2,814		1,407		1,407	

Source: Authors' analysis.
Note: Four employment categories: household farm, household non-farm enterprise, wage labor, and casual labor.
T-test for difference from 2010 at significance level: *p<0.1; **p<0.05; ***p<0.01.

Table 4.5 offers insights in how those working hours in different employment categories are spread out. The values in the first row of each grouping of three rows shows the percentage of individuals working in the employment category (category y) listed in the column header of the table, who also work in the employment category (category x) listed in the row header of the table. The diagonal cells indicate the proportion of people working *only* in the given employment category. The second number in each grouping of three rows represents hours worked in employment category y conditional on working at least one hour in both employment category y and x. The third number in each grouping of three rows is the total hours worked , conditional on working at least one hour in both categories. For the cells at the intersection of the same employment category, only a single value for hours worked is reported, as the two estimates are the same, by definition.

Table 4.5. Participation and hours worked across different employment categories, pooled sample

Are also active or only active in:	Are active in:	Household farm	Household non-farm enterprise	Wage labor	Casual labor
Household farm	Individuals active in above category, also active in category at left, %	41	89	80	92
	Hours worked annually in above ctgy, if also worked in category at left	238.0	849.5	1,092.0	492.9
	Total hours worked annually in the two employment categories		1,194.7	1,397.2	808.9
Household non-farm enterprise	Individuals active in above category, also active in category at left, %	17	9	10	11
	Hours worked annually in above ctgy, if also worked in category at left	192.5	1,269.1	937.6	413.8
	Total hours worked annually in the two employment categories	1,194.7		1,715.9	1,257.1
Wage labor	Individuals active in above category, also active in category at left, %	7	5	16	5
	Hours worked annually in above ctgy, if also worked in category at left	174.8	505.9	1,678.7	284.4
	Total hours worked annually in the two employment categories	1,397.2	1,715.9		1,458.5
Casual labor	Individuals active in above category, also active in category at left, %	42	27	22	6
	Hours worked annually in above ctgy, if also worked in category at left	205.5	644.0	926.1	757.3
	Total hours worked annually in the two employment categories	808.9	1,252.6	1,458.5	

Source: Authors' analysis.
Observations: 2,814.

Of the people working on the household farm, 41 percent only engaged in household farm labor, 17 percent also engaged in a non-farm enterprise, 7 percent in wage labor, and 42 percent in casual labor (*ganyu*). Those only working on the household farm work on average 238 hours a year on the farm—of course, for these individuals, this also equals their total hours worked. Those working both on the household farm and *ganyu* work on average 205 hours on the farm and 809 hours in total, and so on.

From Table 4.5, we see that hours worked on the household farm are quite independent of whether someone is engaged in another employment category or not. People who solely rely on, or combine, household farming and *ganyu* (the top right cells in Table 4.5) work the lowest number of hours per year. Outside that combination of employment categories, among people who also rely on non-farm self-employment and wage labor, total hours worked are noticeably higher. Individuals only engaged in household non-farm employment and wage employment are able to achieve hours worked that come very close to full employment (31 40-hour work weeks and 41 40-hour work weeks annually, respectively), something which does not seem possible with farm labor or *ganyu* alone. The numbers from Table 4.4 and Table 4.5 thus show that engaging in labor outside of the household farm is an important strategy to raise the amount of hours worked.

5 ESTIMATION STRATEGY

In the previous section, we documented large and widespread levels of rural underemployment. For those farming their own household plots, the seasonal agricultural calendar restricts the amount of labor that they can supply throughout the year. Only rural workers who work in at least one additional employment category realize higher hours of work. Those engaged in household non-farm enterprises and wage labor have the lowest levels of underemployment. Off-farm employment might therefore be important to put labor, one of the most important assets of the rural poor, to full utilization.

In the remainder of this paper, we will investigate if, and how, nearby urbanization can provide opportunities for increased labor supply by rural individuals. This section explains our empirical strategy to do so. We start by describing our urban access variable that measures how exposed a rural area is to urban areas around it. We then describe how we relate urban access to hours worked in a panel set-up. The next section will provide the results of this analysis.

5.1 Urban access

Our urban access variable is a measure of urban influence on a rural area. For each rural enumeration area *k*, this urban access variable (UA) is the weighted sum of the population size of the surrounding urban areas, with the weights inversely proportional to the distance between the rural area and the urban areas:

$$\mathrm{UA}_{\mathrm{k,t}} = \sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{size}_{\mathrm{j,t}} * \mathrm{dist}_{\mathrm{k,j}}^{-\beta} \qquad (2)$$

with $size_{j,t}$ being the population of agglomeration *j* at time *t*, $dist_{k,j}$ being the distance between rural location *k* and the nearest boundary of agglomeration *j*, and β being a discount factor that weights this distance.[10] The larger the β, the less urban influence further away agglomerations are assumed to have on a rural area. This urban access variable is similar to classic ‘accessibility’ indicators derived from a gravity model, which is often used in studies assessing interaction between entities based on their distance as well as size. This interaction is often trade, but also consumption, migration, financial flows, and flows of knowledge have been studied using some form of this gravity model (Bigman & Fofack, 2000; Gutiérrez, Condeço-Melhorado, & Martín, 2010). For our specifications we use β=1, which is the standard distance discount value in gravity models (Head & Mayer, 2014). In our analysis, we will use one year lagged urban access to warn against reverse causality (see section 6.4 for a reflection on causality). In Table 5.1 you can find the summary statistics of the one year lagged urban access variable for baseline, endline and changes over time.

[10] For the enumeration areas that are classified as rural, but which fall inside the boundary of an Africapolis agglomeration because of the randomly imposed offset imposed on household locations in the IHSP dataset for confidentiality reasons (National Statistical Office 2012), we set the minimum distance to one km.

Table 5.1. Urban access, summary statistics

	Mean	Std. dev.	Min	Max
Urban access, 2009	50247.61	104618.1	5639.926	651978.9
Urban access, 2018	70988.15	154029.3	7985.099	1004688
∆ Urban access 2009-2018	20740.53	52131.74	2345.172	375098.9
Urban access, logged, 2009	10.27576	.8202392	8.637627	13.38777
Urban access, logged, 2018	10.59228	.832271	8.985332	13.82019
∆ logged Urban access 2009-2018	.3165215	.0762224	.1967363	.4673653

Source: Authors' analysis.
Observations: 1407.

5.2 Regression specification

Armed with measures of rural labour allocation in each rural enumeration area (EA) and urban access for those same EAs, we can now explore how changes in urban access are associated with changes in labour patterns. To examine the link between urban growth over time and changes in hours worked in different employment categories, we use a first-difference regression specification:

$$\Delta l_{i,k,t} = \alpha_0 + \beta_1 \Delta lnUA_{k,t-1} + \beta_n c_{i,n,k,t-1} + \varepsilon_{i,k,t} \tag{3}$$

with $\Delta l_{i,k,t}$ being the change in labour supply over the period between *t*-1 and *t* for individual *i* living in rural area *k*, $\Delta lnUA_{k,t-1}$ being a measure of change of urban access in rural area *k* between *t*-2 and *t-1*, $c_{i,n,k,t-1}$ being a set of *n* baseline (individual) characteristics, and $\varepsilon_{i,k,t}$ as a random error term.[11] For ease of interpretation, we use a level-log specification estimated with OLS.

We are concerned that unobserved heterogeneity can be spatially correlated. That spatial correlation may not be restricted to observations from the same village, the primary sampling unit. Some of our clusters lie close enough to each other to worry about spatially correlated errors across individuals from other (nearby) sampled villages.[12] We draw a circle with a 50km radius around each unit and follow Colella, Lalive, Sakalli, and Thoenig (2019)—making use of their user-written Stata command *acreg*—to allow for clustering of errors of all units that fall within overlapping circles. The weights are one for individuals in the same village and then decrease linearly with distance from one to zero for individuals from other villages that lie within the 50km circle. Weights of zero are assumed for individuals outside the circle.

[11] The variables used as baseline controls are age, male dummy, went to school dummy, can read or write dummy, and eight household size controls based on four age categories (0-5, 5-15, 15-65, >65 years of age) for both sexes.

[12] The average minimum distance between two rural clusters is 19.6 km. 64 percent of all EA's have at least one other EA at 20km distance, which increases to 92 percent for 50km distance. See Annex Table 9 in Annex D.

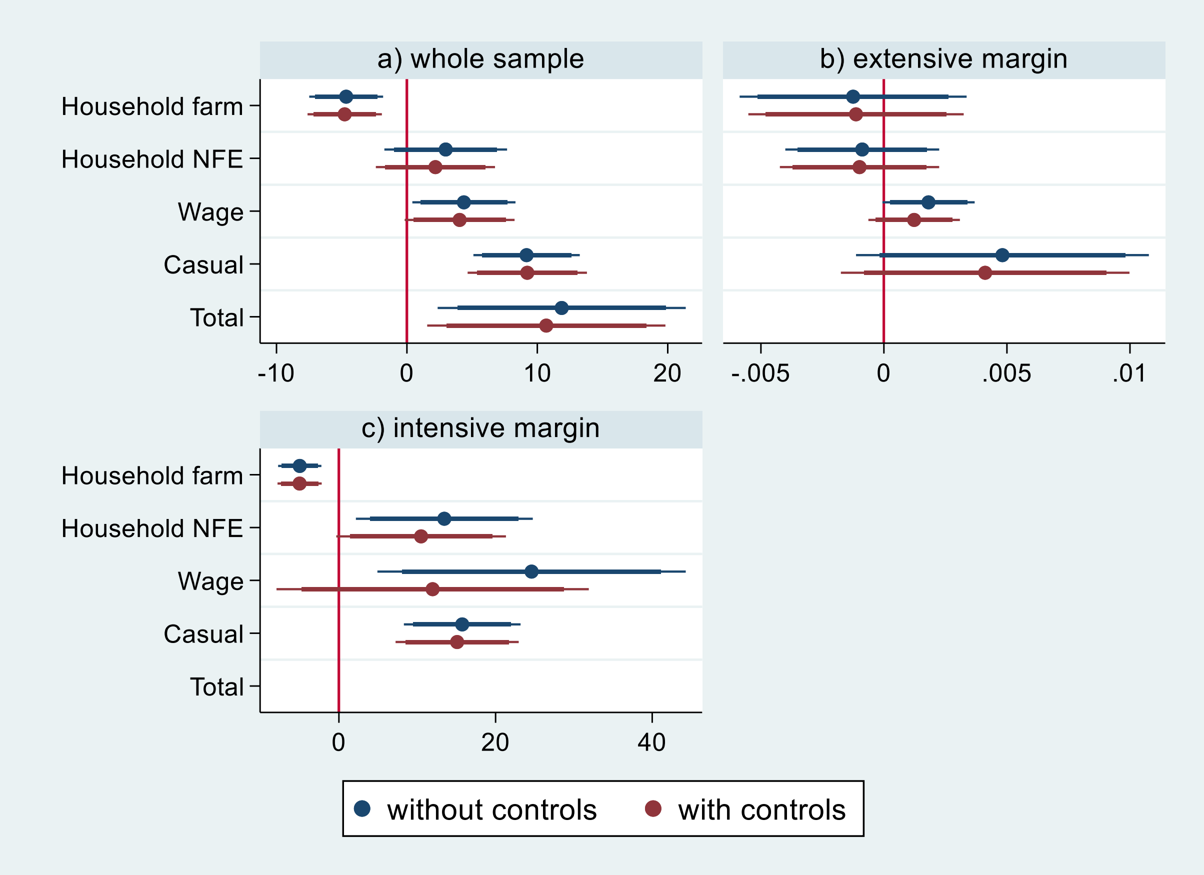
6 RESULTS

6.1 Main results

Panel a) of Figure 6.1 presents point estimates of β_1 from Equation (3) and their corresponding 95 percent confidence intervals, from five different regressions, with and without the control variables. Each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various employment categories. Due to the level-log specification, the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked. It is useful to keep in mind, as a point of reference, the median increase in urban access is 32 percent for the whole sample over the period between 2010 and 2019.

We see that urban growth has a net positive effect on total hours worked. An individual who experienced a 10 percent increase in urban access will, on average, work about 110 hours—about 14 days—more a year. The positive effect of urbanization on total hours worked comes both from casual labor (*ganyu*) and from wage labor. A 10 percent increase in urban access increases hours worked in *ganyu* by 92 hours a year and in wage labor by 44 hours a year. However, increasing urban access has a negative effect on hours worked on the household farm. An individual who experiences a 10 percent increase in urban access will work on average 48 hours per year less on their household farm. That own farm labor goes down as wage labor and *ganyu* goes up could indicate a constraint on total hours worked, for example because of household chores, which are disproportionately done by women (Section 3.1). Section 6.3 looks at the heterogeneity of the effects by gender. Another explanation might be that urbanization increases the demand for farm land from nearby urban residents who complement their urban-based activities with farming.

Figure 6.1. Effect of urbanization between 2010 and 2019 on hours worked



Source: Authors' analysis.
Note: This figure represents the point estimates of β_1 from Equation (3) and their corresponding 95% confidence intervals. In panel a), each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in the various employment categories.
Panel b) estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin).
Panel c) restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories.
Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a) and panel b): 1,407 for estimations without controls, 1,390 for estimations with controls as seventeen individuals dropped from sample due to missing education variables. Panel c): 1,371/1,356 for household farm labor, 400/401 for household non-farm enterprise labor, 184/181 for wage labor and 871/862 for casual labor, for estimations with/without controls.

Panels b) and c) of Figure 6.1 split those results up into their intensive and extensive margins. First, in panel b), Equation (3) is estimated as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin). The coefficients are signed consistently with the results for total hours worked from panel a) of Figure 6.1, except for household non-farm employment, which switches sign. However, the standard errors are very large and none of the effects are significant at the 95 percent level. It does not seem that the main driver of the increase in hours worked by rural individuals between 2010 and 2019 comes from people entering new categories of work.

Panel c) in Figure 6.1 restricts the sample for each regression to people who have worked at least one hour in the employment category in question at either baseline or endline. The coefficients therefore reflect the effect of urban access on hours worked, conditional on a person being active in that category (the intensive margin). We see that increased urban access affects hours worked in the same direction as in the main specification, but the effects are more pronounced. It seems, therefore, that urbanization positively affects hours worked for those individuals already working in a specific (non-farm) sector.

6.2 Heterogeneity in urban growth

Urban growth comes in two flavors. The first form is the expansion of existing urban agglomerations, while a second form involves previously rural areas emerging as new urban centers. According to the Africapolis database, 46 new agglomerations emerged in Malawi between 2000 and 2010, to total 65 in 2010. Between 2010 and 2015, 12 new agglomerations emerged. New urban agglomeration have, in addition to the population size effect, also an effect on distance: for each new urban town that arises, the average rural dweller will be in closer proximity to an urban center, enhancing various rural-urban linkages (De Weerdt et al., 2021). To get an idea of the relative importance of the growth of existing urban agglomerations against the emergence of new agglomerations, we can split the change in urban access in two components: the growth stemming from the 'older' agglomerations—those 19 agglomerations that already existed before 2000—and the growth stemming from the 58 new agglomerations that only emerged after 2000.[13] We run regression (3) separately for the change in urban access for, on the one hand, the subset of old agglomerations and, on the other, the subset of new agglomerations.[14] That way, urban access as specified in (2) is calculated as the discounted sum of the 19 (58) agglomerations belonging to the old (new) agglomeration classification.

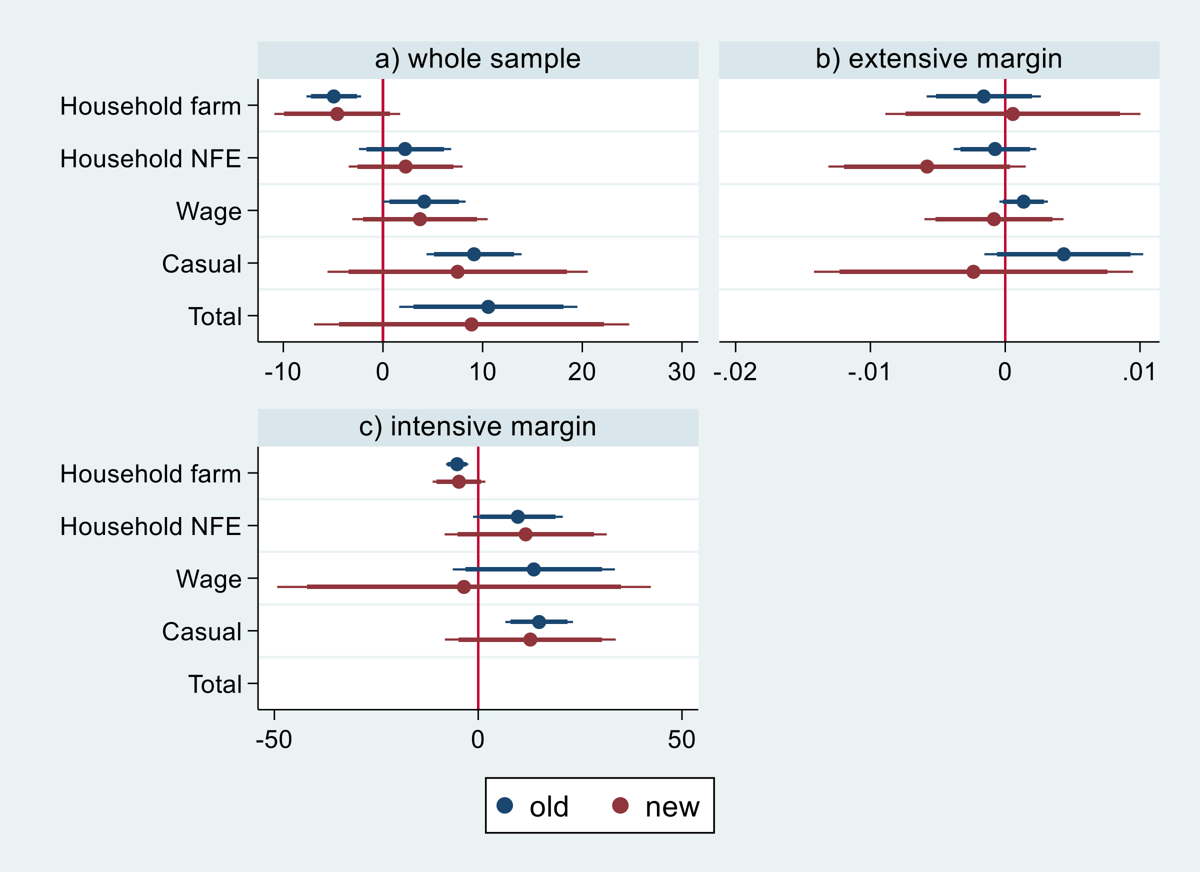
Figure 6.2 shows the results for these two separate regressions. The results show that both growth in old agglomerations and new agglomerations are signed consistently with what we found in the baseline specification (Figure 6.1). However, the standard errors for the coefficients on growth in new agglomerations are large, so emerging new agglomerations do not seem to have a significant effect on hours worked in any category. Note that the total growth in urban access stemming from old agglomerations is almost twice as large as the growth in new agglomerations. When we do run the model with both UAs included in the same regression, we get similar coefficients for old cities, while the coefficients for new towns remain insignificant[15] and sometimes switch signs (see Annex Figure 9 in Annex D). Taken together, we can conclude that the associations between urbanization and labor markets identified in Figure 6.1 are mainly driven by larger cities.

[13] Those 19 well-established, 'older' agglomerations are Balaka, Blantyre, Dedza, Karonga, Kasungu, Lilongwe, Liwonde, Lumbadzi, Mangochi, Mchinji, Mponela, Mulanje, Mzimba, Mzuzu, Nkhotakota, Nsanje, Rumphi, Salima, and Zomba.

[14] We opt for running the regressions separately for two main reasons. The change in these two separate definitions of urban access are highly correlated, with a correlation coefficient of 0.58, and we do not want to control for the part of the effect of new agglomerations that happens through cities (or vice versa). The results of the alternative specification with both subsets in the same regression can be found in Annex Figure 9 in Annex D.

[15] Except for casual labor (*ganyu*) on the extensive level, which becomes significantly negative for new agglomerations.

Figure 6.2. Effect of urbanization between 2010 and 2019 on hours worked, for 'old' and 'new' agglomerations



Source: Authors' analysis.
Note: This figure represents the point estimates of β_1 from Equation (3) and their corresponding 95% confidence intervals, estimated with controls. Urban access is split into two components: growth stemming from the 'older' agglomerations—those 19 agglomerations that already existed before 2000 ('old')—and growth stemming from the 58 new agglomerations that only emerged after 2000 ('new'). The regressions are run separately for each of the two components.
In panel a) each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various aggregations of employment categories.
Panel b) estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin).
Panel c) restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories.
Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a) and panel b): 1,390. Seventeen individuals dropped from sample due to missing education variables. Panel c): 1,356 for household farm labor, 400 for household non-farm enterprise labor, 181 for wage labor and 862 for casual labor.

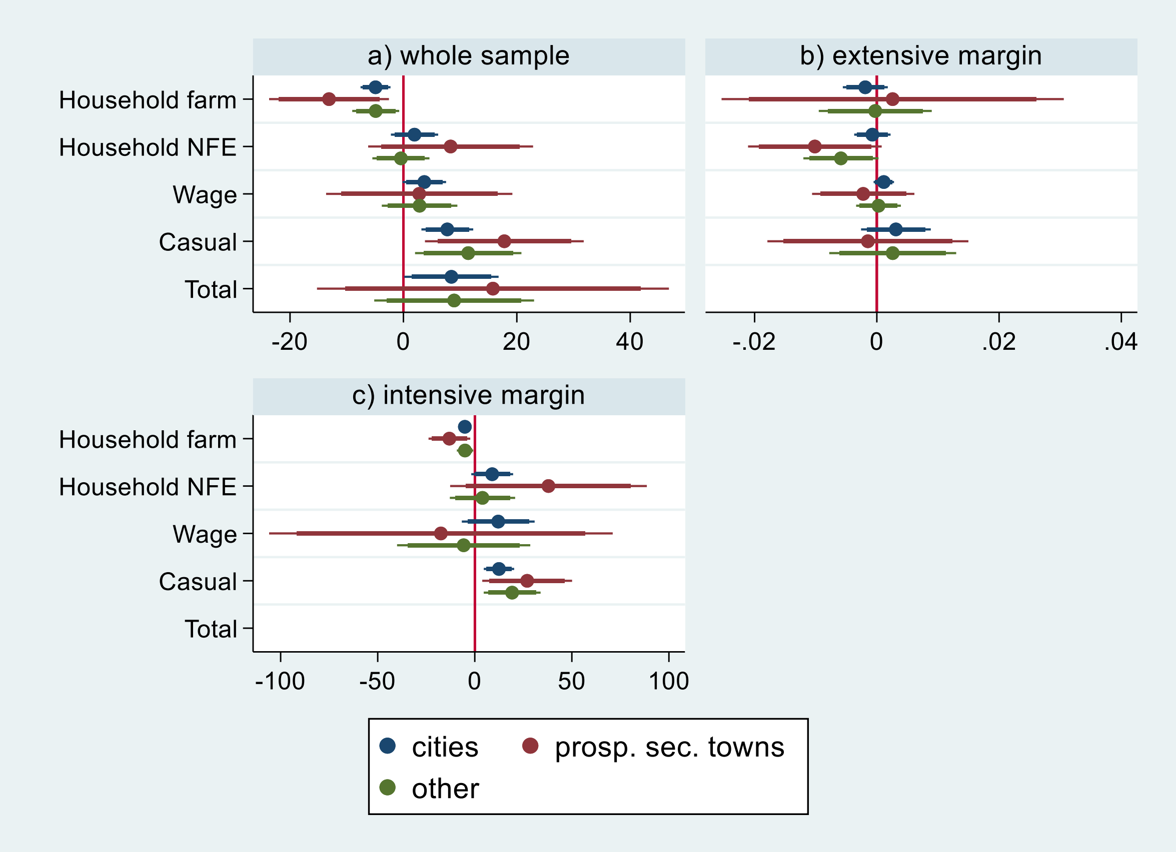
Another way to decompose the effect of urban growth stemming from different types of cities is taking an administrative approach. Malawi currently has four cities. However, the government of Malawi has identified urbanization as one of the three pillars of their most recent long-term development plan, called Malawi 2063 (MW2063). In light of this, they have identified eight urban agglomerations to become secondary cities, based on their (potential) level of activity in terms of governance, industry, agriculture, tourism, and mining (National Planning Commission, 2020). This allows us to split the agglomerations in our sample into three categories: the four cities (Mzuzu, Zomba, Blantyre, Lilongwe), the prospective secondary cities (Karonga, Kasungu, Nkhata Bay, Chipoka, Liwonde, Luchenza, Bangula),[16] and other urban agglomerations. It is worth noting that agglomerations identified as having potential for prospective secondary towns are not necessarily the 'bigger' towns. Annex Table 8 in Annex D lists the agglomerations by these three categories together with their 2019 populations as estimated using WorldPop.

We run regression (3) for the change in urban access for the subset of cities, prospective secondary towns and other agglomerations separately.[17] The results for these three subsets of agglomerations are presented in Figure 6.3. We see that cities have an effect on labor similar to the general effect of urban growth as shown in our baseline specification (Figure 6.1). In prospective secondary towns, the negative effect on household farm labor and the positive effect on hours worked in casual labor (*ganyu*) is especially pronounced—it is about twice as large as the effect of growth in the four cities, although also measured more imprecisely. Growth in the 66 other agglomerations also has a significant negative effect on household farm labor and a significant positive effect on *ganyu* —both about the size of the general effect.

[16] Monkey Bay is also identified as a prospective secondary town by the government of Malawi, but it is not recognized as an agglomeration in the Africapolis database, so we are not able to include Monkey Bay in our analysis.

[17] The coefficient of correlation between urban access growth stemming from cities and prospective towns is 0.57, 0.70 for cities and other agglomerations, and 0.72 for secondary towns and other agglomerations.

Figure 6.3. Effect of urbanization between 2010 and 2019 on hours worked, for three types of agglomerations: cities, prospective secondary towns, and other (smaller) agglomerations



Source: Authors' analysis.
Note: This figure represents the point estimates of β_1 from Equation (3) and their corresponding 95% confidence intervals, estimated with controls. Urban access is split into three components: growth stemming from cities ('cities'), prospective secondary towns ('prosp. sec. towns') and other urban agglomerations ('other'). The regressions are run separately for each of the three components.
In panel a) each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various aggregations of employment categories.
Panel b) estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin).
Panel c) restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories.
Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a) and panel b): 1,390. Seventeen individuals dropped from sample due to missing education variables. Panel c): 1,356 for household farm labor, 400 for household non-farm enterprise labor, 181 for wage labor and 862 for casual labor.

When looking at the extensive margin (panel b) in Figure 6.3, we see that growth in prospective secondary towns and smaller agglomerations have a negative effect on participation in non-farm enterprise labor, significant at the 90% level, although again with large standard errors. Looking at the effect on the intensive margin (panel c) in Figure 6.3, we see that growth in cities increases hours worked in non-agriculture, while growth in secondary towns and other agglomerations increases hours worked in *ganyu*.

Annex Figure 10 in Annex D shows the results for when all three urban access components are included in the same regression.[18] The coefficients on the growth in cities remain roughly the same when including all three components in the same regression. For towns and other agglomerations the positive effect on hours worked in *ganyu* disappears, both in the general specification and on the intensive margin. Growth in other agglomerations becomes negatively correlated with hours worked in non-farm enterprises. This again suggests that cities are the main driving force of the associations between urbanization and labor markets we showed in Figure 6.1.

Taken together these results suggest that growth in big cities induces more hours of work off the household farm, which, as shown in section 4, is key to boosting total hours worked. In the concluding discussion (Section 7), we will discuss the relevance of these results for policy.

6.3 Heterogeneity: sex, age, and education

The positive effect of urbanization on hours worked may differ with sex, age, education and other individual characteristics. For example, in Malawi the production of household goods, like water and firewood collection, childcare, cooking and the like, fall disproportionately on women, potentially imposing additional constraints to their ability to increase hours worked when urban access increases. de Janvry et al. (2022) calculate that women spend 23 hours a week on such activities, while men spend only 4. As detailed in section 3.1, we found for our sample that women spend about 10 times longer on collecting firewood and fetching water than men. Alternatively, urban access may provide more opportunities for women because they start from a lower baseline participation rate. Similarly, education and age may affect the ability to supply more hours of work.

To test for these effects, we include interaction terms in our regressions (see Annex Table 5, Annex Table 6, and Annex Table 7 in Annex D). While urban growth has no differential effects in terms of sex, age, and educational status for the general specification, we do see some interesting effects on the extensive margin. The effect of urban growth on the propensity to move into casual labor (*ganyu*) is almost six times larger for women than for men: a 25 percent increase in urban access increases the female participation rate in *ganyu* with 17 percent, compared to under 2 percent for men. We know from Table 4.2 that at the baseline men were significantly more likely to participate in *ganyu,* so urban growth seems to have an equalizing effect with respect to participation in *ganyu*. We also find evidence that as people get older, urban growth is less likely to lead to more *ganyu* and more likely to reduce work on their household farm. Finally, we find that younger people are more likely to move into wage labor.

6.4 Direction of causality

We have shown how changes in urban access are associated with changes in hours worked in rural labor markets, but what can we say about the causality of the effects? Asher, Campion, Gollin, and Novosad (2022) show how canal building led to large increases in agricultural productivity in India and how this, in turn, spurred nearby urbanization. However,

[18] On the intensive margin in this alternative specification, the effect of growth in cities decreases for *ganyu*, but becomes positively statistically significant for wage labor.

other research has suggested effects that go the other way. Calì and Menon (2013); Gibson et al. (2017) find a positive effect of urban growth on rural poverty reduction in India. Gibson et al. (2017) find that it is especially growth in towns (and not necessarily cities) that reduces rural poverty, while Calì and Menon (2013) show that the causal poverty-reducing effect of urban growth comes mainly through rural economic growth and not necessarily from rural-urban migration. Binswanger-Mkhize et al. (2016) show for Kenya a positive effect of urban growth on agricultural and rural non-farm income growth, mainly through its effect on education and commercialization. Vandercasteelen et al. (2018); Vandercasteelen et al. (2021) investigate the effect of urbanization on farmer's behavior in Ethiopia and find a positive effect on agricultural productivity in the teff and dairy sector.

Although we have no iron-clad proof, we believe that an important channel of causality runs from urbanization to rural labor markets. First, the period we have studied has been one of stagnating agricultural productivity in Malawi so that the particular channel documented by Asher et al. (2022) is unlikely to hold in our context.[19] Second, we use one-year-lagged urban growth in all our specifications. We also perform a robustness check in which we lag urban growth by two years and find the results to be robust to this. Third, we add controls for population growth in the rural area. If rural population growth is triggering nearby urban growth we expect to see a reduced effect once rural growth is controlled for. The results are not consistent with this and remain robust to including rural growth. Annex B elaborates on these robustness checks.

7 CONCLUDING DISCUSSION

Rural labor markets in Malawi are characterized by underemployment and significant amounts of seasonality. A farming household operating in such an environment could increase income by diversifying agricultural production, investing in technology to reduce dependence on seasonal rainfall, diversifying into off-farm activities, or engaging in seasonal or permanent migration of some of its members. But if the supply of labor is restricted by the imperative of farming to ensure subsistence and the irregularity of the agricultural schedule, then the challenge is for the labor demand to fit that schedule.

In this paper we constructed a measure for urban access. This measures the extent to which a rural individual is exposed to urban areas, combining population size of surrounding urban areas discounted by the distance to them. We then look at how labor supply changes when urban access changes for a nine-year panel of 1,407 rural working age individuals. We find that increasing urban access by 10 percent leads to 14 extra days of work. Urbanization comes with increased hours of work in casual labor (*ganyu*) and in non-agricultural sectors, but at the expense of work on the household farm, an effect that is more pronounced for younger people. Increased hours worked in off-farm enterprises and, to some extent, in wage labor, drive the results for non-agricultural labor. These effects are primarily at the intensive rather than the extensive margin. In other words, we mainly see people who are already taking part in an activity providing more hours to that activity, but very little new entry.

[19] Caruso and Cardona Sosa (2022) show that value-added yield, defined as the value of the harvest minus total costs, and total factor productivity have increased but very slowly over the period 2010-2019. Benson (2021) shows that growth in agriculture is volatile: it fell six times from year to year between 2000 and 2019.

We found that the linkages with urban labor markets are strongest for cities and long-established agglomerations, compared to the newly emerged ones. This is an important point for policy in Malawi because there are several dozen emerging new urban agglomerations, which all hold potential to provide opportunities to nearby rural areas. Policies and investments that promote linkages between newly emerging urban centers and their rural hinterlands are important to ensure the benefits of urbanization are geographically spread widely. A promising plan in this respect formulated in Malawi 2063 (National Planning Commission, 2020), the document outlining Malawi's development visions and aspirations, is to transform eight current urban centers into an interconnected network of secondary cities to complement the existing four major cities of Lilongwe, Blantyre, Mzuzu, and Zomba. The proposed secondary cities are already spread along the country's north-south axis, which can help to ensure an equitable geographical distribution of the benefits of urbanization. Investments in these cities will be anchored in economic activities that are locally appropriate, such as value addition for cities that lie in areas with high agricultural potential, fisheries for those by the lake, and hubs for cross-border trade for those connected to borders. Our data show that over the past ten years the effect of growth in the proposed sites on the labor markets around them could not be precisely measured. The economic anchoring of these cities provides an opportunity for policy makers to explicitly take impacts on employment in their surrounding rural areas into consideration.

At baseline in 2010-11, we see that men and youth work more in non-agricultural activities and *ganyu*, and that people with higher levels of formal education work more in non-agricultural jobs. While women start from very low levels of participation in *ganyu* at baseline, they are six times more likely to switch into *ganyu* than men due to urban growth. The effect of urban growth on rural labor supply is also larger for young people. Clearly urban growth has broad-based effects on labor participation across socio-economic groups in rural areas.

A particularly striking finding is the strong effect of urbanization on increasing *ganyu*. As a flexible labor agreement that is steadily increasing over time, it plays an important role in job diversification and is positively affected by nearby urbanization. While *ganyu* might offer flexible informal labor contracts, its increasing importance might also be a sign of rural distress, income insecurity, or tensions with household farm labor (Bryceson, 2008). We find that combining *ganyu* with own-farm labor is common, but does not lead to a full working schedule, on average. Further exploration of the scope and relevance of *ganyu* in current day Malawi is needed to clarify these issues.

It is important to bear in mind that our results come with a number of caveats. First, our definition of labor is narrowly defined to income generating activities and we cannot take into account labor focused on production of household goods. Second, although our assumption is that working more hours leads to higher incomes in a context where underemployment is pervasive, we do not have information on wages, prices, and consumption that would be needed to calculate the welfare effects. If households are working more for less pay, then they may be becoming what Bick, Fuchs-Schündeln, and Lagakos (2018) call 'both leisure poor and consumption poor'. The data needed to conduct this analysis are not yet available.

Steady urbanization is currently one of the most important demographic trends in Africa, and its effects on rural areas, where the majority of the poor live, is poorly understood. This lacuna in our knowledge on the link between the urbanization process and rural transformation holds the danger of foregoing policy opportunities to ensure urban growth benefits rural poverty reduction. Understanding urbanization in a spatial context, particularly in its relation to rural areas, is key to building effective institutional and policy frameworks to guarantee the process is advantageous for the economy and society (Bloom, Canning, Fink, Khanna, & Salyer, 2010).

In this paper we have shown that, on the one hand, urbanization provides opportunities to increase hours worked, but, on the other, that the effect is primarily driven by increases in casual labor and has not come with any significant movement away from agriculture. For now, we do not observe structural transformation in the classical sense in Malawi, but the country is only at the beginning of its urbanization process with scope for policy to leverage future urbanization for inclusive development.

ABOUT THE AUTHORS

Joachim De Weerdt is a senior research fellow with the Development Strategy and Governance Unit of the International Food Policy Research Institute (IFPRI). He is based in Lilongwe where he heads IFPRI's Malawi Country Office.

Hanne Van Cappellen is a PhD Candidate in the Institute of Development Policy at the University of Antwerp in Belgium.

ACKNOWLEDGMENTS

We are grateful for comments on earlier drafts by Todd Benson and Jan Duchoslav. Financial support for this work was graciously provided by the Government of Flanders and the Irish Embassy in Malawi.

REFERENCES

Africapolis. 2022. Country Report: Malawi. Retrieved from https://africapolis.org/en/country-report/Malawi

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:10.1016/j.foodpol.2018.07.005

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:10.1016/j.jdeveco.2017.10.005

Asfaw, S., Orecchia, C., Pallante, G., & Palma, A. (2018). *Soil and nutrients loss in Malawi: an economic assessment*. Rome, Italy: FAO, UNEP and UNDP.

Asher, S., Campion, A., Gollin, D., & Novosad, P. (2022). The Long-Run Development Impacts of Agricultural Productivity Gains: Evidence from Irrigation Canals in India.

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:10.1016/S0306-9192(01)00014-8

Beall, J., Guha-Khasnobis, B., & Kanbur, R. (2010). *Urbanization and development: multidisciplinary perspectives*: Oxford University Press.

Beegle, & Christiaensen, L. (2019). *Accelerating poverty reduction in Africa*. Washington, DC: World Bank Publications.

Beegle, 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. doi:https://doi.org/10.1162/REST_a_00105

Benson, T. (2021). *Disentangling food security from subsistence agriculture in Malawi*. Washington, DC: International Food Policy Research Institute (IFPRI).

Benson, T., & De Weerdt, J. (2023). *Employment options and challenges for rural households in Malawi: An agriculture and rural employment analysis of the fifth Malawi Integrated Household Survey, 2019/10* (Vol. 40): International Food Policy Research Institute (IFPRI).

Bick, A., Fuchs-Schündeln, N., & Lagakos, D. (2018). How Do Hours Worked Vary with Income? Cross-Country Evidence and Implications. *American Economic Review, 108*(1), 170-199. doi:10.1257/aer.20151720

Bigman, D., & Fofack, H. (2000). *Geographic Targeting for Poverty Alleviation: Methodology and Applications*. Washington, D.C.: World Bank.

Binswanger-Mkhize, H. P., Johnson, T., Samboko, P. C., & You, L. (2016). *The impact of urban growth on agricultural and rural non-farm growth in Kenya*. Paper presented at the 5th International Conference of the African Association of Agricultural Economists, Addis Ababa, Ethiopia. http://ageconsearch.umn.edu/record/249274/files/329.%20Urban%20growth%20on%20agricultural%20and%20non-farm%20income%20in%20Kenya.pdf

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.

Brenner, N., & Schmid, C. (2014). The 'Urban Age' in Question. *International Journal of Urban and Regional Research, 38*(3), 731-755. doi:10.1111/1468-2427.12115

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

Bryceson, D. F. (2008). Moral economy inversion: Ganyu Labour in rural Malawi. In I. N. Kimambo (Ed.), *Contemporary perspectives on African moral economy* (pp. 87-106).

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

Caruso, G. D., & Cardona Sosa, L. M. (2022). *Malawi Poverty Assessment : Poverty Persistence in Malawi - Climate Shocks, Low Agricultural Productivity, and Slow Structural Transformation (English).* Retrieved from Washington, D.C.: World Bank Group: http://documents.worldbank.org/curated/en/099920006302215250/P174948072f3880690afb70c20973fe214d

Chamberlain, H. R., Lazar, A. N., & Tatem, A. J. (2022). High-resolution estimates of social distancing feasibility, mapped for urban areas in sub-Saharan Africa. *Scientific Data, 9*(1), 711. doi:10.1038/s41597-022-01799-0

Christiaensen, L. (2013). Introduction: rural diversification, secondary towns and poverty reduction: do not miss the middle. *Agricultural Economics, 44*(4-5), 433-434. doi:10.1111/agec.12026

Cockx, L., Colen, L., & De Weerdt, J. (2018). From corn to popcorn? Urbanization and dietary change: Evidence from rural-urban migrants in Tanzania. *World Development, 110*, 140-159. doi:10.1016/j.worlddev.2018.04.018

Colella, F., Lalive, R., Sakalli, S. O., & Thoenig, M. (2019). Inference with arbitrary clustering. *IZA Discussion Paper No. 12584*. doi:10.2139/ssrn.3449578

de Brauw, A., Mueller, V., & Lee, H. L. (2014). The Role of Rural–Urban Migration in the Structural Transformation of Sub-Saharan Africa. *World Development, 63*, 33-42. doi:10.1016/j.worlddev.2013.10.013

de Janvry, A., Duquennois, C., & Sadoulet, E. (2022). Labor calendars and rural poverty: A case study for Malawi. *Food Policy, 109*, 102255. doi:https://doi.org/10.1016/j.foodpol.2022.102255

De Weerdt, J., Christiaensen, L., & Kanbur, R. (2021). When distance drives destination, towns can stimulate development. *Policy Research Working Paper; No. 9622*. Retrieved from https://openknowledge.worldbank.org/handle/10986/35444

Diao, X., Magalhaes, E., & Silver, J. (2019). Cities and rural transformation: A spatial analysis of rural livelihoods in Ghana. *World Development, 121*, 141-157. doi:10.1016/j.worlddev.2019.05.001

Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M., & Schiavina, M. (2021). Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics, 125*, 103312. doi:https://doi.org/10.1016/j.jue.2020.103312

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

Dorosh, P., & Thurlow, J. (2014). Can Cities or Towns Drive African Development? Economywide Analysis for Ethiopia and Uganda. *World Development, 63*, 113-123. doi:10.1016/j.worlddev.2013.10.014

Dorward, N., & Fox, S. (2022). Population pressure, political institutions, and protests: A multilevel analysis of protest events in African cities. *Political Geography, 99*, 102762. doi:https://doi.org/10.1016/j.polgeo.2022.102762

Elvidge, C. D., Zhizhin, M., Ghosh, T., Hsu, F.-C., & Taneja, J. (2021). Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019. *Remote Sensing, 13*(5), 922. Retrieved from https://www.mdpi.com/2072-4292/13/5/922

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), 147-160. doi:10.2174/1874923201003010147

Gibson, 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

Gibson, J., Olivia, S., & Boe-Gibson, G. L., Chao. (2021). Which night lights data should we use in economics, and where? *Journal of Development Economics, 149*, 102602. doi:10.1016/j.jdeveco.2020.102602

Gibson, J. K., & Boe-Gibson, G. (2020). Three facts about night lights data. *Working Papers in Economics*. Retrieved from https://EconPapers.repec.org/RePEc:wai:econwp:20/03

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

Gunning, J. W., Krishnan, P., & Mengistu, A. T. (2018). Fading Choice: Transport Costs and Variety in Consumer Goods. *CSAE Working Paper WPS/2018-052*.

Gutiérrez, J., Condeço-Melhorado, A., & Martín, J. C. (2010). Using accessibility indicators and GIS to assess spatial spillovers of transport infrastructure investment. *Journal of Transport Geography, 18*(1), 141-152. doi:10.1016/j.jtrangeo.2008.12.003

Head, K., & Mayer, T. (2014). Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of International Economics* (Vol. 4, pp. 131-195): Elsevier.

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

Krishnan, P., & Zhang, P. (2020). Restricting trade and reducing variety: Evidence from Ethiopia. *World Development, 126*, 104695. doi:10.1016/j.worlddev.2019.104695

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

Lucas, R. E. B. (2001). The effects of proximity and transportation on developing country population migrations. *Journal of Economic Geography, 1*(3), 323-339. doi:10.1093/jeg/1.3.323

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

Meredith, H. R., Giles, J. R., Perez-Saez, J., Mande, T., Rinaldo, A., Mutembo, S., . . . Wesolowski, A. (2021). Characterizing human mobility patterns in rural settings of sub-Saharan Africa. *eLife, 10*, e68441. doi:10.7554/eLife.68441

Nagler, P., & Naudé, W. (2013). Non-farm entrepreneurship in rural Africa: Patterns and determinants. *IZA Discussion Paper Series No. 8008*.

Nagler, P., & Naudé, W. (2014). Labor productivity in rural African enterprises: empirical evidence from the LSMS-ISA. *IZA Discussion Paper Series No. 8524*.

National Planning Commission. (2020). *Malawi 2063 - Transforming Our Nation*. Retrieved from https://malawi.un.org/sites/default/files/2021-01/MW2063-%20Malawi%20Vision%202063%20Document.pdf

National Statistical Office. (2012a). *Malawi Third Integrated Household Survey (IHS3) 2010-2011: Basic Information Document*. Retrieved from https://microdata.worldbank.org/index.php/catalog/3819/download/49100

National Statistical Office. (2012b). *Malawi Third Integrated Household Survey, 2010/11: Enumerator Manual for the Household Questionnaire*. Retrieved from https://microdata.worldbank.org/index.php/catalog/1003/download/40804

National Statistical Office. (2019). *2018 Malawi Population and Housing Census: Main Report*. Retrieved from https://malawi.unfpa.org/sites/default/files/resource-pdf/2018%20Malawi%20Population%20and%20Housing%20Census%20Main%20Report%20%281%29.pdf

OECD. (2020). *Cities in the World: A New Perspective on Urbanisation*. Paris: OECD Publishing.

Pezzulo, C., Hornby, G. M., Sorichetta, A., Gaughan, A. E., Linard, C., Bird, T. J., . . . Tatem, A. J. (2017). Sub-national mapping of population pyramids and dependency ratios in Africa and Asia. *Scientific Data, 4*(1), 170089. doi:10.1038/sdata.2017.89

Reardon, T., Boughton, D., Tschirley, D., Haggblade, S., Dolislager, M., Minten, B., & Hernandez, R. (2015). Urbanization, diet change, and transformation of the downstream and midstream of the agrifood system: effects on the poor in Africa and Asia. *Faith & Economics, 66*(1), 43-63. Retrieved from http://christianeconomists.org/2016/01/01/urbanization-diet-change-and-the-transformation-of-the-downstream-and-midstream-of-the-agrifood-system-reardon-boughton-tschirley-haggblade-dolislager-minten-hernandez/

Reed, F. J., Gaughan, A. E., Stevens, F. R., Yetman, G., Sorichetta, A., & Tatem, A. J. (2018). Gridded Population Maps Informed by Different Built Settlement Products. *Data, 3*(3), 33. Retrieved from https://www.mdpi.com/2306-5729/3/3/33

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:10.1016/j.jag.2012.02.009

Soto, J. D., Vargas, M., & 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. *Working Paper Series of the Latin American Center for Rural Development- Rimisp, 239*, 1-41.

Steel, G., & van Lindert, P. (2017). Mobility and connectivity: driving rural livelihood transformations in Africa. *IIED Briefing*. Retrieved from http://pubs.iied.org/10814IIED

Stevens, F. R., Gaughan, A. E., Nieves, J. J., King, A., Sorichetta, A., Linard, C., & Tatem, A. J. (2020). Comparisons of two global built area land cover datasets in methods to disaggregate human population in eleven countries from the global South. *International Journal of Digital Earth, 13*(1), 78-100. doi:10.1080/17538947.2019.1633424

Tacoli, C. (1998). Rural-urban interactions: a guide to the literature. *Environment and Urbanization, 10*(1), 147-166. doi:10.1177/095624789801000105

Tatem, A. J. (2017). WorldPop, open data for spatial demography. *Scientific Data, 4*(1), 170004. doi:10.1038/sdata.2017.4

Tschirley, D., Reardon, T., Dolislager, M., & Snyder, J. (2015). The Rise of a Middle Class in East and Southern Africa: Implications for Food System Transformation. *Journal of International Development, 27*(5), 628-646. doi:10.1002/jid.3107

U.N. (2018). *World Urbanization Prospects: The 2018 Revision.* Retrieved from https://esa.un.org/unpd/wup/Publications/Files/WUP2018-KeyFacts.pdf

Vandercasteelen, J., Beyene, S. T., Minten, B., & Swinnen, J. (2018). Cities and agricultural transformation in Africa: Evidence from Ethiopia. *World Development, 105*, 383-399. doi:10.1016/j.worlddev.2017.10.032

Vandercasteelen, J., Minten, B., & Tamru, S. (2021). Urban proximity, access to value chains, and dairy productivity in Ethiopia. *Agricultural Economics, 52*(4), 665-678. doi:https://doi.org/10.1111/agec.12641

Westlowski, A., Bengtsson, L., Buckee, C., Wetter, E., & Tatem, A. (2014). West Africa human mobility models. Retrieved from https://eprints.soton.ac.uk/408677/1/Flowminder_Mobility_Data_21.08.14.pdf

Wodon, Q., & Beegle, K. (2006). Labor shortages despite underemployment? Seasonality in time use in Malawi. In Q. Wodon & C. M. Blackden (Eds.), *Gender, Time Use, and Poverty in Sub-Saharan Africa.* Washington: The World Bank.

Yin, X., Li, P., Feng, Z., Yang, Y., You, Z., & Xiao, C. (2021). Which Gridded Population Data Product Is Better? Evidences from Mainland Southeast Asia (MSEA). *ISPRS International Journal of Geo-Information, 10*(10). doi:10.3390/ijgi10100681

ANNEXES

Annex A: Alternative urbanization proxies

In this section, we check whether our results are robust against alternative urbanization proxies. A potential concern is the possibility that population estimates only capture one part of urbanization or that they exhibit measurement errors. To investigate this, we compared the performance of the WorldPop data with that of the Visible Infrared Imaging Radiometer Suite (VIIRS) night time lights (NTL) data.

Instead of measuring changes in population in the agglomerations in our sample, we proxy urbanization by measuring changes in light emitted at night within the boundaries of the agglomerations. To measure night time lights (NTL) we use the Median Masked Radiance product of the new VIIRS satellite Version 2 (V2) data, which has been shown to be superior in measuring economic activity and urbanization compared to the NTL data measured with the Defense Meteorological Satellite Program Operational Linescan System (DMSP OLS) (J. Gibson, Olivia, & Boe-Gibson, 2021). The newly available V2 VIIRS Nighttime Lights (VNL) yearly time series that is based on the monthly VIIRS Version 1 (V1) products was recently produced and includes outlier removal (Elvidge, Zhizhin, Ghosh, Hsu, & Taneja, 2021). Despite its high potential, the availability of the data remain a disadvantage for time-series analysis: VIIRS data collection only started in April 2012 and stray light corrected yearly data is only available from 2014 onwards.

The light intensity per pixel is measured as radiances (in nano Watts per cm^2 per steradian) on a 15 arc-second grid, which is approximately a 500 meter square at the equator. As with the WorldPop data, the degree of 'urbanness' of each agglomeration is calculated by summing the light intensity scores of each 15 arc second grid raster that falls inside the Africapolis shapefile boundaries. 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 light intensity into account (Binswanger-Mkhize et al., 2016; Ghosh et al., 2010; Gibson et al., 2017; Henderson, Squires, Storeygard, & Weil, 2018; Small & Elvidge, 2013; Soto et al., 2018). Annex Table 1 provides VIIRS V2 statistics for Malawi for the period 2014 to 2019.

Annex Table 1. Statistics on sum-of-lights within Africapolis urban agglomeration boundaries, by year between 2014 and 2019

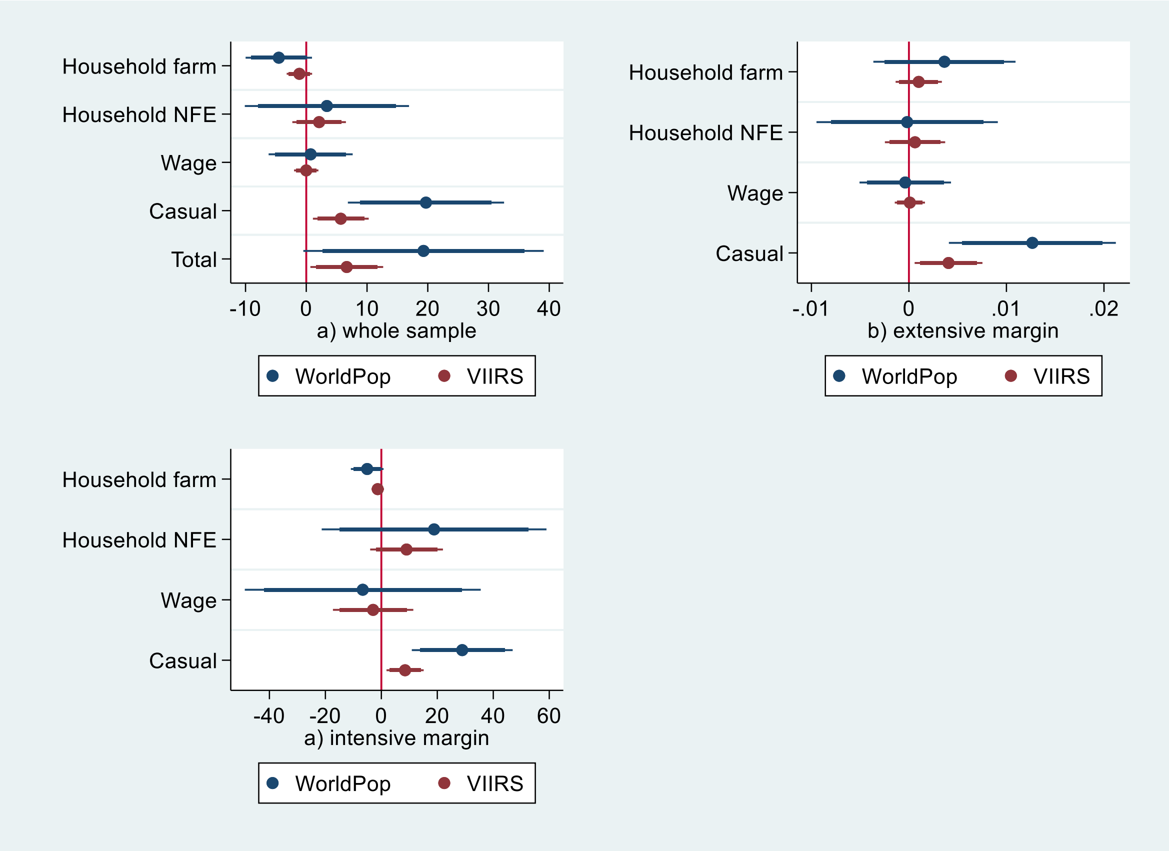
Sum of Lights	2014	2015	2016	2017	2018	2019
Average	211.8	233.7	200.8	198.5	207.0	204.1
Minimum	0	0	0	0	0	0
Maximum	6331.4	6724.6	5870.7	5954.8	5933.5	5823.3

Source: Authors' analysis.
Note: Based on analysis of VIIRS NTL V2 median masked radiance spatial data product.
Total agglomerations: 77.

Due to the fact that stray light corrected data is only available from 2014 onwards, we can only compare WorldPop and VIIRS performance for the shorter time period of 2016-2019. We therefore use the 2016-2017 IHS round for baseline labor data. The results can be found in

Annex Figure 1. We see that WorldPop and NTL data as a proxy for urbanization for the period 2016-2019 give similar results, although the coefficients for NTL are smaller. Urban growth over the shorter time period (2016-2019) has similar effects than those for the long time period (2010-2019) found in our main specification (Figure 6.1), which suggests that urbanization has a positive effect on total hours worked. Especially hours worked in casual labor (*ganyu*) are susceptible to urban growth. These results show that it is unlikely that our results are driven by measurement error and that population is likely to be a consistent proxy for urbanization, as they perform similarly to VIIRS NTL data. However it is important to note that, although the VIIRS NTL data have been hailed for their potential to study urbanization, concerns have been raised about their performance to predict changes over time (J. K. Gibson & Boe-Gibson, 2020). This could be especially the case for Malawi, which can be seen to experience fluctuating NTL emissions over time (see Annex Table 1). These fluctuations are due in part to seasonal disruptions in electricity supply caused by Malawi's heavy reliance on hydropower for electricity generation. Additionally, in 2016 Malawi experienced a drought (see Annex Figure 7 and Annex Figure 8 in Annex D). This could have impacted both labor supply figures as well as electricity supply.

Annex Figure 1. Effect of urbanization between 2016 and 2019 on hours worked, comparison of WorldPop versus VIIRS data as proxy for urbanization



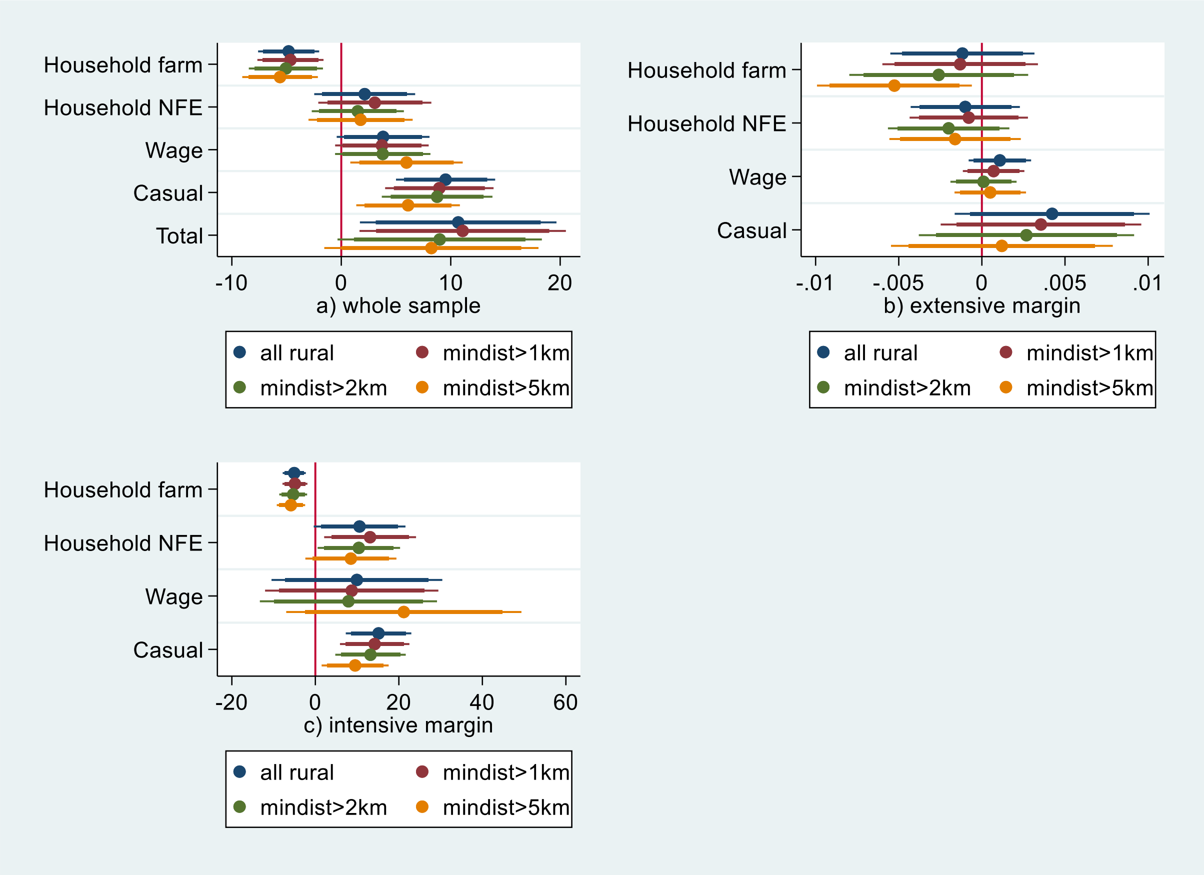
Source: Authors' analysis.
Note: This figure represents the point estimates of β_1 from Equation (3) and their corresponding 95% confidence intervals estimated with controls, comparing the performance of WorldPop and VIIRS as a proxy for urbanization.
In the panel at upper-left, each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various aggregations of employment categories.
The panel at upper-right estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin).
The panel at lower-left restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories.
Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a) and panel b): 2,351. Panel c): 2,261 for household farm labor, 698 for household non-farm enterprise labor, 276 for wage labor and 1,655 for casual labor.

As a second robustness check, we test if our results are robust to different definitions of what constitutes an urban area. The total sample of the long-term (2010-2019) IHPS panel consists of 102 EAs. 30 out of those 102 EAs are considered urban, while 72 (those in our sample) are rural. The categorization of these EAs is however based on administrative definitions of what constitutes an urban area. Officially, Malawi has four cities, six towns, 21 bomas (district capitals) and four designated urban areas (National Statistical Office, 2019). Africapolis, however, uses an algorithm based on continuously built-up area and a population threshold of 10 000 (based on disaggregated population statistics) to identify an urban agglomeration. As a consequence, 45 out of the 77 Africapolis agglomerations are not considered urban according to official administrative definitions. As we are using the Africapolis agglomerations as our definition of urban, it is informative to investigate how the rural/urban administrative status of the EAs in our sample behave according to Africapolis.

As we have the district and the sub-district Traditional Authority (TA) name (and code) for each EA, we can check the location of the EAs. Out of the 30 urban EAs, 12 are located in Lilongwe, six in Blantyre, five in Zomba, five in Mzuzu, one in Karonga, and one in Mwanza. These locations are all linked to an Africapolis agglomeration, so dropping those EAs from our sample is correct as we only want to investigate rural individuals. The other 72 EAs are not urban according to administrative definitions, but could be located in one of the (small) agglomerations not recognized as such. However, linking each EA to its exact location and thus checking whether it is located with certainty in an Africapolis agglomeration is not possible because of two reasons. First, we are not able to link them by administrative names as the Africapolis agglomerations do not follow administrative boundaries. Second, rural EAs are subjected to an offset of up to 5 km for sample household confidentiality purposes (National Statistical Office, 2012a). We could thus argue that an EA that is classified as rural according to the IHPS, falls within one of the non-official Africapolis agglomerations (or reasonably close to it to be directly influenced by it) when the minimum distance to the closest non-official agglomeration is smaller than a certain distance. For example, seven out of 72 ‘rural’ EAs have the closest non-official agglomeration within a radius of 2 km and ten out of 72 within a radius of 5 km.

Annex Figure 2 presents the comparison of the regression results for different subsets of rural individuals. When we impose a more strict definition of what constitutes rural the results are in line with our baseline specification, although they become smaller and less significant with stricter definitions. This finding could also be caused by the fact that the sample is reduced by 57 individuals (from 3 EAs) for a 1 km minimum distance, by 142 individuals (from 7 EAs) for a 2 km minimum distance, and by 209 individuals (from 10 EAs) for a 5 km minimum distance. We see, however, a stronger positive effect on wage labor for the more restricted sample of (more remote) rural individuals. By eliminating individuals that are located near a (smaller) agglomeration, we reduce the importance of smaller nearby agglomerations in the urban access variable. This finding is in line with what we found in section 6.2 that access to cities especially has a positive effect on hours worked off the household farm.

Annex Figure 2. Effect of urbanization between 2010 and 2019 on hours worked for different subsets of rural individuals based on different definition of 'rural'



Source: Authors' analysis.
Note: This figure represents the point estimates of β_1 from Equation (3) and their corresponding 95% confidence intervals estimated with controls, comparing the results for different subsets of rural individuals. The results for 'all rural' present the results for all individuals considered rural according to official administrative definitions. The other results apply a more strict definition of what constitutes as rural: only the individuals that are considered rural according to official definitions and are at a minimum distance of at least 1/2/5 km from any other (non-official) Africapolis agglomeration are included in the sample.
In the panel at upper-left, each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various aggregations of employment categories.
The panel at upper-right estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin).
The panel at lower-left restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories.
Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a) and panel b): 1,390. Seventeen individuals dropped from sample due to missing education variables. Panel c): 1,356 for household farm labor, 400 for household non-farm enterprise labor, 181 for wage labor and 862 for casual labor.

Annex B: Direction of causality

In this annex, we investigate the causality of the effects of our empirical analysis by checking whether our main results are robust to controlling for local population growth and to lagging the urban access variable by two years.

One potential concern with our main results is that they are partly driven by growth in local population density in rural areas. To investigate this, we calculate the population growth for the Traditional Authority (TA) in which the household is located. To circumvent the offset imposed on the survey location coordinates, we match the household TA with a TA boundaries shapefile,[20] not by using the geographic location of the household, but by linking the official TA code that is found in both datasets. The population is calculated each year for each TA by calculating the total WorldPop population within the TA boundaries. Calculating the difference between the end- and baseline logged value of the TA population provides us with a measure of local population growth.

We first note that the correlation coefficient between urban access and local population is 0.29 at the baseline, expressed in levels. Expressed in differences between 2019 and 2010, the correlation between local population growth and growth in surrounding urban areas is 0.48. The descriptive statistics of local population growth can be found in Annex Table 2.

Annex Table 2. Population growth at Traditional Authority (TA) level, 2010 to 2019

	Observations	Mean	Standard deviation	Minimum	Maximum
TA population, 2010	1,407	92,515	55,,853	11,340	244,652
TA population, 2019	1,407	117,110	70,505	15,065	313,239
Change in (∆) TA population, logged	1,407	0.24113	0.06966	0.08620	0.46601
Change in (∆) TA population, logged, lagged 1 year	1,407	0.24355	0.07073	0.08473	0.46565

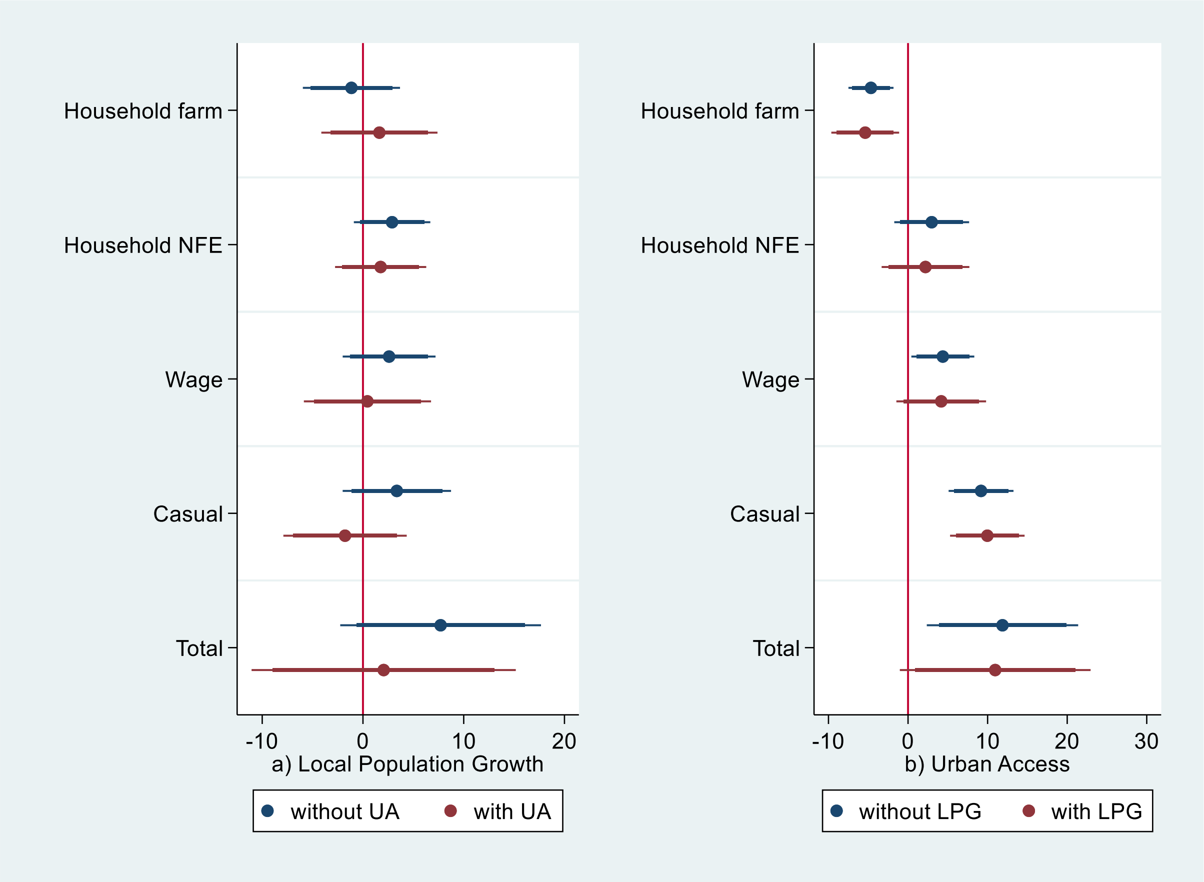
Source: Authors' analysis.

We re-estimate Equation (3), replacing the urban access variable with local population growth. The results, plotted in panel a), the left hand side panel, of Annex Figure 3 show that local population growth has no significant effect on hours worked, both with and without controlling for urban access. This provides evidence that it is unlikely that the findings from Figure 6.1 are driven by local population growth.

We also re-estimate the main regression plotted in panel a) of Figure 6.1, controlling for growth in local population density, and see that the effect on labor on the household farm, casual labor (*ganyu*) and total labor are robust to the inclusion of local population growth (panel b - the right hand side panel) in Annex Figure 3. The effect on wage labor is more sensitive to the inclusion of control variables (as seen as well in Figure 6.1), which partly reflects multicollinearity. On the intensive and extensive margin no robust significant effects of local growth could be found either (Annex Figure 4). It is important to note, though, that this robustness check leaves other potential omitted variables unaddressed.

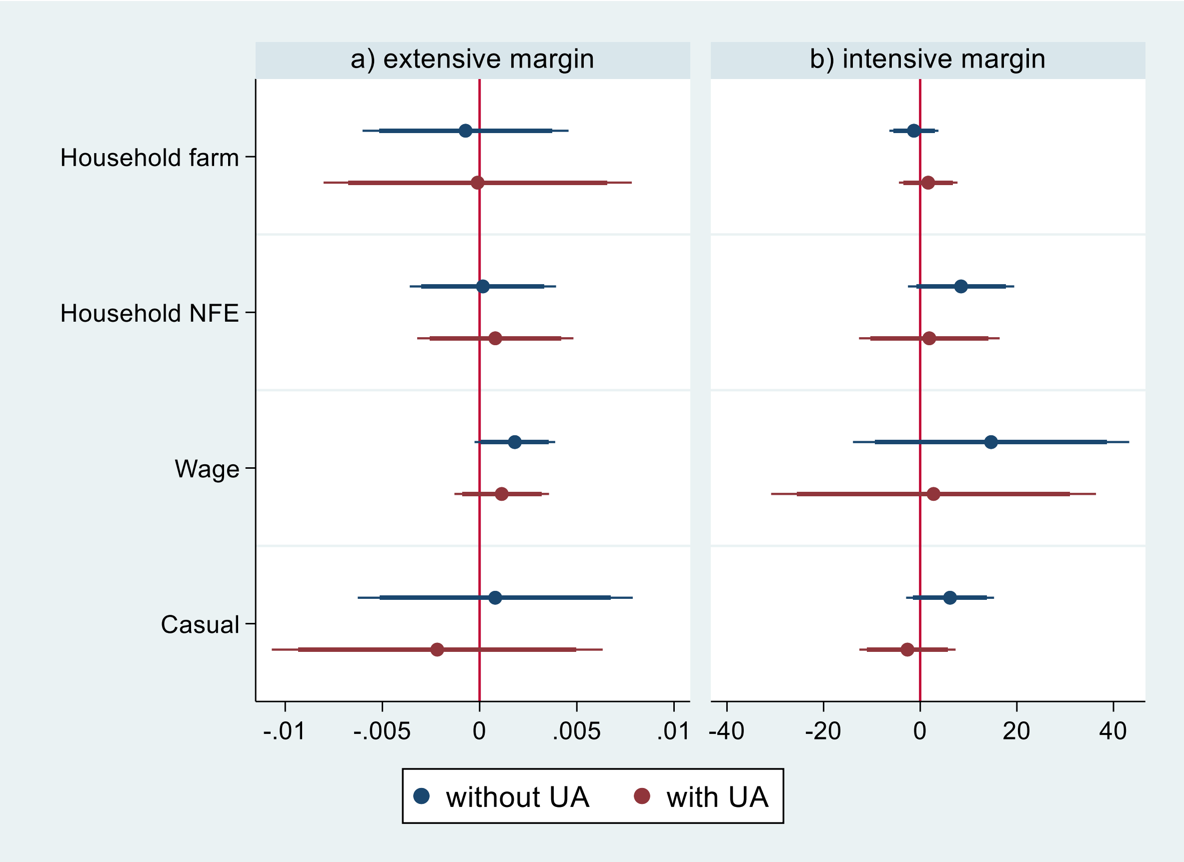
[20] We use shapefiles from the OCHA database: https://data.humdata.org/dataset/cod-ab-mwi

Annex Figure 3. Effect of local (rural) population growth and surrounding urbanization between 2010 and 2019 on hours worked



Source: Authors' analysis.
Note: This figure represents the point estimates of β_1 from Equation (3) and their corresponding 95% confidence intervals, estimated without controls.
In panel a) however, urban access is replaced with a measure of local population growth. The effect of local population growth is estimated both without and with the inclusion of the urban access variable as control variable.
Panel b) compares the effect of urban access without the inclusion of local population growth (as in our main specification in Figure 6.1) and with the inclusion of local population growth as a control variable.
Each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various aggregations of employment categories. Due to the level-log specification, the coefficients can be interpreted as the effect of a one percent increase in local population growth or urban access on hours worked respectively.
Observations: 1,407.

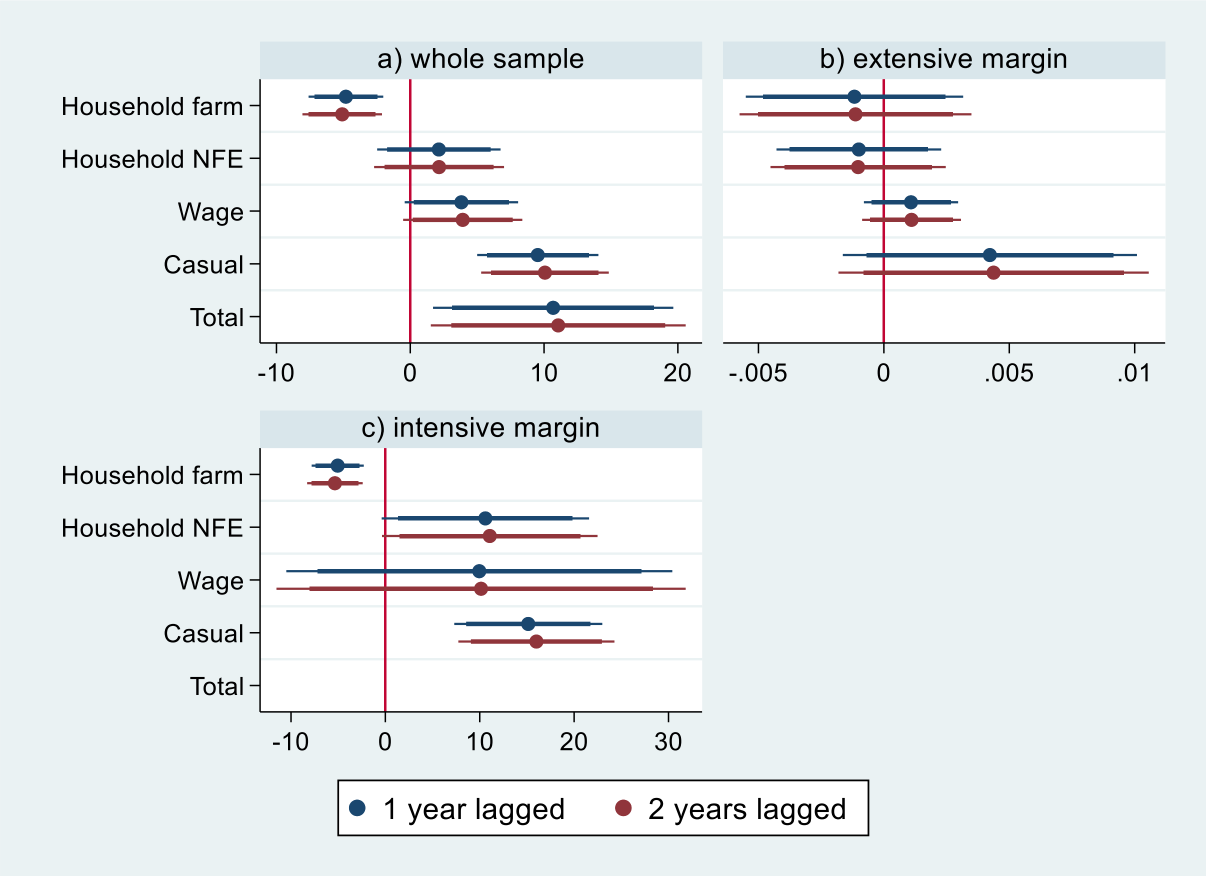
Annex Figure 4. Effect of local population growth between 2010 and 2019 on hours worked, extensive and intensive margins



Source: Authors' analysis.
Note: This figure represents the point estimates of β_1 from Equation (3) and their corresponding 95% confidence intervals, replacing urban access with a measure of local population growth. Both panels compare the effect of local population growth without and with the inclusion of urban access as a control variable. Panel a) estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing local population (the extensive margin). Panel b) restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories. Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a): 1,407. Panel b): 1,371 for household farm labor, 401 for household non-farm enterprise labor, 184 for wage labor and 871 for casual labor.

A second potential concern stems from the direction of causality. As explained in section 6.4, the reverse of our claim has been true in other contexts, which is that, instead of urban growth impacting rural labor patterns, rural development is causing nearby urban growth. Keeping in mind the fact that the period of analysis was one of stagnating agricultural productivity and the findings on local population growth above, this hypothesis is unlikely. A third way of assessing this concern is using lagged urbanization variables. In all our specifications, we use a one-year lagged urban growth variable. Additionally, in Annex Figure 5, we compare the results of the main specification in Figure 6.1 for one-year lagged and two-year lagged urban growth. We find that the results do not change qualitatively across the different urban growth specifications. The findings in this section thus show that it is unlikely that our results suffer from reverse causality.

Annex Figure 5. Effect of urbanization between 2010 and 2019 on hours worked using one-year lagged and two-year lagged urbanization variables



Source: Authors' analysis.
Note: This figure represents the point estimates of β_1 from Equation (3) and their corresponding 95% confidence intervals estimated with controls, comparing the results for contemporaneous urban access, one year lagged urban access and two year lagged urban access.
In panel a), each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various aggregations of employment categories.
Panel b) estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin).
Panel c) restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories.
Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a) and panel b): 1,390. Seventeen individuals dropped from sample due to missing education variables. Panel c): 1,356 for household farm labor, 400 for household non-farm enterprise labor, 181 for wage labor and 862 for casual labor.

Annex C: Comparing different population measures

Error! Reference source not found. shows a comparison between the Africapolis population for 2015, the population count based on WorldPop for 2015 and 2018, and the population count according to the census of 2018 for 29 agglomerations with the administrative title of ‘city’, ‘town’, ‘boma’ or ‘urban’ in the 2018 Census and for which the spatial boundaries are comparable. Note that Africapolis agglomerations do not follow any official administrative boundaries, so comparison with census statistics is not perfect from a spatial perspective. For Africapolis agglomerations that are not officially recognized, only TA wide population statistics are available in the 2018 Census. For Mpana/Likoswe (the urban outskirt of Blantyre), the populations of the two TAs with the respective names are summed. For an example on the comparison of Africapolis and administrative boundaries, see **Error! Reference source not found.** for the case of Mzuzu.

From these comparisons, we can conclude that the population statistics of Africapolis likely overstate the actual population, while those of WorldPop (slightly) underestimate the population. This is most likely to be the result of the different estimation techniques used for the different data products, something also noted by Dijkstra et al. (2021) and Dorward and Fox (2022). The WorldPop data product that we use applies an unconstrained approach that divides the population estimates over all 3 arc-second land-grid cells, in contrast with the constrained product that divides population estimates only within areas containing built settlements. The advantage of the unconstrained method is that it does not depend on measurement of built-up area for division of population counts. This is especially useful in contexts where built up area is unlikely to be accurately measured (Dijkstra et al., 2021). The consequence however is that all grid cells have non-zero population counts. The trade-off of false positives (non-zero population counts where there are actually no people living) against false negatives (built settlements that remain undetected) likely leads to an underestimation of urban populations. The Africapolis dataset on the other hand can be thought of as being a constrained dataset of agglomerations, since the condition for demarcating agglomeration boundaries is having a population that exceeds 10,000 people and a built environment that contains no unbuilt spaces greater than 200 meters (Africapolis, 2022). These built environment demarcations are then used as boundaries to which to designate the census statistics of urban population. Because spatial agglomeration boundaries are different and often more coarse than administrative boundaries (especially for smaller or ‘unrecognized’ agglomerations), the population in these agglomerations is higher than the underlying administrative population counts (see also **Error! Reference source not found.**).

One way to check for dependence of the results on the choice of population measure would be to do a robustness check based on the constrained WorldPop data product. Unfortunately, this is not possible as the product is available only for 2020. However, the believe that our results are not dependent on the underlying WorldPop data is strengthened by the following two arguments. First of all, Reed et al. (2018) investigate the effectiveness of constrained and unconstrained measures for six countries among which Malawi. It shows that for Malawi the unconstrained method has the lowest minimum error for different measures of error, but that constrained and unconstrained methods generate comparable accurate results. Secondly, in our analysis, we look at differences. The results would depend on the product used insofar that relative changes based on different population statistics for the given Africapolis agglomerations differ in a non-random way. Comparing Africapolis and WorldPop population estimates for 2010 and 2015 reveal that the contemporary correlation is 99.25 for 2010 and

99.30 for 2015, and the correlation between the 2018 census data and the 2018 WorldPop data for the 45 agglomerations with comparable boundaries in Annex Table 3**Error! Reference source not found.** is 0.9951. The correlation in changes over 2010-2015 is 0.97. This makes sense as both WorldPop and Africapolis interpolate census data to get yearly/5-yearly/10-yearly population statistics.

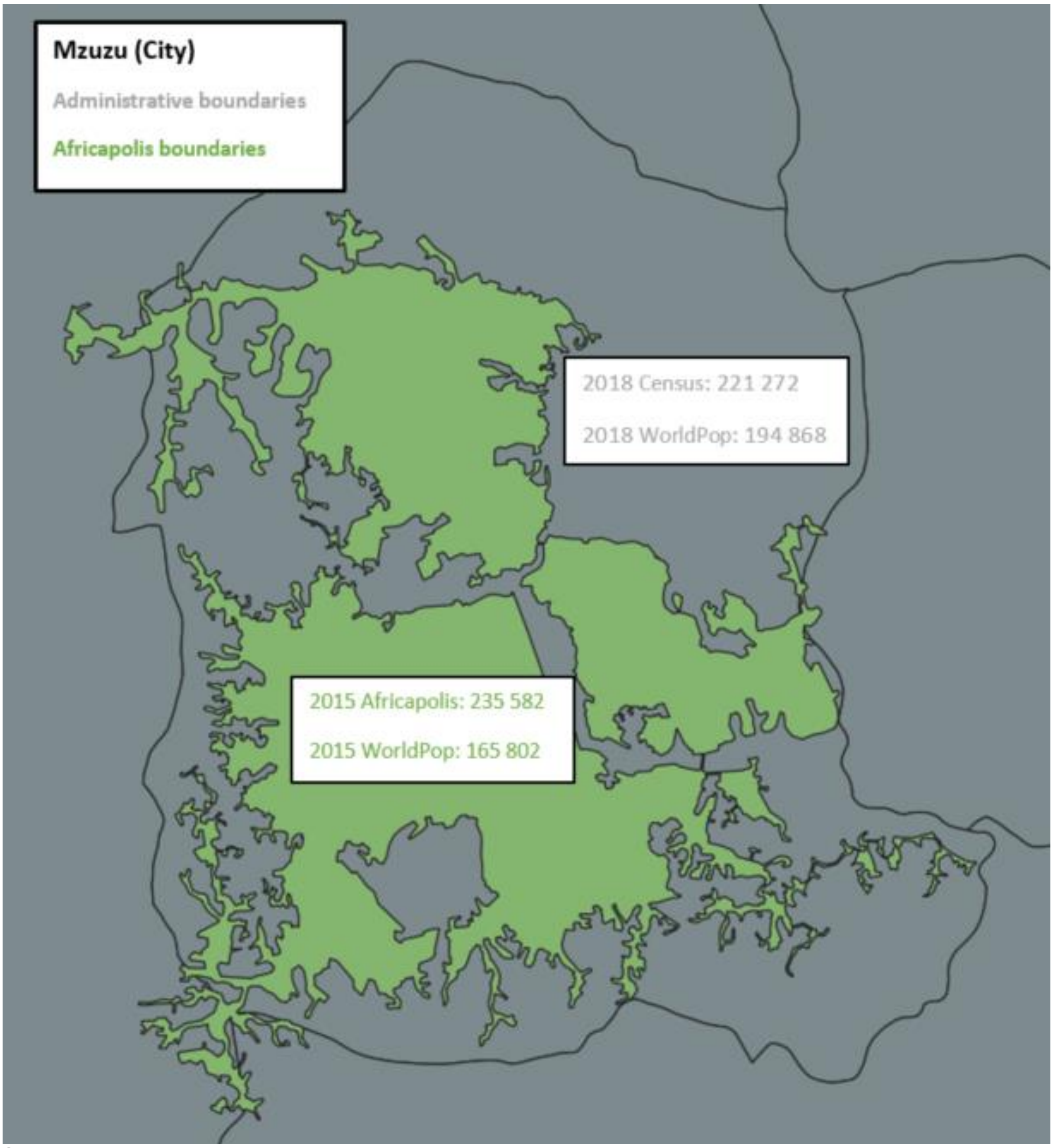
Annex Table 3. Comparison of WorldPop, Africapolis and Census population for the 77 agglomerations in Malawi as defined by Africapolis

Name	WorldPop 2015	Africapolis 2015	World Pop 2018	Census 2018	Census 2018 adm. unit
Balaka	35,862	61,833	39,789	36,308	town
Bangula UA	16,932	30,830	18,546		
Bereu UA	2,729	13,053	2,972		
Blantyre	748,682	1,057,790	805,150	800,264	city
Bunda UA	2,461	10,718	2,715		
Bvumbwe UA	16,473	51,899	17,535		
Bweteka	2,003	10,412	2,205		
Chikwawa	11,363	35,293	12,538		
Chintheche UA	5,940	31,540	6,684		
Chipoka UA	2,718	10,486	2,851		
Chitipa	26,941	30,127	31,460		
Dedza	18,086	37,826	19,208	30,928	Boma
Dowa	3,234	21,023	3,183	7,135	Boma
Dwangwa UA	8,903	17,820	9,228		
Dzaleka UA	4,813	12,571	4,850		
Dzoole UA	6,838	10,691	7,434		
Ekwendeni UA	10,252	20,181	11,341		
Jombo UA	3,452	10,343	3,978		
Kadozo UA	6,133	14,002	6,496		
Kamange UA	1,074	13,419	1,194		
Karonga	42,164	66,954	46,038	61,609	town
Kasankha UA	1,800	13,084	1,919		
Kasungu	36,387	85,824	39,767	58,653	Boma
Lilongwe	883,352	1,124,965	1,054,810	989,318	city
Lirangwe UA	4,334	10,880	4,607		
Liwonde	19,441	66,178	21,283	23,374	town
Luchenza	13,245	55,320	13,442	12,600	town
Lukwa UA	5,378	15,789	5,842		
Lumbadzi	18,548	40,811	20,921		
MWI3081388	1,499	11,714	1,597		
Madisi	5,763	18,755	6,036		
Maganaa UA	3,808	22,053	4,060		
Makanjila UA	3,194	11,547	3,839		
Makoko UA	5,099	12,114	5,258		
Malindi/Mizingo UA	8,712	20,126	9,531		
Malomo UA	2,214	13,243	2,342		
Mangochi	42,376	58,824	46,471	53,498	town
Mchinji	21,217	38,667	23,857	28,011	boma
Miseu Folo UA	3,253	11,488	3,397		

Mitundu UA	12,716	23,271	13,742		
Mkanda UA	3,390	15,608	3,749		
Mpama/Likoswe UA	20,972	104,215	22,109	50,846 68,321	TA mpama TA Likoswe
Mponda UA	985	10,319	1,081		
Mponela	17,747	33,507	19,829	24,543	Urban
Mposa UA	2,184	11,697	2,359		
Mulanje	7,845	37,172	8,099	14,782	Boma
Mwanza	11,452	27,086	13,288	18,039	Boma
Mzimba	19,578	42,784	21,810	26,096	Boma
Mzuzu	145,651	235,582	165,802	221,272	City
Nayuchi UA	5,236	22,722	6,033		
Nchalo UA	11,975	15,918	12,499		
Ndamera UA	6,063	20,491	6,238		
Neno	3,168	22,952	3,390	2,283	Boma
Ngabu	8,653	19,125	9,130	7,032	Urban
Ngabu UA	2,822	14,728	2,990		
Nkanda/Chikumbu UA	26,247	90,190	27,327		
Nkhata bay	10,171	14,172	10,759	14,274	Boma
Nkhoma UA	2,458	10,273	2,401		
Nkhotakota	23,322	42,786	24,686	28,350	Boma
Nkopola UA	23,109	51,050	24,398		
Nkotakata UA	13,417	16,350	15,102		
Nkwazi UA	1,009	11,985	1,094		
Nsangwe UA	1,339	11,193	1,511		
Nsanje	18,581	30,679	19,429	26,844	Boma
Ntcheu	16,573	43,998	18,308	21,241	Boma
Ntchisi	9,353	29,906	10,070	9,357	Boma
Phalombe	14,440	51,809	16,020	6,242	Boma
Phodgoma UA	11,803	47,607	11,203		
Rhumpi	13,820	25,458	14,444	22,358	Boma
Salima	38,928	73,778	42,458	36,789	Town
Senga UA	8,718	22,259	9,609		
South of Mabuka UA	6,478	39,067	6,706		
Thornwood UA	15,749	18,391	17,215		
Thyolo	8,455	22,161	8,784	7,843	Boma
Timbiri	1,155	10,522	1,290		
Uliwa UA	9,881	31,111	10,967		
Zomba	115,607	283,884	124,609	105,013	City
Total	2,727,720	4,552,115	2,930,301		
Average	35,425	59,118	38,056		
Median	8,903	22,259	9,609		

Source: Authors' analysis.

Annex Figure 6. Comparison of the administrative boundaries of Mzuzu City district and Mzuzu agglomeration as defined by Africapolis



Source: Authors' composition.

Annex D: Additional figures and tables of statistics and analytical results

Annex Table 4. Annual labor allocation of rural workers at baseline and endline for different age groups

	Baseline (2010)			Endline (2019)		
	Hours worked, average	Hours worked, conditional on working, avg.	Individuals that worked at least one hour, share	Hours worked, average	Hours worked, conditional on working, avg.	Individuals that worked at least one hour, share
Individuals aged less than 25 years (n=376)						
Total Labor	415.6	459.6	0.90	545.1	605.7	0.90
Household farm	186.2	215.5	0.86	76.7***	85.2***	0.90
Household non-farm enterprise	82.6	887.7	0.09	50.8	508.0	0.10
Agricultural	7.4	466.6	0.02	50.8	508.0	0.10***
Non-agricultural	78.4	951.5	0.08	0.0	.	0.00
Wage labor	28.1	660.5	0.04	16.8	336.0	0.05
Agricultural	12.2	767.3	0.02	0.0	.	0.00
Non-agricultural	15.9	596.4	0.03	16.8	336.0	0.05
Casual labor	118.6	345.7	0.34	400.8	801.6	0.50
Individuals aged 25 to 45 years (n=774)						
Total Labor	599.9	616.7	0.97	785.6***	799.6***	0.98
Household farm	256.8	276.8	0.93	126.4***	141.1***	0.90**
Household non-farm enterprise	95.4	716.6	0.13	225.3***	930.3*	0.24***
Agricultural	6.3	377.9	0.02	17.7**	460.9	0.04***
Non-agricultural	89.0	757.1	0.12	207.5***	1,001.6**	0.21***
Wage labor	126.7	1,420.9	0.09	101.8	1,197.6	0.08
Agricultural	27.7	1,193.3	0.02	26.5	1,337.4	0.02
Non-agricultural	96.8	1,498.6	0.06	75.3	1,155.1	0.07
Casual labor	121.1	370.6	0.33	332.1***	587.1***	0.57***
Individuals aged over 45 years (n=257)						
Total Labor	594.1	598.8	0.99	716.8**	743.6**	0.96**
Household farm	308.4	319.5	0.96	170.9***	192.4***	0.89***
Household non-farm enterprise	61.6	494.7	0.12	207.9***	1,076.2***	0.19**
Agricultural	1.2	156.0	0.01	28.4**	651.6	0.04***
Non-agricultural	60.4	517.3	0.12	179.5***	1,170.2***	0.15
Wage labor	128.8	1,226.4	0.11	97.7	,1097.5	0.09
Agricultural	37.5	1,204.0	0.03	26.2	1,384.0	0.02
Non-agricultural	80.9	1,155.1	0.07	71.5	1,020.1	0.07
Casual labor	95.3	429.7	0.22	240.3***	592.9*	0.41***

Source: Authors' analysis.
Note: t-test for difference from 2010 at significance level: *p<0.1; **p<0.05; ***p<0.01.

Annex Table 5. Differential effects of urban access, whole sample

	Hours worked on household farm	Hours worked in household non-farm enterprise	Hours worked in wage labour	Hours worked in casual labour	Total hours worked
Urban access by male interaction					
Urban access	-505.4***	450.1	53.11	1,086***	1,083**
	(143.0)	(336.2)	(86.70)	(249.9)	(429.2)
Urban access * male	57.5	-553.4	771.0	-310.5	-35.5
	(150.6)	(525.4)	(494.7)	(344.4)	(843.2)
Urban access by went to school interaction					
Urban access	-714.5***	145.1	391.3	687.9	509.8
	(239.1)	(574.3)	(238.6)	(520.4)	(641.7)
Urban access * went to school	296.7	87.3	-11.6	336.9	709.3
	(223.4)	(573.0)	(294.9)	(612.6)	(745.1)
Urban access by higher education than primary school interaction					
Urban access	-543.0***	229.9	292.0	942.1***	921.0**
	(145.5)	(263.2)	(246.1)	(235.1)	(431.8)
Urban access * higher education than primary school	332.5*	-85.8	482.8	59.1	788.7
	(199.0)	(679.4)	(992.2)	(306.7)	(1,403)
Urban access by can read or write interaction					
Urban access	-388.0**	95.2	521.9*	907.9***	1,137**
	(170.6)	(317.6)	(302.8)	(293.3)	(552.0)
Urban access * can read or write	-161.5	206.2	-243.0	78.7	-119.5
	(151.0)	(438.7)	(407.3)	(384.3)	(849.4)
Urban access by age interaction					
Urban access	-34.6	-251.8	290.2	1,688**	1,692
	(308.1)	(838.7)	(588.9)	(682.7)	(1,284)
Urban access * age	-13.5	14.0	2.8	-22.2	-18.8
	(9.3)	(25.5)	(18.2)	(17.7)	(39.8)

Source: Authors' analysis.
Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Observations: 1,390. Seventeen individuals dropped from sample due to missing education variables. All regression include baseline control variables: age, male dummy, went to school dummy, higher education than primary school dummy, can read or write dummy, and eight household size controls based on four age categories (0-5, 5-15, 15-65, >65 years of age) for both sexes.

Annex Table 6. Differential effects of urban access, extensive margin

	Hours worked on household farm	Hours worked in household non-farm enterprise	Hours worked in wage labour	Hours worked in casual labour
Urban access by male interaction				
Urban access	-0.078	-0.131	0.091	0.684**
	(0.235)	(0.216)	(0.091)	(0.320)
Urban access * male	-0.091	0.072	0.040	-0.611*
	(0.240)	(0.295)	(0.239)	(0.315)
Urban access by went to school interaction				
Urban access	-0.488	0.060	0.091	0.242
	(0.366)	(0.396)	(0.105)	(0.632)
Urban access * went to school	0.471	-0.203	0.021	0.229
	(0.310)	(0.412)	(0.210)	(0.755)
Urban access by higher education than primary school interaction				
Urban access	-0.160	-0.110	0.174*	0.333
	(0.251)	(0.181)	(0.096)	(0.323)
Urban access * higher education than primary school	0.232	0.053	-0.351	0.480
	(0.392)	(0.367)	(0.328)	(0.434)
Urban access by can read or write interaction				
Urban access	-0.337	-0.156	0.258**	0.574
	(0.289)	(0.224)	(0.115)	(0.386)
Urban access * can read or write	0.381	0.097	-0.261	-0.264
	(0.253)	(0.256)	(0.235)	(0.466)
Urban access by age interaction				
Urban access	0.629	-0.583	0.704*	1.569*
	(0.398)	(0.480)	(0.422)	(0.919)
Urban access * age	-0.022***	0.015	-0.018	-0.035
	(0.008)	(0.014)	(0.013)	(0.024)

Source: Authors' analysis.
Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Observations: 1,390. Seventeen individuals dropped from sample due to missing education variables. All regression include baseline control variables: age, male dummy, went to school dummy, higher education than primary school dummy, can read or write dummy, and eight household size controls based on four age categories (0-5, 5-15, 15-65, >65 years of age) for both sexes.

Annex Table 7. Differential effects of urban access, intensive margin

	Hours worked on household farm	Hours worked in household non-farm enterprise	Hours worked in wage labour	Hours worked in casual labour
Urban access by male interaction				
Urban access	-524.6***	2,253**	-94.53	1,675***
	(140.6)	(884.3)	(1,459)	(397.2)
Urban access * male	46.7	-2,480*	1,350	-380.9
	(157.9)	(1,499)	(2,099)	(595.7)
Urban access by went to school interaction				
Urban access	-719.3***	1,352	6,282**	931.3
	(239.9)	(2,219)	(2,908)	(722.3)
Urban access * went to school	273.9	-347.1	-5,597*	770.2
	(222.7)	(2,452)	(2,879)	(871.3)
Urban access by higher education than primary school interaction				
Urban access	-560.5***	1,300*	473.7	1,397***
	(143.2)	(696.6)	(1,671)	(387.1)
Urban access * higher education than primary school	318.4	-794.7	1,251	863.3
	(224.1)	(1,427)	(2,985)	(566.6)
Urban access by can read or write interaction				
Urban access	-394.2**	1,123	3,426	1,320***
	(172.9)	(1,045)	(2,335)	(457.9)
Urban access * can read or write	-196.2	-91.2	-3,193	370.3
	(152.3)	(1,423)	(2,582)	(644.9)
Urban access by age interaction				
Urban access	-125.1	769.9	1,039	1,753*
	(326.2)	(3,362)	(2,456)	(976.8)
Urban access * age	-11.4	8.9	-1.3	-7.5
	(9.9)	(100.5)	(59.4)	(27.6)

Source: Authors' analysis.
Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Observations: 1,390. Seventeen individuals dropped from sample due to missing education variables.
All regression include baseline control variables: age, male dummy, went to school dummy, higher education than primary school dummy, can read or write dummy, and eight household size controls based on four age categories (0-5, 5-15, 15-65, >65 years of age) for both sexes.

Annex Table 8. 2019 population of urban agglomerations (UA) in the sample, based on WorldPop

Population agglomeration name (Africapolis)	2019 Population estimate (WorldPop)	Population agglomeration name (Africapolis)	2019 Population estimate (WorldPop)
Cities (N=4)		Makoko UA	5,280
Blantyre	825,358	Malindi/Mizingo UA	9,659
Lilongwe	1,124,959	Malomo UA	2,481
Mzuzu	173,324	Mangochi	48,045
Zomba	128,290	Mchinji	24,837
		Miseu Folo UA	3,483
Prospective secondary towns (N=7)		Mitundu UA	14,148
Bangula UA	18,831	Mkanda UA	3,863
Chipoka UA	2,942	Mpama/Likoswe UA	22,495
Karonga	47,419	Mponda UA	1,024
Kasungu	41,219	Mponela	20,519
Liwonde	21,868	Mposa UA	2,388
Luchenza	13,581	Mulanje	8,145
Nkhata Bay	10,958	Mwanza	14,034
		Mzimba	22,544
Other agglomerations (N=66)		Nayuchi UA	6,155
Balaka	41,182	Nchalo UA	12,598
Bereu UA	3,030	Ndamera UA	6,338
Bunda UA	2,752	Neno	3,541
Bvumbwe UA	17,889	Ngabu	9,358
Bweteka	2,185	Ngabu UA	3,016
Chikwawa	12,477	Nkanda/Chikumbu UA	27,760
Chintheche UA	6,731	Nkhoma UA	2,444
Chitipa	33,127	Nkhotakota	25,130
Dedza	19,622	Nkopola UA	25,023
Dowa	3,188	Nkotakata UA	15,584
Dwangwa UA	9,370	Nkwazi UA	1,085
Dzaleka UA	5,205	Nsangwe UA	1,543
Dzoole UA	7,606	Nsanje	19,430
Ekwendeni UA	11,681	Ntcheu	18,896
Jombo UA	4,061	Ntchisi	10,392
Kadozo UA	6,470	Phalombe	16,584
Kamange UA	1,200	Phodgoma UA	10,971
Kasankha UA	2,054	Rumphi	14,769
Lirangwe UA	4,706	Salima	43,547
Lukwa UA	6,020	Senga UA	9,839
Lumbadzi	21,921	South of Mabuka UA	6,816
MWI3081388	1,620	Thornwood UA	17,691
Madisi	6,214	Thyolo	8,924
Maganaa UA	4,091	Timbiri	1,357
Makanjila UA	4,068	Uliwa UA	11,153

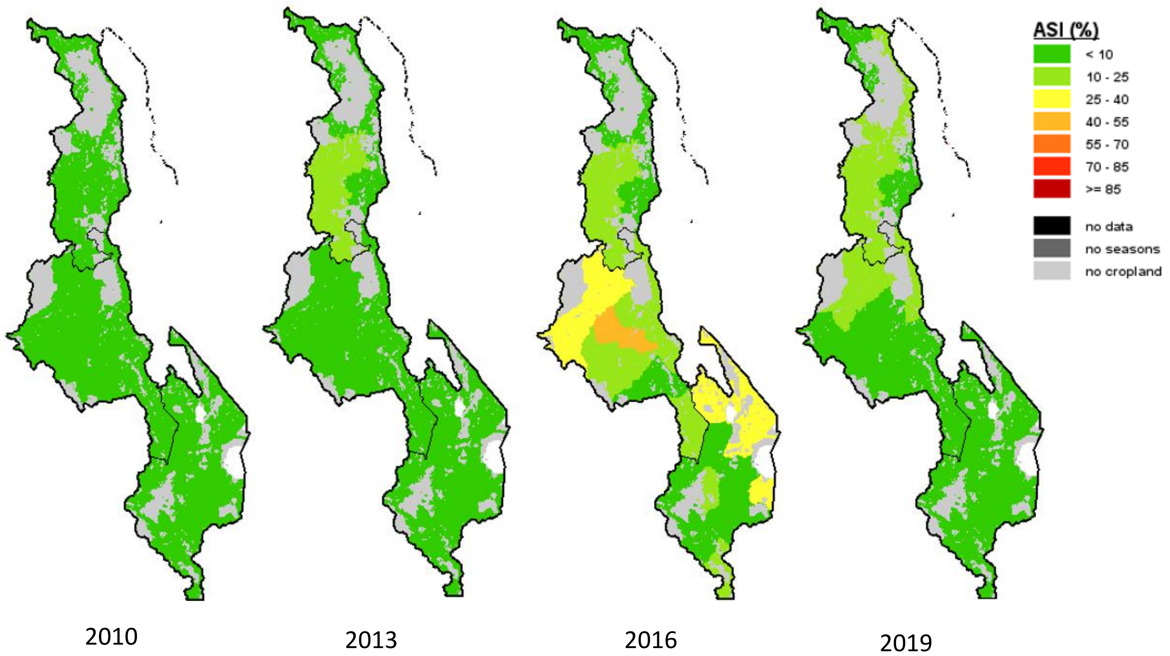
Source: Authors' analysis.

Annex Table 9. Spatial distribution of Enumeration Areas in sample

	Observations	Mean	Median	Standard deviation	Min-imum	Max-imum
Agglomerations within 20 km, no.	72	1.06	1	1.11	0	5
Agglomerations within 50 km, no.	72	6.92	7	3.91	0	18
Agglomerations within 100 km, no.	72	18.81	21	6.83	1	27
Distance to closest EA, km	72	19.6	15.3	13.5	44.8	54.0
Enumeration Areas with at least one other Enumeration Area within:						
20 km, no.	46 (64%)					
50 km, no.	66 (92%)					
100km, no.	72 (100%)					

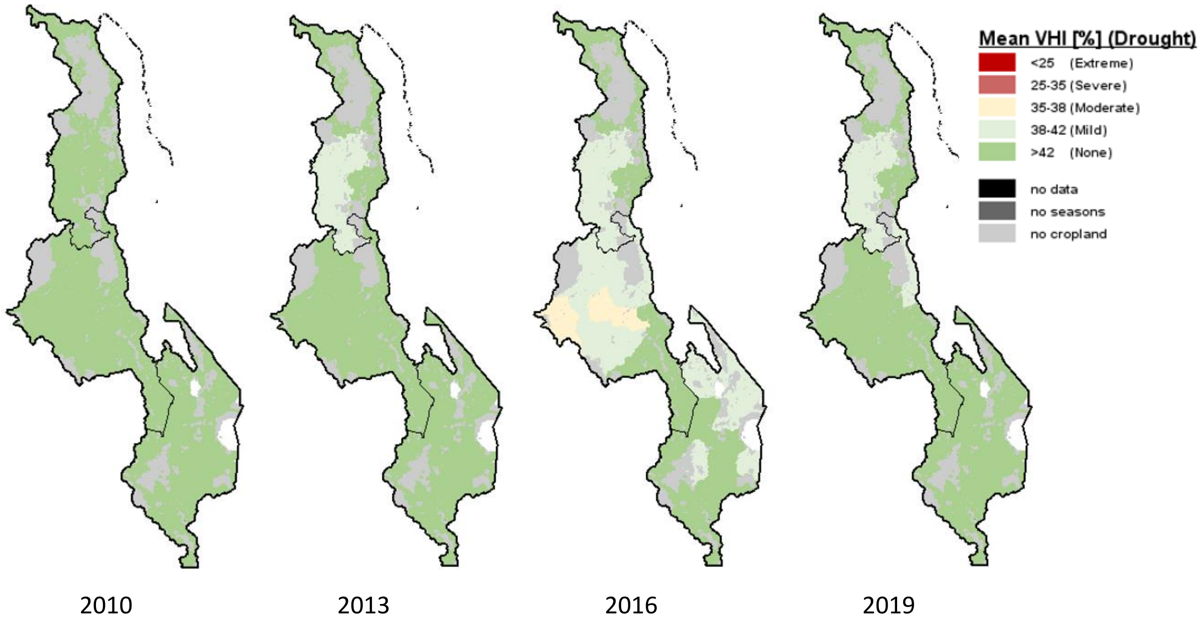
Source: Authors' analysis.

Annex Figure 7. Agricultural Stress Index (ASI), % of cropland area affected by severe drought per GAUL 2 region for complete season 1 of 2010, 2013, 2016 and 2019



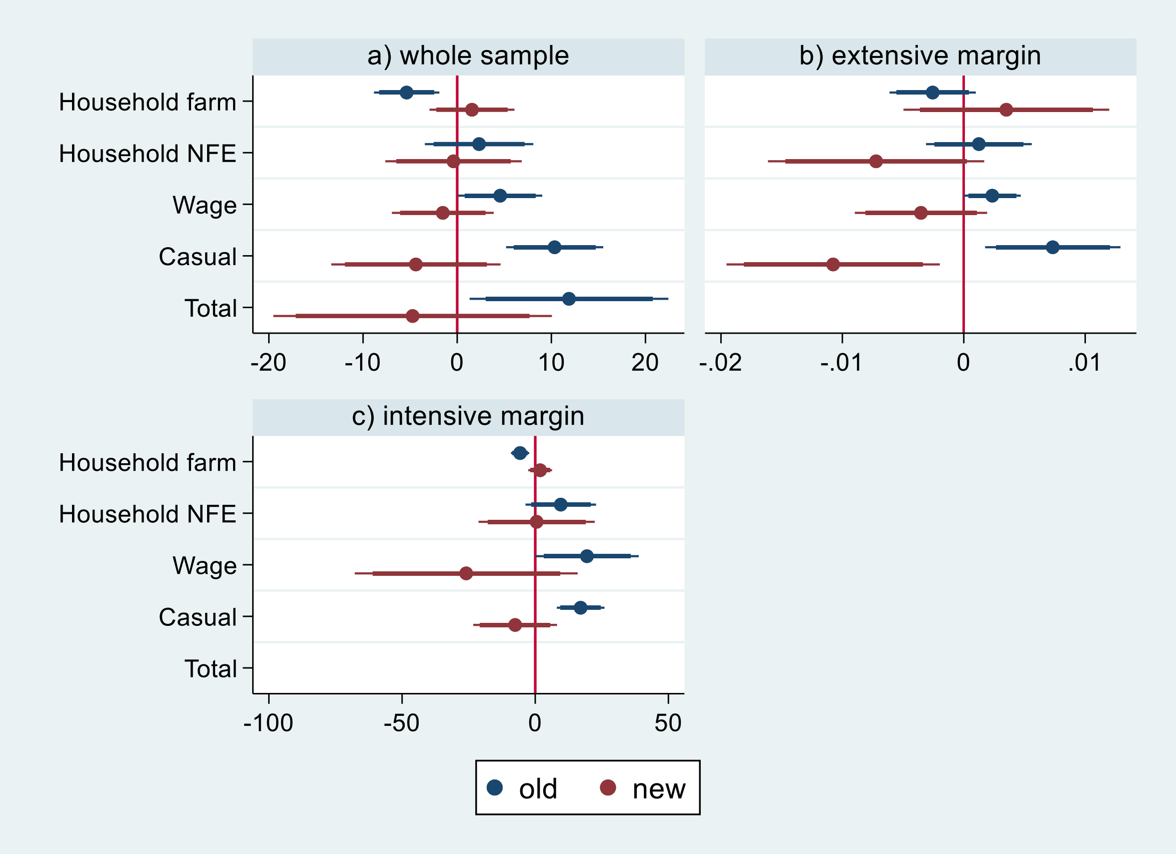
Source: FAO, Global Information and Early Warning Systems (GIEWS).
https://www.fao.org/giews/earthobservation/country/index.jsp?lang=en&code=MWI

Annex Figure 8. Drought Intensity of Cropland, Mean Vegetation Health Index (VHI) averaged per Gaul 2 region for complete season 1 of 2010, 2013, 2016 and 2019



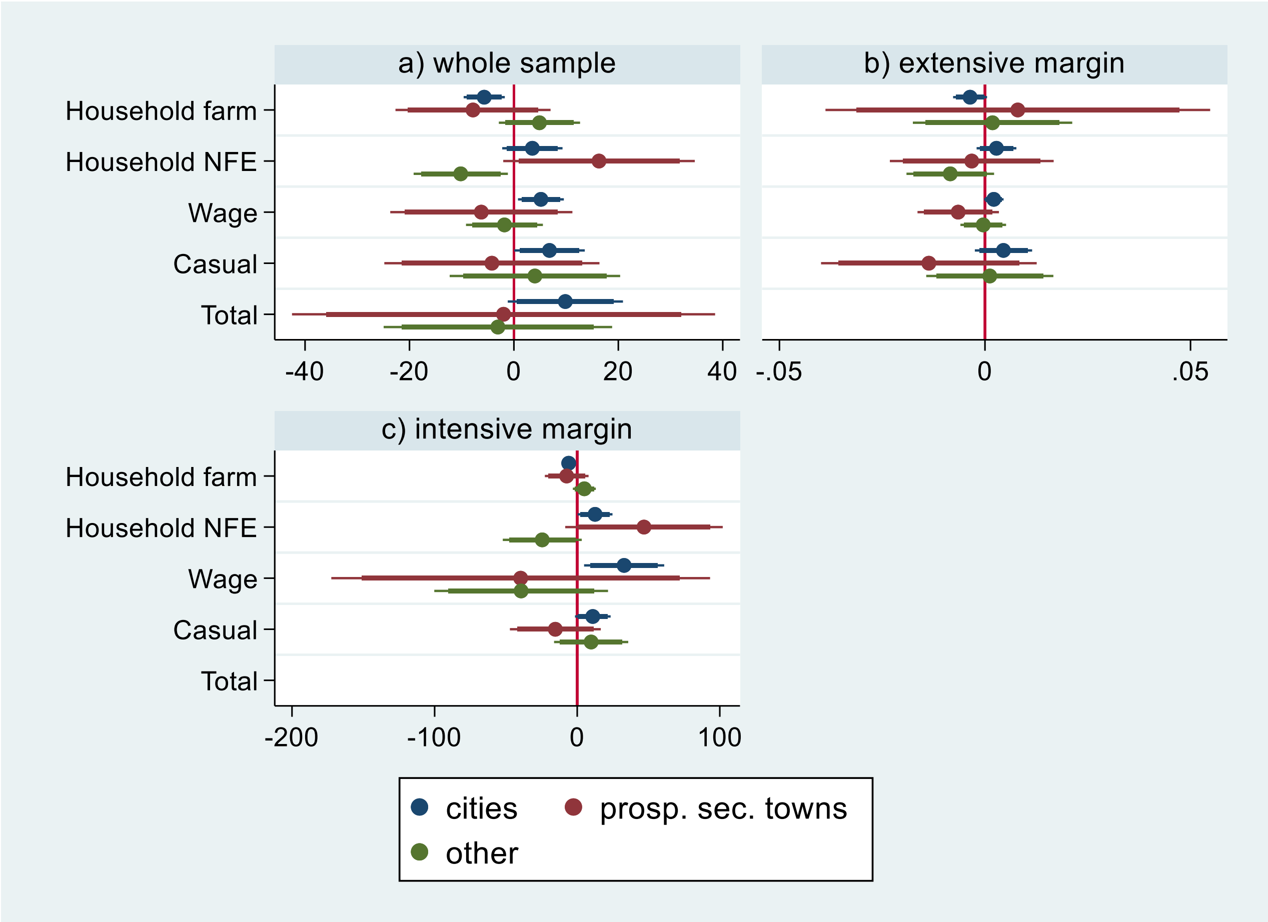
Source: FAO, Global Information and Early Warning Systems (GIEWS).
https://www.fao.org/giews/earthobservation/country/index.jsp?lang=en&code=MWI

Annex Figure 9. Effect of urbanization between 2010 and 2019 on hours worked, decomposed for ‘old’ and ‘new’ agglomerations



Source: Authors’ analysis.
Note: This figure represents the point estimates and 95% confidence intervals of β_1 and β_2 from the following specification, slightly adapted from (3): $\Delta l_{i,k,t} = \alpha_0 + \beta_1 \Delta lnUA_{k,t-1,old} + \beta_2 \Delta lnUA_{k,t-1,new} + \beta_n c_{i,n,k,t-1} + \varepsilon_{i,k,t}$ with $\Delta l_{i,k,t}$ being the change in labour supply over the period between t-1 and t for individual i living in rural area k, $\Delta lnUA_{k,t-1,old}$ being a measure of change of urban access in rural area k between t-2 and t-1 for the subset of ‘old’ cities (those 19 agglomerations that already existed before 2000), $\Delta lnUA_{k,t-1,new}$ being a measure of change of urban access in rural area k between t-2 and t-1 for the subset of ‘new’ cities (the 58 new agglomerations that only emerged after 2000), $c_{i,n,k,t-1}$ being a set of n baseline (individual) characteristics, and $\varepsilon_{i,k,t}$ as a random error term.
In panel a) each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various aggregations of employment categories.
Panel b) estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin).
Panel c) restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories.
Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a) and panel b): 1,390. Seventeen individuals dropped from sample due to missing education variables. Panel c): 1,356 for household farm labor, 400 for household non-farm enterprise labor, 181 for wage labor and 862 for casual labor.

Annex Figure 10. Effect of urbanization between 2010 and 2019 on hours worked, decomposed for three types of agglomerations: cities, prospective secondary towns, and other (smaller) agglomerations



Source: Authors' analysis.
Note: This figure represents the point estimates and 95% confidence intervals of β_1, β_2 and β_3 from the following specification, slightly adapted from (3): $\Delta l_{i,k,t} = \alpha_0 + \beta_1 \Delta lnUA_{k,t-1,city} + \beta_2 \Delta lnUA_{k,t-1,pst} + \beta_2 \Delta lnUA_{k,t-1,oth} + \beta_n c_{i,n,k,t-1} + \varepsilon_{i,k,t}$ with $\Delta l_{i,k,t}$ being the change in labour supply over the period between t-1 and t for individual i living in rural area k, $\Delta lnUA_{k,t-1,city}$ being a measure of change of urban access in rural area k between t-2 and t-1 for the subset of cities, $\Delta lnUA_{k,t-1,pst}$ being a measure of change of urban access in rural area k between t-2 and t-1 for the subset of prospective secondary towns, $\Delta lnUA_{k,t-1,oth}$ being a measure of change of urban access in rural area k between t-2 and t-1 for the subset of other agglomerations, $c_{i,n,k,t-1}$ being a set of n baseline (individual) characteristics, and $\varepsilon_{i,k,t}$ as a random error term.
In panel a) each regression is run on the full analytical sample and the left hand side represents the total number of hours worked in various aggregations of employment categories.
Panel b) estimates equation (3) as a Linear Probability Model (LPM) with on the left a dummy indicating whether or not the individual is engaged in the activity. The plotted coefficient therefore reflects the propensity of entering into (or exiting out of) a certain employment category with increasing urban access (the extensive margin).
Panel c) restricts the sample to individuals who have worked at least one hour in the employment category in question at either baseline or endline. The left hand side represents the total number of hours worked in various aggregations of employment categories.
Due to the level-log specification, all the coefficients can be interpreted as the effect of a one percent increase in urban access on hours worked.
Observations: Panel a) and panel b): 1,390. Seventeen individuals dropped from sample due to missing education variables. Panel c): 1,356 for household farm labor, 400 for household non-farm enterprise labor, 181 for wage labor and 862 for casual labor.

All working papers published by IFPRI's Malawi Strategy Support Program are available at https://massp.ifpri.info/resources/working-papers/

Funding for this work was provided by the Government of Flanders and the Irish Embassy in Malawi. This publication has been prepared as an output of Malawi Strategy Support Program (MaSSP) and has not been independently peer reviewed. Any opinions expressed here belong to the author(s) and are not necessarily representative of or endorsed by IFPRI.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

A world free of hunger and malnutrition

IFPRI is a CGIAR Research Center

IFPRI Malawi, Area 14 Office, Plot 14/205, Lilongwe, Malawi | Mailing Address: PO Box 31666, Lilongwe 3, Malawi

T +265-1-771-780 | Email: IFPRI-Lilongwe@cgiar.org | http://massp.ifpri.info