Digital Infrastructure and Local Economic Growth: Early Internet in Sub-Saharan Africa*

Moritz Goldbeck† Valentin Lindlacher‡

Preliminary Draft
May 23, 2021

Abstract

Digital infrastructure enables widespread access to modern information and communication technologies—most prominently the Internet—promising to stimulate economic growth. We analyze 10 Sub-Saharan African (SSA) countries in the early 2000s and ask if Internet availability even at basic speeds contributes to regional economic development. Exploiting quasi-random variation in Internet availability induced by sub-marine cable arrivals in a difference-in-differences setting, we measure the growth of SSA towns using nighttime light data. Our findings suggest that Internet availability at basic speeds leads to about two percentage points higher economic growth of SSA towns in the years after connection compared to a control group of similar cities not (yet) connected.

Keywords: Internet, regional/urban development, cities, nighttime light, Sub-Saharan Africa

JEL-Codes: O33, O18, R11

---

*We thank Vojtech Bartos, Mathias Bühler, Oliver Falck, Jonas Hjort, Anna Kerkhof, Markus Ludwig, Niklas Potrafke, Helmut Rainer and Maria Waldinger for valuable comments and suggestions. Further, we thank Nicolas Göller for excellent research assistance. We are thankful for valuable comments at the 14th RGS Doctoral Conference in Economics, the third International Workshop "Market Studies and Spatial Economics", and the 10th European Meeting of the Urban Economics Association, especially, to the discussants Kathrin Wernsdorf and Marcelo Sant'Anna. Finally, we thank the participants of the (internal) seminars at the ifo Institute and at Economic Research South Africa. All errors are our own.

†University of Munich & ifo Institute, goldbeck@ifo.de.

‡University of Munich & ifo Institute, lindlacher@ifo.de.

1 Introduction

In the last decades, the provision of digital infrastructure in many countries enabled widespread access and adoption of modern information and communication technologies (ICT), most prominently the Internet. Evidence shows positive effects of (broadband) Internet availability on individual-level economic performance (Akerman et al., 2015) and country-level economic growth (Czernich et al., 2011) for developed countries. Hopes are high that Internet access can foster regional economic growth in the developing world as well (World Bank, 2016). In Sub-Saharan Africa (SSA) for example—where impulses for economic growth are required to fight poverty and deprivation—Governments, public-private partnerships, and companies alike invest large amounts of money to bring the Internet to everyone. To date, SSA countries invested more than 28 billion US-Dollar into their national Internet backbone network (Hamilton Research, 2020). Facebook recently announced an effort to build a new sub-marine Internet cable to Africa for one billion US-Dollar (Bloomberg, 2020). And China plans to invest more than 60 billion US-Dollar in Africa's digital infrastructure as part of its Belt-and-Road initiative (Invesco, 2019).

Despite these enormous investments, a growth effect of Internet in SSA is less obvious than it seems. On the one hand, lacking legacy infrastructure (i.e., fixed-line telephony networks) to build on makes the provision more complex and costly. At the same time, low population density apart from a few mega-cities, missing hardware, financial constraints and a lower willingness to pay lead to lower adoption rates (World Bank, 2016). On the other hand, the potential of Internet seems particularly high in SSA since alternative ICT like fixed-line telephony is largely absent. It is thus crucial to investigate the effect of Internet availability on regional economic development in a developing-country context. With the notable exception of Hjort and Poulsen (2019), who find sizable positive individual-level effects of a broadband Internet speed upgrade on employment in SSA around 2010, evidence on the economic impact of Internet availability in developing countries is surprisingly rare.

In this paper, we ask if there is a causal effect of Internet availability on local economic growth in SSA even at basic speeds. We focus on the initial introduction of Internet in SSA since the early 2000s through the first wave of Internet-enabled sub-marine cables, which made basic Internet speed available to SSA countries. To investigate if potential individual-level effects matter for the economic development of entire localities, we conduct our analysis at the town level. Further, we provide suggestive evidence on the mechanisms driving the effect.

To identify the causal effect of Internet availability on local economic growth, we exploit quasi-random variation in the timing of Internet availability induced by the arrival of the first sub-marine Internet cables (SMCs) in SSA in the early 2000s. This approach was established by Hjort and Poulsen (2019), who use an Internet speed upgrade in SSA induced by SMCs around 2010. In a difference-in-differences setting, we then compare the growth of towns with Internet access at the time of SMC arrival to a control group of similar towns getting access only later. To ensure our results are in fact driven by Internet availability induced by SMC arrival, we control for the roll-out of other (potentially confounding) infrastructure.

We measure economic growth in each SSA town by nighttime light intensity data captured by satellites, a well-established proxy introduced by Henderson et al. (2011) and validated e.g. by

Storeygard (2016) and Bruederle and Hodler (2018). Town-level Internet availability is determined by data on the location of access points (APs) to the national Internet backbone. Because existing data on the location and establishment year of APs in SSA only starts in 2009 (Hamilton Research, 2020), we backdate the establishment year of APs to their actual construction year by hand via an extensive review of network deployment projects for each SSA country.

This is, to the best of our knowledge, the first study investigating the effect of Internet at basic speeds on economic growth in the developing world. At comparable speeds, Czernich et al. (2011) find positive growth effects of Internet in developed countries. In our study, we estimate Internet effects in a setting with no pre-existing fixed-line telephony network. Therefore, our results show that the Internet induces growth even when penetration is low. With our approach utilizing night-time light satellite data we are able to capture towns in 10 SSA countries getting an SMC connection in the 2000s. Thus, our results suggest that the Internet fosters regional development at the town-level in a comprehensive set of SSA towns.

We find that connection to the Internet on average leads to a 8% increase in light intensity of SSA towns in the years after connection, which approximately translates into 2 percentage points higher economic growth.[1] Moreover, we differentiate growth in the number of lit pixels, indicating a spatial expansion of towns ("extensive margin"), and growth in brightness, which is associated with a higher density of economic activity in the towns ("intensive margin"). We find that towns with Internet access are becoming both brighter and larger. This provides suggestive evidence that cities with Internet access grow at the intensive margin as well as the extensive margin, i.e. geographically. Controlling for population further suggests that the observed growth is partly driven by migration. Our findings therefore point to the increased economic activity being a result of both an increase in per capita economic activity and migration into connected towns.

We contribute to two main strands of the literature. First, we add to the literature assessing the impact of infrastructure investments on economic outcomes. There is ample evidence that investments in transportation infrastructure have long-lasting effects on regional growth and development (see e.g., Hornung, 2015; Allen and Arkolakis, 2019; Banerjee et al., 2020). The effect of digital infrastructure and especially (broadband) Internet has been assessed by Czernich et al. (2011) and Akerman et al. (2015), who show that firms and workers in connected regions become more productive when they have access to Internet. Closely related to our work is Hjort and Poulsen (2019), who study this in a developing-country context and find an Internet speed upgrade to have a skill-biased and net positive employment effect in SSA. Our analysis contributes to these findings by showing that the benefits of digital infrastructure are present not only at the individual level but at the more aggregate town level as well. We further show that even the availability of basic Internet speeds adopted by few individuals and businesses is beneficial for regional development.

Second, our work contributes to the literature on urban and regional development. Economic productivity is typically higher in agglomerations for several reasons, for example, thick local labor markets, knowledge spillovers, low transportation costs, but also due to local provision of amenities (see e.g., Albouy, 2016; Clark et al., 2002; Deller et al., 2008). Thinking of digital infrastructure as local

[1] Henderson et al. (2011) finds an elasticity of GDP-to-light of 0.284.

amenity, our findings indicate that the benefits of Internet availability are spatially highly concentrated, accruing to connected locations (towns) only, with important implications for regional inequality both across cities and between cities and rural areas.

The paper proceeds as follows. In Section 2, we provide a brief overview of the early Internet in Sub-Saharan Africa. Section 3 lays out the empirical strategy and in Section 4 the data and spatial methods are described. Results are presented in Section 5. Section 6 concludes.

2 Background

There are three major components of Internet infrastructure determining availability and bandwidth of Internet in a given location. First, international sub-marine fiber-optic cables connect SSA countries to the global Internet backbone. Second, within-country inter-regional fiber cables form the national backbone. Precondition for Internet availability in a location is an access point (AP) to the national backbone. Finally, individual users in a location are reached via the 'last mile' infrastructure.

2.1 International backbone: sub-marine cables

Since the vast majority of web pages and applications is hosted on servers located in North America or Europe, almost all African Internet traffic is routed inter-continentally (Kende and Rose, 2015; Chavula et al., 2015). Before the first fiber-optic sub-marine cables (SMCs) landed on African shores, the only way to connect to the Internet on the continent was via satellite.[2] While being largely unconstrained by geography and local infrastructure, satellite connection is costly and allows only for very narrow bandwidths. With SMCs—often a joint effort of governments, private investors, and/or multinational organizations—Internet connection was first brought to SSA at a noticeable scale.

As shown in Figure 1, the first wave of internet-enabled SMCs arrived in SSA countries only in the early 2000s.[3] These 'first-generation' cables had the capacity to provide Internet at basic speeds. The biggest of them was SAT-3 and started operating in 2001. It featured landing points on the shores of eight SSA countries at the western coast of Africa.[4] These landing points—typically one per country—constitute the starting point for the respective national backbones (cf. Section 2.2). Until the late 2000s, most SSA countries were connected to the Internet via these 'first-generation' SMCs.[5]

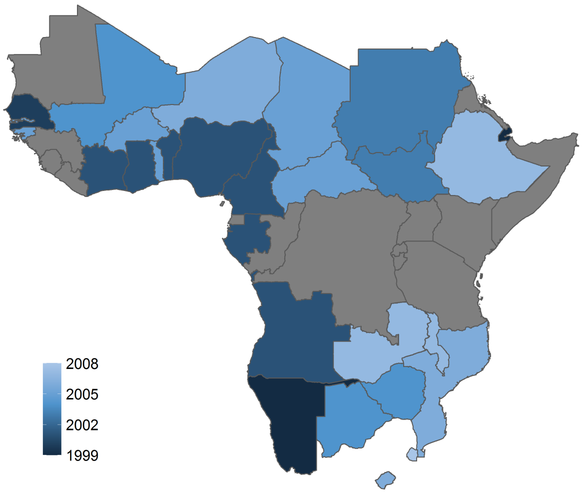
[2] Single-channel and co-axial SMCs for telegraphy and telephony already existed before. The first telegraphy cable ('East coast' cable) started operating as early as 1879.

[3] South Africa was connected in 1993, when the first SMC constructed to enable the use of the Internet (SAT-2) preceded an old co-axial telephone cable established in 1968 (SAT-1).

[4] Benin, Cameroon, Côte d'Ivoire, Gabon, Ghana, Nigeria, Senegal, and South Africa.

[5] The 'second-generation' of SMC landed very similarly between 2009 and 2012.

Figure 1: Sub-marine Cable Connection Years



Notes: Figure shows SSA countries with SSA connection years before 2008. Darker blue colors indicate earlier initial SMC connection years. Gray indicates countries not connected by SMC until 2008.

2.2 National backbone: inter-regional cables

After being routed through a SMC, Internet traffic travels through the national backbone. The national backbone infrastructure consists of inter-regional (fiber) cables. Therefore, as soon as a new SMC arrives at a landing point of a SSA country, Internet becomes available country-wide in every location with access to the national backbone. As Internet capacity (i.e. speed) of the national backbone does not depend substantially on distance to the landing point, this upward shift occurs uniformly across the country's connected locations. In the last decades, national backbone networks were continuously improved and expanded in parallel to the installation of SMCs.[6] This network expansion focused heavily on connecting economically and/or politically important locations since they feature the largest market potential (high population density and GDP per capita).[7] This often lead to a network evolution where the national capital (often a coastal city) was connected first. Then, the network spread out to the next largest (or politically important) cities, which are often regional capitals or other large cities. Due to their role as nodes in the national backbone networks, we call these cities 'nodal cities'.

Inter-regional cables are almost always built along pre-existing infrastructure (roads, but also railroads, the electric grid, and pipelines) to minimize construction costs. Even though the goal was to connect nodal cities, in many cases, towns on the route of inter-regional cables incidentally got Internet access as well. Our empirical strategy (cf. Section 3) focuses on these on-route towns which

[6] Many of these cables were built decades ago as part of the telegraph and telephone infrastructure and were only later used for the transmission of early Internet traffic. They typically have been installed by the national telecom. Each country typically has an own, self-contained backbone network.There are no network operators owning networks in more than one country.

[7] Routes establishing connections to (landlocked) neighboring countries are a focus of network expansion as well.

get Internet connection because of their location next to an inter-regional cable but are not nodal cities themselves.

2.3 Local transmission: 'last mile' infrastructure

Internet traffic transported by inter-regional cables is accessed at 'access points' (APs). There are several technologies transmitting Internet traffic from these access points (so-called fiber nodes) to the user. These 'last mile' transmission technologies include fiber cables (FTTH/B), copper cables, and wireless transmission using cellular towers (e.g. mobile or WiMax). Unlike in many developed nations which rely heavily on transmission to the end user via pre-existing telephony cable infrastructure, in SSA countries the Internet is mostly accessed via wireless and/or mobile devices. For this technology, no local cable network connecting each users exact locality (firm, household) is needed. Traffic data is exchanged wirelessly between cellular towers and the user's device. Relative to the costs to construct an inter-regional cable it is thus cheap to establish Internet access along the cable, making it profitable for the network operator to establish APs even in on-route towns, which are typically much smaller than nodal cities.

3 Empirical strategy

We are interested in the relationship between Internet availability and local economic growth. However, their correlation is not informative about the causal effect of Internet availability on local economic growth due to endogeneity concerns. In particular, towns with and without Internet access might be very different. Internet access is not randomly assigned and likely driven by commercial interest and/or political and administrative planning. Decision makers might prioritize the connection of a priori larger towns on a higher growth path above smaller stagnating towns.

To address these endogeneity concerns, we leverage two distinct features of Internet infrastructure evolution in SSA countries. First, we use plausibly exogenous time variation in connections to submarine cables (SMCs), which determine Internet availability nation-wide, to investigate the effect of Internet availability. Following Hjort and Poulsen (2019), we argue that the exact timing of SMC arrival is essentially random. First, for each individual SMC, it is exogenous to national planners when exactly it is put into operation and starts providing (international) Internet connection. This is because each SMC typically connects many countries and the connection years are highly uncertain due to unforeseen delays in construction and coordination difficulties among consortium members.[8] Second, within SMCs the connection years are mainly determined by the geographical location because SMCs come from Europe, either through the west passing Spain or Portugal or trough the east passing Egypt or the Arabic peninsula, and connect SSA countries according to their location at Africa's coastline. Moreover, landlocked countries get their connection through the backbone network of their neighboring countries and rely therefore on the construction speed there. This construction speed again is exogenous for the respective landlocked country.

[8] For example, the cable EASSy was delayed by five years due to coordination difficulties among consortium members (**?**).

Second, we focus on incidentally connected towns, i.e., towns close to an access point (AP) that are not (endogenously connected) nodal cities. We focus our analysis on these towns because they are very similar to each other in key characteristics such as size and infrastructure. Therefore, we exclude nodal cities: the landing point, the capital, regional capitals, and economic centers (cities with a population of more than 100,000 inhabitants). Importantly, all towns in our analysis get connected eventually, mainly because of their favourable location between nodal cities. Hence, we also exclude towns that are still waiting for an Internet connection today.

This enables us to use a difference-in-differences (DiD) design. We compare on-route towns that already have access to the national backbone when the SMC arrives to a control group of similar on-route towns getting an access point in later years (first difference) before and after the arrival of a SMC (second difference). Since national backbones are self-contained and (coastal) countries have own landing points (cf. Section 2.2) each country has a specific treatment date. Therefore, depending on the nation-wide connection date, towns in different countries are treated at different points in time.

The basic model used to identify the average treatment effect on the treated (ATT) of Internet availability on local economic growth is given by

$$y_{ict} = \beta_0 + \beta_1 \left(smc_{ct} \times access_{ic} \right) + \boldsymbol{X}_{ict}\beta + \alpha_{ic} + \delta_{ct} + \epsilon_{ict} \tag{3.1}$$

where y_{ict} is economic growth of town i in country c in calendar year t as proxied by nighttime-light luminosity measured by satellites (cf. Section 4). The dummy variable smc_{ct} indicates if country c has access to a SMC in calendar year t. The variable $access_{ic}$ is one if town i in country c is located within 10km distance to an access point that was established in the year of SMC arrival or before. Contrary, the indicator is zero if town i in country c is located close to an access point that was established in the years after the SMC arrival.[9] Thus, the interaction term $smc_{ct} \times access_{ic}$ indicates Internet availability in town i in country c in calendar year t. The coefficient of interest is β_1. It captures the effect of Internet availability on local economic activity. $\boldsymbol{X}_{ict}$ contains time-varying control variables, such as mobile internet coverage. We include two types of fixed effects into the model. Time-constant differences across towns are captured by town fixed effects α_{ic}. Differences across calendar years common to all towns within a country are absorbed by country-year fixed effects δ_{ct}. Note that this allows for country-specific time trends and variations in satellite sensor quality over years. ϵ_{ict} is an error term. Like in many other DiD applications, our panel data are serially correlated in the time dimension. Hence, we use cluster-robust standard errors whereby we cluster at the town level.

The key identifying assumption for this DiD model is that treatment and control towns would have evolved similarly in the absence of treatment (parallel trends assumption). This assumption cannot be tested. Its plausibility can, however, be examined by testing for pre-treatment differences in time trends between the treatment and control group. Therefore, we look at the dynamic impact of Internet

[9] Towns within 10km distance to an access point that was established in the three years under observation after the arrival of the SMC are excluded from the control group. Otherwise, due to getting treated in the observation period they would confound our analysis.

availability on local economic activity using event studies:

$$y_{ict} = \beta_0 + \sum_{j=\underline{T}}^{\overline{T}} \beta_{1j}\left(t_j \times access_{ic}\right) + \boldsymbol{X}_{ict}\beta + \alpha_{ic} + \delta_{ct} + \epsilon_{ict} \tag{3.2}$$

where t_j indicates the year relative to treatment year, i.e. the year when the SMC arrives, starting in relative year $j = \underline{T}$ and ending with relative year $j = \overline{T}$. The treatment year is normalized to $j = 0$. We exclude $j = -1$ as the reference point. Thus, the interaction $t_j \times access_{ic}$ indicates if town i in country c is part of the treatment group and restricts the coefficient to one particular relative year j. The coefficients β_{1j} inform about the dynamic effect of Internet availability. Thereby, each coefficient captures relative-year-specific treatment effects. We expect to see no effect before the treatment. Thus, if we cannot distinguish the estimates of the coefficients on the pre-treatment relative-year dummies (i.e. $\hat{\beta}_{1j} \forall j < 0$) from zero, the treatment and control group follow similar trends before the treatment, supporting the common trends assumption.

4 Data and spatial methods

We analyze the effect of Internet availability on local economic growth in SSA.[10] To this end, we tap two main data sources. First, local economic activity is measured by nighttime-light (NTL) satellite data. Second, locations connected to the Internet are identified via the geo-location of access points (APs). Moreover, we use data on towns' built-up area, on towns' administrative and economic importance, whether they are a (regional) capital and their number of inhabitants, and on infrastructure, such as roads. This section describes the data sets we use, the processing steps we apply, and shows descriptive statistics relevant for our analysis.

4.1 Local economic activity: nighttime-lights

Since geographically and chronologically granular data on economic activity in SSA is lacking, we deploy nighttime-light (NTL) data. This data makes it possible to measure human-caused nighttime-light emissions in a geographically high resolution and on a yearly basis. There have been two major programs that collected NTL data. First NTL data was collected in the *Defense Meteorological Satellite Program* (DMSP) *Operational Linescan System* (OLS) between 1992 and 2013. The follow-up program *Visible Infrared Imaging Radiometer Suite* (VIIRS) on the *Suomi National Polar-orbiting Partnership* satellite started in April 2012. We use the harmonization of the two sources by Li et al. (2020) to get consistent yearly NTL data from 1992 to 2018.

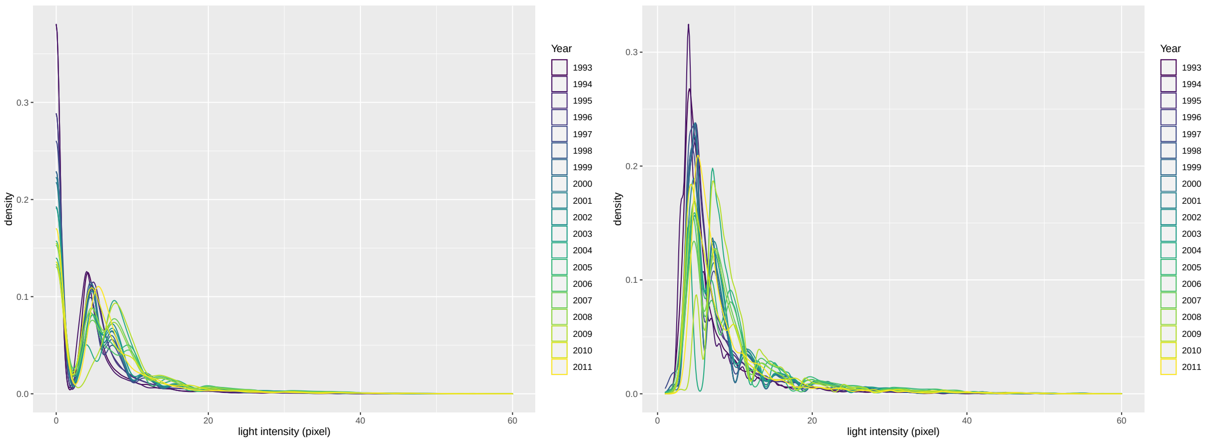
The instruments of DMPS-OLS satellites measure light intensity on an integer scale from 0 to 63 with pixels covering 30 arc-second grid-cells (an area of .86 km^2 at the equator). The data is then combined to yearly composite images. With VIIRS, both the spatial (pixels are smaller) and radiometric resolutions (both dark and bright spots are recorded better) as well as the temporal resolution have been improved. To harmonize these two sources, Li et al. (2020) recalculated the newer and better VIIRS data to the (spatial and radiometric) resolution of the original DMPS-OLS data. Therefore, the

[10] We exclude all islands, Northern African states, and South Africa.

lower DMPS-OLS resolutions apply in our case. An advantage of this procedure is that noise from aurora, fires, boats, and other temporal lights are excluded. Another advantage of this data is that the DMPS-OLS data has been inter-calibrated globally from 1992 to 2013 as well, making it temporally consistent. Unlike in the developed world, very high light intensities (i.e. top-coded pixels) are less a concern in the context of SSA (Bluhm and Krause, 2018).[11]

Figure 2 shows the distribution of light intensity by year for our estimation sample. In the left panel non-lit pixels are included, while in the right panel the distribution of lit pixels only is shown. For earlier years, the maximum density is typically smaller than for later years. Hence, the towns in our analysis are generally getting brighter over time. Most lit pixels have a light intensity of 10 or lower, reflecting the large share of rural locations with small- and mid-sized towns. Hence, as stated before, top coded pixels are not an issue in our setting.

Figure 2: Distribution of pixels' light intensity by year



(1) including non-lit pixels

(2) excluding non-lit pixels

Notes: Each figure depicts kernel density plots pixels' light intensities by year.

On the country level, NTL data is well established as a measure of economic activity and widely used by economists (Henderson et al. (2012) and Chen and Nordhaus (2011) among the first ones). Closely related to our work, Storeygard (2016) established this data on city level. On a broader area, Bruederle and Hodler (2018) added the relation to household wealth, education, and health for DHS cluster locations as well as for grid cells of roughly 50×50 km.

Within each town, we define several outcome measures. Local economic activity is measured by summing the light intensity of all pixels within a town in each year. This measure was established by Storeygard (2016) and accounts for both geographical extension and light intensity. As alternative measures, we calculate the sum of all lit pixels, instead of their intensity, in each year within a town (extensive margin) and the average light intensity of pixels within each town and year (intensive margin). We interpret the sum of lit pixels as a proxy for spatial extension of a town and the average light intensity as a proxy for density in terms of population or per capita economic activity.[12]

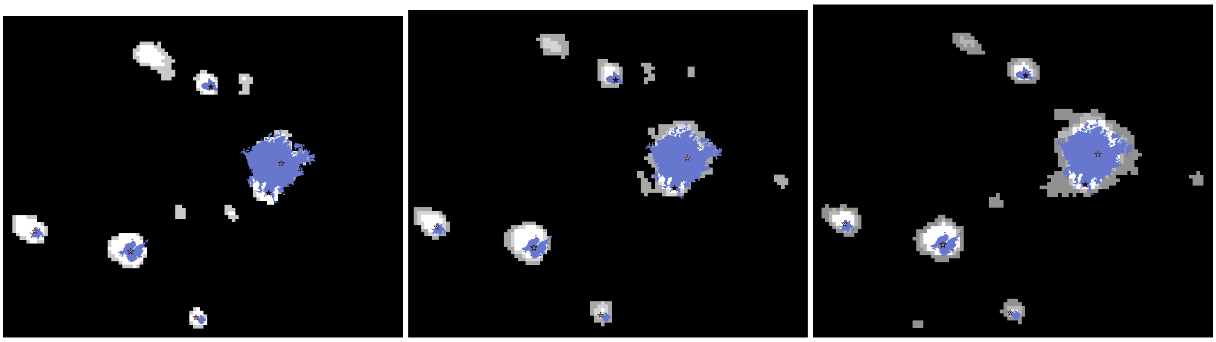
[11] Looking at pixels in our sample that only contains cities and towns, less than 2% of pixels are assigned light intensities of 60 or more.

[12] As specified in Section 3, we apply the logarithm of each outcome measure.

We measure economic activity at the town level. There exist several approaches to define towns spatially.We rely on established data from *Africapolis.*[13] This database contains the geographical delineation of 5,811 SSA agglomerations with more than 10,000 inhabitants in 2015. The delineation is based on a classification algorithm detecting built-up areas. The median size is around 20.000 inhabitants and about 90 percent have less than 100.000 inhabitants. This approach defines the extent of an agglomeration in 2015. Though agglomerations might have been growing after our observation period, for some agglomerations the NTLs still might blur over the extent of the built-up areas. Therefore, we account for blurring by adding a radius of 2km to the built-up areas so that extensive margin growth is properly captured.[14]

Figure 3 shows the evolution of towns in Senegal (which was connected by a SMC in 2000). The panels depict the years 1995, 2000, and 2005, respectively. Darker gray reflects lower light intensity values. The blue area is the built-up area from *Africapolis* (without the 2 km radius). This area is not changing over time. The yellow stars represent APs. In this example, the town with an AP gets brighter and also grows spatially over time while the other towns stagnate, shrink or even disappear entirely.

Figure 3: Evolvement of night-time light in Senegal over time



(1) 1995 (2) 2000 (3) 2005

Notes: Panels (1)-(3) show night-time lights and built-up areas in the same region of Senegal for the years 1995, 2000, and 2005, respectively. The black-to-white scale indicates light intensity, with brighter colors reflecting higher light intensities. Built-up areas from *Africapolis* are shown in blue. Access points are indicated by a yellow star.

4.2 Backbone infrastructure: access points and sub-marine cables

The geo-location of the access points in the national backbone comes from *Africa Bandwidth Maps.*[15] The database contains the most comprehensive set of APs for Africa and covers the period starting from 2009 and is updated on a yearly basis. The data in this database is directly sourced from the network operators. As APs existing in 2009 were largely established earlier, we conduct an extensive review of network deployment projects for each country. Thereby, we are able to determine the construction years of the access points from 2009 going back to the late 1990s.This makes it possible to identify

[13] https://africapolis.org

[14] For robustness, we also show the results using different radii on the built-up areas, including a specification without a radius.

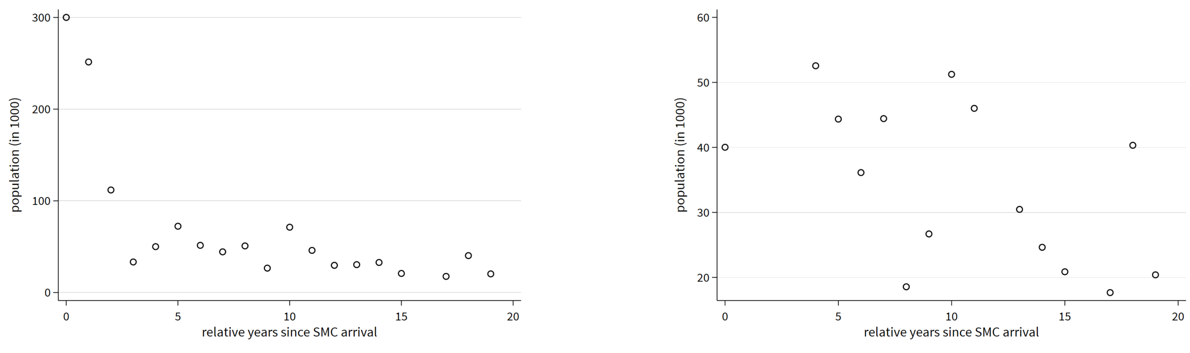
[15] http://www.africabandwidthmaps.com

which towns already were connected to the national Internet backbone when the first wave of SMCs arrived. We match the APs to the towns via their geo-location: We calculate the distance between the towns' border and the closest AP. We assign national Internet backbone connection to towns within a distance of less than 10 km.[16]

To date, there are 2.708 APs in SSA countries.[17] About half of them were constructed since 2013. Especially in bigger cities, more than one AP is usually built to account for the limited capacities of each AP. Around 900 APs were built in different cities and towns. In 2019, for example, although 189 new APs were constructed, only 27 new cities and towns were connected. Before 2009, 87 (regional) capitals were connected, that is about one third of the connected cities and towns. Additionally, 26 cities with more than 100.000 inhabitants were connected before 2009. These cities fall into our definition of nodal cities (c.f. Section 3) and are not considered in the analysis due to endogeneity concerns. This leaves 207 towns in the main specification of our analysis.

Figure 4 compares the average size of cities and towns by the year they get their AP (relative to the SMC connection year). In the early years, notably bigger cities are connected. These cities are usually nodal cities (i.e. the capital, other cities close to the landing point of the SMC, and other important political and economical centers like regional capitals). In subsequent years, cities and towns are on average a lot smaller and vary for all subsequent years around a population of 50,000 inhabitants. However, when only examining treated and control towns—i.e. excluding nodal cities—there is no clear pattern of size over time: treated towns have on average a population of 40,000 inhabitants and control towns also vary around this value (in some years up to more than 50,000 inhabitants and in other years down to less than 30,000 inhabitants). This suggest that treated towns are not selected to get treatment because of their population size.

Figure 4: Population size of connected cities and towns (without nodal cities) by year (relative to connection year)



(1) Whole sample (including nodal cities)

(2) Estimation sample (excluding nodal cities

Notes: Panels (1) and (2) show the average population size of connected cities and towns by year relative to connection year. Panel (1) shows all connected cities and towns whereas in Panel (2) nodel cities are excluded.

[16] We conducted interviews with industry experts to verify this decision. In addition, in a robustness check we vary this distance.

[17] Their location and establishment years are mapped in Figure A.1 in the Appendix.

To measure the time of treatment, we use information on SMCs landing dates on the shores of SSA countries.The data comes primarily from *Submarine Cable Map*.[18] In particular, we use the country-specific connection dates and the geo-location of the landing points.[19]

4.3 Descriptive statistics

Before discussing the estimation results, we provide descriptive statistics to compare treated and untreated towns. We focus on early SMCs bringing basic Internet speeds to SSA, and therefore drop countries which were connected after 2008 for the first time.[20] This leaves 27 SSA countries. Among the first countries are Djibouti, where an SMC landed in 1999, Namibia, which was connected by a trans-national fiber cable from South Africa also in 1999, and Senegal, which was connected in 2000 with an individual-country SMC. In 2001, seven more countries were connected by a single SMC, the SAT-3 cable. In the following years, 17 more countries got an SMC connection or were connected through their neighboring countries until 2008.

However, not all countries that were connected until 2008 had built a national backbone infrastructure when the respective SMC or the connection through a neighboring country arrived. In this case, the treatment group is missing as there are no towns with national backbone connection right after the connection. This reduces the number of countries in our analysis to 24.[21] Moreover, 11 countries built at least one AP before the SMC arrival but only in 'nodal cities'.[22] Finally, Namibia is dropped as it did not construct APs after getting Internet access so that we cannot define a control group. This leaves 12 countries for our analysis.

Our sample represents an unbalanced panel. From the sample, the first connected country is Senegal, which was connected in 2000, which leaves us with 7 pre-treatment years. Hence, on the left with a truncation at -7 years, we have no data loss on the country-side. Though, it limits our pre-treatment comparison to a period of seven years. Malawi and Mozambique only have 2 post-treatment years. Hence, estimating on a balanced sample with 3 post-treatment years leaves us with a sample of 10 countries.[23].

We first plot light intensity for the whole sample (Figure A.2) and then focus on the treatment and control group (Figure 5). In Figure A.2, besides the treatment and control group, all from our analysis the excluded cities are shown. It can be seen that the capitals are by far the biggest cities and that cities that do not get an access point by the end of our data are the smallest on average. However, they are not a lot smaller than the cities in the control group. Additionally, the other excluded cities, the regional capitals and the cities with an population of at least 100,000 inhabitants are rather similar in comparison to the cities in our analysis. It can be noted that the regional capitals on average do not

[18] https://www.submarinecablemap.com

[19] A broader overview is given in Section 2.1

[20] The next generation SMCs landed between 2009 and 2012 and brought much higher Internet speeds.

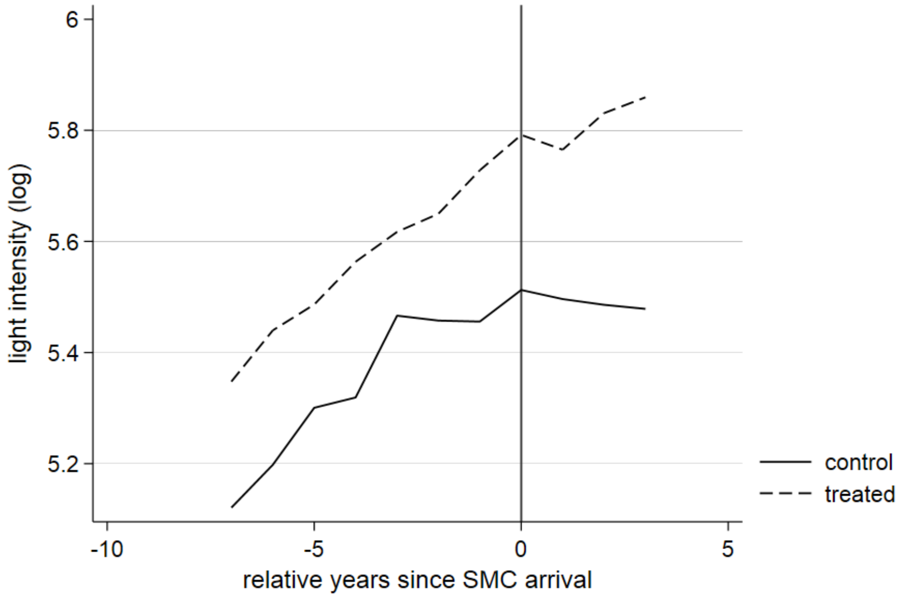
[21] Central African Republic has not yet built a national backbone infrastructure. In Djibouti, the first APs were built in 2007, which is 8 years after the first SMC connection. Nigeria built its first APs in 2003, which is 2 years after the arrival of the first SMC. This means especially that the biggest two SSA economies, South Africa and Nigeria, are not included in our analysis.

[22] These countries are: Angola, Chad, Côte d'Ivoire, Cameroon, Gabon, Gambia, Ghana, Niger, Guinea-Bissau, Lesotho, and Swaziland. The last three built all APs until today only in 'nodal cities'. Guinea-Bissau, for instance, only built APs in 2005, the connection year.

[23] These countries are Angola, Benin, Botswana, Ethiopia, Mali, Sudan, Senegal, Togo, Zambia, and Zimbabwe

exceed the bigger cities category. It is hard to see growth patterns in this aggregated figure. Therefore, we show a figure of just the treatment and control group next. Figure 5 shows that in the early pre-treatment years both groups grew with similar rates. In more recent years before the treatment, this growth stagnated. This stagnation holds on for the control group, while the treated cities grew again in the second and third year after the treatment.

Figure 5: Time Trends of Treatment and Control Group



Next, we compare both groups in their characteristics (Table 1). Variables are either geographic, whether they are located at a national border, at the coast, or at a river, or they compare the local infrastructure, whether there is an airport, a port, a bank, a specific type of (rail)road, or educational or health institutions. We can show that both groups are very similar. The biggest difference is that in the treatment group, more cities are located at the coast than in the control group. The variables come either from *Natural Earth* or *Open Street Map*. Therefore, the data is time-invariant and not dated before the treatment but was collected in 2010 or more recently.

We focus our analysis on mid-sized cities. Therefore, we set a focus on the comparison of the population between the two groups. The population data is coming with the built-up areas from *Africapolis*. Figure 6 shows, first, that both groups are very similar in their population distribution, and second, that mid-sized cities have a population of around 30,000 inhabitants. There are 285 cities excluded from the analysis, either because they are 'nodal cities' or cities that did not build an AP until 2019. These cities are on average (144.000 inhabitants) bigger than the cities of our analysis. 67 cities of them have more than 100.000 inhabitants.

Finally, we show the geographic distribution of the treatment and control group cities (Figure 7). For many countries, there is no specific pattern. Sudan is a special case as 10 cities are in the control group but only one city is in the treatment group. One should also note that of the 10 countries 4 are coastal, while 6 are landlocked.

Table 1: Means Comparison between Treatment and Control Group

	Mean		
	Control	Treated	Δ
Area	9.910	10.832	0.923
Population			
2015	39,766	38,316	-1,450
2010	30,140	29,742	-397
2000	20,918	19,569	-1,349
Health			
Hospital	0.162	0.115	-0.047
Pharmacy	0.206	0.246	0.040
Doctor	0.029	0.131	0.102**
Transportation			
Railroads_ne	0.294	0.410	0.116
Roads (natural earth)	0.721	0.738	0.017
Roads (major 1)	0.000	0.033	0.033
Roads (major 2)	0.456	0.492	0.036
Border crossings	0.000	0.016	0.016
Coast	0.044	0.180	0.136**
River	0.059	0.049	-0.010
Port	0.000	0.000	0.000
Education & Other			
School	0.324	0.410	0.086
College	0.088	0.115	0.027
University	0.015	0.049	0.034
Bank	0.382	0.410	0.027
Cities	120	87	

Notes: Asterisks in column Δ report t-test for equality of means: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 6: Population Density in Treated and Control Cities

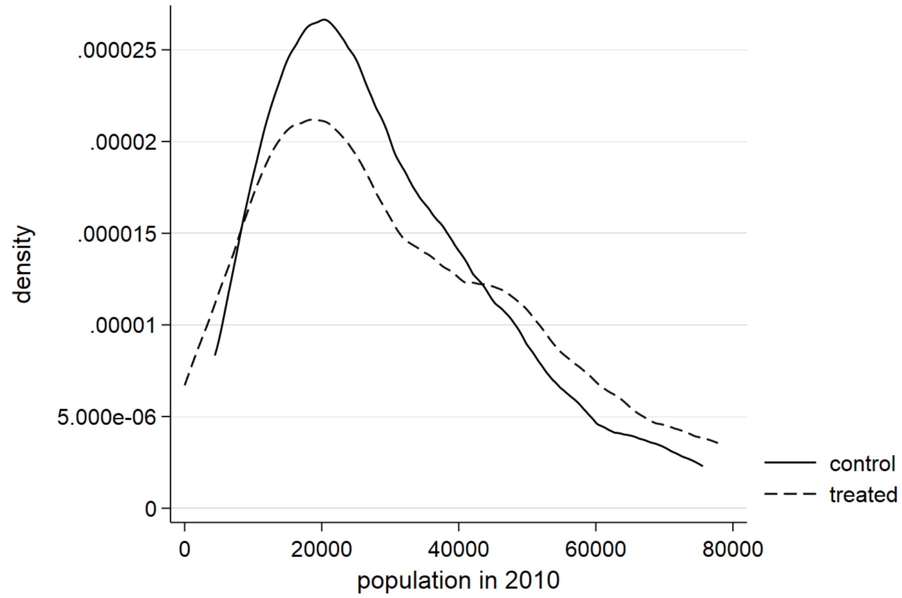
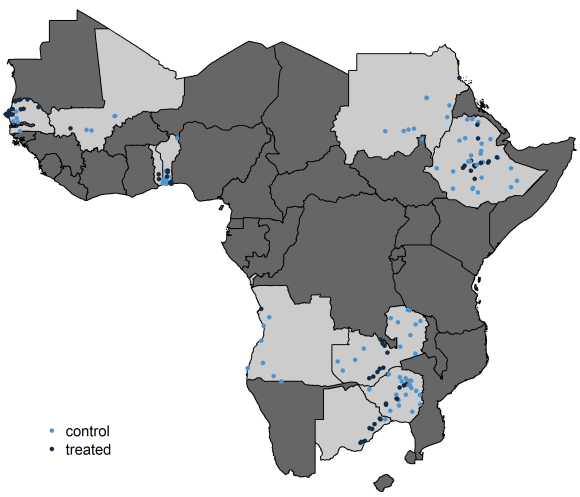


Figure 7: Town Locations in the Estimation Sample



Notes: This figure depicts the countries in our analysis (brighter gray) and for each country the town in the treatment and control group.

5 Results

5.1 Main effects

We estimate the effect of Internet availability on local economic growth. We are particularly interested in the effect of early Internet availability brought by the 'first-generation' SMCs. We estimate a linear model on a balanced panel by difference-in-differences.

In our main specification, we measure economic activity by the logarithm of the sum of light intensities. Table 2 shows the main results. In the first column, the raw specification, nodal cities and the landing point are still included. We then step-wise eliminate nodal cities until we reach our preferred specification where the remaining towns are comparable. In the second column, we remove the city of the landing point and the national capital. In the third column, we also remove regional capitals. In the fourth column, we remove cities of more than 100.000 inhabitants. Finally, we add as a control variable the share of the towns' area with GSM mobile network coverage (column 5) and its population (column 6).

We find a economically and statistically positive effect of basic Internet on local economic growth. In our preferred specification (column 6), towns which were connected to the Internet in the year of SMC arrival are 8% brighter than towns without Internet availability. For an approximation of the implied economic effect, we perform a back-of-the-envelope calculation using the GDP-luminosity elasticity of $\epsilon_{GPD,light} = 0.284$ from Henderson et al. (2012). The calculation translates the increase in light intensity of 8% into about 2 percentage points higher GDP growth.

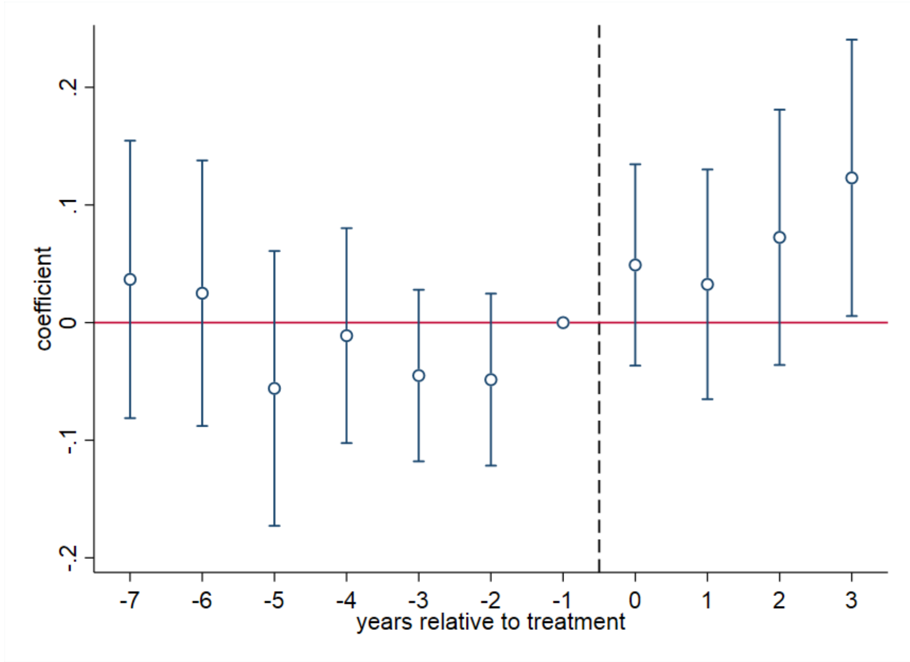
Table 2: The Effect of Internet on Economic Growth of Cities

Dep. var.: light intensity	(1)	(2)	(3)	(4)	(5)	(6)
post × treated	0.0462	0.0532*	0.0591*	0.0803**	0.0875**	0.0831**
	(0.0286)	(0.0302)	(0.0328)	(0.0355)	(0.0360)	(0.0326)
GSM coverage					0.0502	0.0476
					(0.0356)	(0.0350)
Population						0.329*
						(0.197)
Town FE	✓	✓	✓	✓	✓	✓
Country × Year FE	✓	✓	✓	✓	✓	✓
Sample restrictions						
Capitals & landing points		✓	✓	✓	✓	✓
Regional capitals			✓	✓	✓	✓
Population >100k				✓	✓	✓
# Observations	3,190	3,069	2,563	2,277	2,277	2,277
# Countries	10	10	10	10	10	10
# Towns	290	279	233	207	207	207
Share treated	.493	.473	.468	.42	.42	.42
Adj. R^2	0.971	0.961	0.958	0.937	0.937	0.937

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area. GSM mobile phone coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100k inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

For our preferred specification (column 6), Figure 8 presents event study coefficients. Before the SMC connection, the point estimates are close to zero and insignificant. In the years after SMC arrival, the point estimates are between .05 and .1 and have slightly the tendency to increase to nearly .15 in year three after the connection, when the effect gets significant at the 5%-level. We cannot estimate effect for a longer time period as second-generation SMCs arrive in the years after initial SMC arrival to upgrade Internet capacity and speeds. From this dynamic perspective, there is no evidence for a potential fading-out of the effect.

Figure 8: Event Study Coefficients



Notes: Coefficients for event study specification of model (6) from Table 2. Robust standard errors clustered by town for 95% confidence interval reported as bars.

5.2 Mechanism

Next, we investigate the intensive and extensive margin. Therefore, we take the mean luminosity (Table A.1, column 3 & 4) and the sum of lit pixels (Table A.1, column 5 & 6) as alternative outcome measures to the sum of luminosity (Table A.1, column 1 & 2). All outcomes are logged. In odd columns, we estimate specifications as in column 5 of Table 2. In even columns, we add population as a control. We observe that towns having Internet available are getting brighter (columns 3 and 4) and also that they increase in their size (columns 5 and 6). Adding population as a control, the main effect remains robust. The coefficient of the population control is specifically significant for the extensive margin (column 6).[24] We can state the effect is not driven by migration as we see the strongest effect in the intensive margin. Furthermore, the population is insignificant there. Hence, we are assured that we mainly estimate an effect of productivity increase.

[24] Taking an additional radius around the built-up area of the towns, we assume that lit pixels are inside the towns' border. Therefore, blurring of night-time lights on the satellite image are not an issue anymore and 'size' measures the actual growth of the towns. Hence, the growth can mainly be reduced to an increase in light intensity.

Table 3: Population, Intensive, and Extensive Margin

	combined		intensive		extensive	
	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.:	sum	sum	mean	mean	# pixels	# pixels
post × treated	0.0875**	0.0831**	0.0667***	0.0659***	0.0576**	0.0530**
	(0.0360)	(0.0326)	(0.0237)	(0.0235)	(0.0291)	(0.0249)
GSM coverage	0.0502	0.0476	0.0461*	0.0457*	0.0328	0.0301
	(0.0356)	(0.0350)	(0.0249)	(0.0250)	(0.0277)	(0.0269)
Population		0.329*		0.0555		0.344**
		(0.197)		(0.115)		(0.157)
Town FE	✓	✓	✓	✓	✓	✓
Country × Year FE	✓	✓	✓	✓	✓	✓
Sample restrictions						
Capitals & landing points	✓	✓	✓	✓	✓	✓
Regional capitals	✓	✓	✓	✓	✓	✓
Population >100k	✓	✓	✓	✓	✓	✓
Observations	2,277	2,277	2,277	2,277	2,277	2,277
# Countries	10	10	10	10	10	10
# Towns	207	207	207	207	207	207
Share treated	.42	.42	.42	.42	.42	.42
Adj. R^2	0.937	0.937	0.941	0.941	0.917	0.917

Notes: Light intensity in the combined models is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area. The measure for the intensive margin is logarithmic mean light intensity and for the extensive margin logarithmic sum of pixels. GSM mobile phone coverage is calculated as the percentage share of town area covered with signal. Population is measured as inhabitants per square kilometer within town area from GPW. All specifications are estimated on a sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100k inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.3 Heterogeneity and Robustness

First, we analyze different connection years. In Table 4, we first remove later connection years. Here, the estimate increases slightly. Moreover, investigating only countries that where connected in 2001, the year where several countries were connected by one bigger SMC, shows a strong positive effect of Internet availability on local economic growth. Also, when dropping countries with an early connection year step-wise the effect size remains but drops slightly in the level of significance.

Table 4: Heterogeneity by SMC Connection Year

Dep. var.: light intensity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
post × treated	0.100**	0.142***	0.129**	0.115**	0.123**	0.196**	0.110**	0.0876*	0.0876*	0.0907*
	(0.0385)	(0.0466)	(0.0509)	(0.0548)	(0.0591)	(0.0776)	(0.0430)	(0.0503)	(0.0503)	(0.0529)
City FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country × year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
City restriction	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Connected before	2008	2006	2005	2004	2003	–	–	–	–	–
Connected in	–	–	–	–	–	2001	2001	2001	2001	2001
Connected after	–	–	–	–	–	–	2000	2001	2002	2003
Observations	1,548	1,212	1,092	636	504	216	1,260	1,044	1,044	912
Adj. R-squared	0.974	0.977	0.978	0.962	0.963	0.960	0.972	0.972	0.972	0.976

Notes: Light intensity in the combined models is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area. GSM mobile phone coverage is calculated as the percentage share of town area covered with signal. Population is measured as inhabitants per square kilometer within town area from GPW. Sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100k inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We define the treatment for a city as having an AP within 10 km. In Table 5, we vary this distance. The estimate remains around .1.

Table 5: Distance to Access Point

Dep. var.: light intensity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
post × treated	0.100**	0.0766	0.116**	0.113**	0.110**	0.0922**	0.112***	0.111***	0.0876**	0.102***	0.101***
	(0.0385)	(0.0606)	(0.0462)	(0.0435)	(0.0430)	(0.0403)	(0.0361)	(0.0359)	(0.0376)	(0.0354)	(0.0359)
Distance to AP	10 km	0 km	2 km	4 km	6 km	8 km	12 km	14 km	16 km	18 km	20 km
City FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
City restriction	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1,548	768	1,128	1,320	1,404	1,488	1,620	1,644	1,572	1,620	1,644
Adj. R-squared	0.974	0.976	0.976	0.973	0.973	0.974	0.973	0.974	0.974	0.974	0.974

Notes: Light intensity in the combined models is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area. GSM mobile phone coverage is calculated as the percentage share of town area covered with signal. Population is measured as inhabitants per square kilometer within town area from GPW. Sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100k inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

6 Conclusion

Digital infrastructure is a key precondition for locations to harvest digital dividends from Internet connectivity. We investigate if the availability of even basic Internet speeds fosters economic development in developing-countries. In particular, we study the arrival of the first sub-marine Internet cables (SMCs) in 10 Sub-Saharan African (SSA) countries in the 2000s. To learn about the causal effect of basic Internet on local economic growth, we compare economic activity—measured by night-time light satellite data—of towns connected to the national Internet backbone at the time of SMC arrival to a control group of similar towns not (yet) connected to the national digital infrastructure.

We find that the connection of towns to the world wide web, on average, leads to an increase in light intensity of about 8%, relative to similar towns not (yet) connected. This translates into 2 percentage points higher growth in terms of GDP. Moreover, we differentiate growth in more pixels, where towns increase in their area (extensive margin), and in a higher average of the light intensity, which is associated with a higher density of economic activity (intensive margin). We find that towns with Internet availability due to access to digital infrastructure typically grow on both margins, i.e. become brighter and geographically bigger. Further, our results suggest that this growth is partly driven by growing populations in connected towns.

References

Akerman, A., I. Gaarder, and M. Mogstad (2015, nov). The skill complementarity of broadband internet. *Quarterly Journal of Economics 130*(4), 1781–1824.

Albouy, D. (2016). What are cities worth? land rents, local productivity, and the total value of amenities. *The Review of Economics and Statistics 98*(3), 477–487.

Allen, T. and C. Arkolakis (2019). The Welfare Effects of Transportation Infrastructure Improvements. *National Bureau of Economic Research*.

Banerjee, A., E. Duflo, and N. Qian (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics 145*(January), 102442.

Bloomberg (2020). Faster internet coming to africa with facebook's $1 billion cable.

Bluhm, R. and M. Krause (2018). Top lights - bright cities and their contribution to economic development. *CESifo Working Paper No. 7411*.

Bruederle, A. and R. Hodler (2018). Nighttime lights as a proxy for human development at the local level. *PLoS ONE 13*(9), 1–22.

Chavula, J., N. Feamster, A. Bagula, and H. Suleman (2015). Quantifying the effects of circuitous routes on the latency of intra-africa internet traffic: A study of research and education networks. In A. Nungu, B. Pehrson, and J. Sansa-Otim (Eds.), *e-Infrastructure and e-Services for Developing Countries*, Cham, pp. 64–73. Springer International Publishing.

Chen, X. and W. D. Nordhaus (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences of the United States of America 108*(21), 8589–8594.

Clark, T. N., R. Lloyd, K. K. Wong, and P. Jain (2002). Amenities Drive Urban Growth. *Journal of Urban Affairs 24*(5), 493–515.

Czernich, N., O. Falck, T. Kretschmer, and L. Woessmann (2011). Broadband Infrastructure and Economic Growth. *Economic Journal 121*(552), 505–532.

Deller, S. C., V. Lledo, and D. W. Marcouiller (2008). Modeling regional economic growth with a focus on amenities. *Review of Urban and Regional Development Studies 20*(1), 1–21.

Hamilton Research (2020). Africa bandwidth maps.

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

Henderson, V., A. Storeygard, and D. N. Weil (2011). A bright idea for measuring economic growth. *American Economic Review 101*(3), 194–199.

Hjort, J. and J. Poulsen (2019). The arrival of fast internet and employment in Africa. *American Economic Review 109*(3), 1032–1079.

Hornung, E. (2015). Railroads and growth in Prussia. *Journal of the European Economic Association 13*(4), 699–736.

Invesco (2019). China embarks on digital silk road.

Kende, M. and K. Rose (2015). Promoting Local Content Hosting to Develop the Internet Ecosystem. Technical report, Internet Society.

Li, X., Y. Zhou, M. Zhao, and X. Zhao (2020). A harmonized global nighttime light dataset 1992–2018. *Scientific Data 7*(1), 1–9.

Storeygard, A. (2016). Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa. *Review of Economic Studies 83*(3), 1263–1295.

World Bank (2016). World Development Report 2016: Digital Dividends. Technical report.

A Appendix

Table A.1: Population, Intensive, and Extensive Margin

	population		GSM coverage	
	(1)	(2)	(3)	(4)
post × treated	0.0123	0.0134	-0.145***	-0.147***
	(0.0188)	(0.0200)	(0.0402)	(0.0394)
GSM coverage		0.00770		
		(0.0125)		
Population				0.152
				(0.199)
Town FE	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓
Sample restrictions				
Capitals & landing points	✓	✓	✓	✓
Regional capitals	✓	✓	✓	✓
Population >100k	✓	✓	✓	✓
Observations	2,277	2,277	2,277	2,277
# Countries	10	10	10	10
# Towns	207	207	207	207
Share treated	.42	.42	.42	.42
Adj. R^2	0.999	0.999	0.815	0.815

Notes: Light intensity in the combined models is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area. The measure for the intensive margin is logarithmic mean light intensity and for the extensive margin logarithmic sum of pixels. GSM mobile phone coverage is calculated as the percentage share of town area covered with signal. Population is measured as inhabitants per square kilometer within town area from GPW. Sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100k inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure A.1: Fiber Nodes and Their Year of Construction

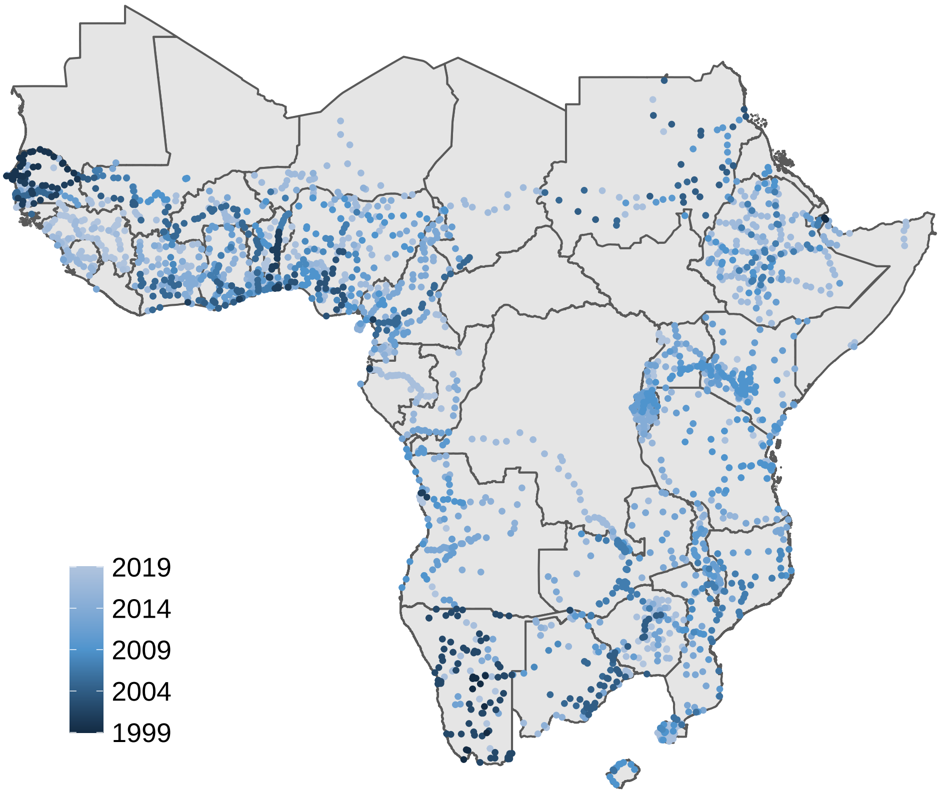


Figure A.2: Time Trends by Sample