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

Moritz Goldbeck† Valentin Lindlacher‡

January 30, 2023

Abstract

We study whether internet availability at basic speeds fosters local economic growth in remote areas of developing countries by analyzing remote towns in about 10 Sub-Saharan African countries. We measure local economic growth of each town by tracking its nighttime light emissions. In a difference-in-differences setting, we exploit plausibly exogenous nationwide variation in internet availability induced by submarine cable arrivals and use the rollout of within-country inter-regional fiber cables to design comparable treatment and control groups. We find that basic Internet availability induces economic growth. Compared to a control group of similar but later connected towns, connected towns experience 10 percent higher light intensity which translates to about 3 percentage points higher annual economic growth in the years after the arrival of submarine cables. Internet availability is accompanied by a shift from agriculture to manufacturing in regional employment. Further analyses suggest this result is mainly driven by per capita productivity growth and not by migration into connected towns. The effect is stronger in towns closer to ports, i.e., with higher market access, indicating that trade is an important mechanism.

Keywords: Internet, regional development, nighttime light, Sub-Saharan Africa
JEL-Codes: O33, O18, R11

---

*We thank Vojtech Bartos, Mathias Bühler, Oliver Falck, Thomas Fackler, Jonas Hjort, Anna Kerkhof, Markus Ludwig, Niklas Potrafke, Helmut Rainer, Claudia Steinwender, 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", the 10th European Meeting of the Urban Economics Association, the EBE CRC Summer School (Rationality and Competition), the 9th Meeting of the Society for the Study of Economic Inequality (ECINEQ), the UEA PhD Student Workshop, the ifo Dresden Workshop on Regional Economics, the Verein für Socialpolitik (VfS), the Doctoral Workshop L'Association Française de Recherche en Economie Numerique (AFREN), and the CESifo Area Conference on the Economics of Digitization, especially, to the discussants Kathrin Wernsdorf, Marcelo Sant'Anna, and Tobias Korn. Finally, we thank the participants of seminars at the ifo Institute, Economic Research South Africa, University of Munich, and TU Dresden. All errors are our own.

†University of Munich & ifo Institute Munich

‡Corresponding author. Helmholtzstr. 10, 01069 Dresden (Germany). TU Dresden, ifo Institute Munich & CESifo, valentin.lindlacher@tu-dresden.de.

1 Introduction

In the last decades, the provision of digital infrastructure in many countries enabled widespread access and adoption of 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 fosters regional economic growth in the developing world as well (World Bank, 2016). For example, in Sub-Saharan Africa (SSA), where impulses for economic growth are required to fight poverty and deprivation, governments, public-private partnerships, and companies alike invest large amounts of money in Internet infrastructure projects. However, provision is complex and costly due to lacking legacy infrastructure like fixed-line telephony networks. Until 2020, SSA countries invested more than 28 billion US-Dollar into their national Internet backbone (Hamilton Research, 2020).[1] Despite these enormous investments in Africa's digital infrastructure, a growth effect of Internet in SSA is not assured. Low population density apart from the mega-cities, missing hardware, financial constraints, and a lower willingness to pay lead to low 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. Given the large investment requirements and unclear economic benefits, it is crucial to understand how Internet availability affects regional economic development in SSA, especially in rural areas where provision is particularly costly.

In this paper, we examine if there is a causal effect of Internet availability on local economic growth in remote towns of SSA even at basic speeds. We focus on the initial introduction of Internet in SSA through the 'first generation' of Internet-enabled submarine cables (SMCs) starting in the early 2000s. We investigate this effect at the town level to analyze whether potential individual-level effects, found by Hjort and Poulsen (2019) for later high-speed Internet availability, materialize on a more aggregate level. Our focus on remote areas allows us to explore whether Internet availability has an effect beyond political and economic centers and thus whether it can affect countries' regional development. We capture the evolution of 220 remote towns in 10 SSA countries provided with (international) Internet connection in the early 2000s and a pre-existing national backbone. We measure growth of towns by spatial expansion (extensive margin) and density of economic activity (intensive margin) and interpret these components as pointing more towards population or productivity growth, respectively. Furthermore, we investigate changes in the industry composition as a potential mechanism.

We tap two main data sources. First, we measure local economic growth, the key outcome of interest, using nighttime light (NTL) intensity captured by satellites, a well-established proxy introduced by Henderson et al. (2011) at the country level and validated by Storeygard (2016) on the city level for SSA. To get the local town-level measure, we assign NTLs to individual agglomerations by linking lit pixels to built-up

[1] Facebook announced an effort to build a new Internet-enabled submarine cable (SMC) to Africa for one billion US-Dollar in 2020 (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).

areas of SSA cities and towns from *Africapolis*. Second, we use data on the roll-out of national backbones to measure internet availability of individual cities and towns from Hamilton Research (2020). The data comprises the geo-location of all Internet access points in SSA. Because data only starts in 2009, hand-collect actual construction years via an extensive review of national backbone deployment projects for each SSA country. This enables us to study the 'first generation' of SMC arrivals, which introduced the Internet in SSA for the very first time on a noticeable scale.

To identify the causal effect of Internet availability on local economic growth, we exploit quasi-random variation in the timing of country-wide Internet access induced by the arrival of the 'first generation' of SMCs in SSA in the early 2000s. This approach was established by Hjort and Poulsen (2019), who exploit an Internet speed upgrade induced by SMCs with higher capacities between 2009 and 2012. We focus on remote towns that are located between nodal cities (political and economic centers). These towns are relatively small and are primarily connected due to their fortunate location. In a difference-in-differences setting, we define treatment and control group towns using the roll-out of national backbones, which make Internet available through local access points. We assign treatment status to towns that were connected to the national backbone when the Internet became available country-wide, while the control group consists of similarly sized towns getting Internet connection through an access point only some years later. In a two-way fixed effects (TWFE) model with town and country-year fixed effects we then compare the growth of towns with Internet access at the time when broadband Internet at basic speeds becomes available country-wide for the first time to a control group of similar towns getting access only later. Our key identifying assumption is that treatment and control group towns would have evolved similarly in the absence of treatment. Although this assumption cannot be tested directly, we perform a dynamic event-study specification of our model to show that there are no differences in pre-treatment trends of economic activity between treatment and control group towns.

We find that connection to the Internet through an access point leads on average to a 10 percent increase in NTL intensity of SSA towns in the first four years after country-wide connection compared to a control group of similar towns not connected through an access point at that time. Applying the established light-to-GDP elasticity from Henderson et al. (2012), this translates into about 3 percentage points higher economic growth. We then differentiate between growth in the average brightness of lit pixels, which is associated with a higher productivity or density in the towns (intensive margin), and growth in the number of lit pixels, indicating a spatial expansion of towns (extensive margin). We find both increased brightness and size of towns with Internet access after connection. Together with our the fact that we do not find effects on population growth this points towards economic development rather than a spatial redistribution of economic activity. Further, we find this effect accompanied by a shift in regional industry shares. In regions with connected towns, manufacturing employment shares increase by around 2 percentage points in comparison to regions getting connected later. Manufacturing employment seems to be growing mainly at the expense of agricultural employment in these regions, although not statistically significantly. This provides suggestive evidence that the increase in economic activity is at least partly a result of a changing regional industry structure induced by Internet availability.

To ensure that our results are driven by Internet availability, we control for the roll-out of mobile GSM coverage.[2] Additionally, we perform placebo tests controlling for access to other potentially confounding infrastructure, such as roads, railroads, and the electricity grid. Further, we perform placebo exercises on the timing of countrywide Internet connections. We use 1,000 simulations with randomly selected country-connection years prior to the countries' actual connection year to show that the effect is only present for the actual connection years. We test robustness of our results under alternative assumptions about the variance-covariance matrix, including changing the level of fixed effects and standard errors, adding linear time trends at the town level, and applying novel event-study estimators. Furthermore, we can show that the results cannot be explained by [[variable used]], suggesting that ethnic favoritism does not drive results. Finally, we extend our sample by relaxing assumptions to assure the external validity of our results and investigate heterogeneous effects by focusing on coastal countries only.

With the notable exception of Hjort and Poulsen (2019), who find sizable positive individual-level effects of an Internet speed upgrade on individual-level employment probability in SSA between 2009 and 2012, causal evidence on the economic impact of Internet availability in developing countries is scarce. This is the first study investigating the causal effect of the introduction of (fixed-line) Internet availability at basic speeds on local economic growth in developing countries. Further, we study early Internet effects in a rural setting with no pre-existing fixed-line telephony network, low penetration rates, and labor-intensive local economies.

We contribute to two main strands of the literature. First, we add to the broad literature assessing the impact of infrastructure on economic outcomes. Our study is the only one investigating the overall impact of (fixed-line) Internet availability on local economic growth for developing countries when Internet becomes available for the first time.[3] Most closely related to our work is Hjort and Poulsen (2019), who study the employment effects of broadband Internet on an individual level when broadband capacity increases. They find a skill-biased and net positive employment effect for an Internet speed upgrade in SSA around 2010. 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 and for an overall measure of economic activity as well and even at basic speeds.[4]

[2]At the time we study, all countries only had basic mobile coverage enabling calls and SMS but not surfing the Internet. Specifically, 3G coverage, and therefore mobile Internet, was not available.

[3]Hjort and Tian (2021) give a comprehensive overview of the effects of Internet connectivity in developing countries, dividing this literature into supply-side and demand-side mechanisms and overall impact of connectivity.

[4]There is also a large body of related literature on the effect of nondigital infrastructure on economic outcomes in developing countries. Assessed infrastructure includes transportation infrastructure (see e.g., Storeygard, 2016; Ghani et al., 2016; Banerjee et al., 2020; Faber, 2014; Donaldson, 2018), electrification (see e.g., Dinkelman, 2011; Grogan and Sadanand, 2013; Rud, 2012).

For developed countries, the effect of digital infrastructure and especially (broadband) Internet has been assessed widely. Czernich et al. (2011) identify an effect of broadband infrastructure on annual per capita growth for OECD countries. For the US, Kolko (2012) finds a positive relationship between broadband expansion and local economic growth, i.e., growth in population, employment, average wage, and employment rate.[5] While Internet speeds and the timing are very comparable to our setting, adoption rates were a lot higher in developed countries, mostly because pre-existing fixed-line telephony infrastructure made household DSL connections a lot easier. We add to this literature by showing that Internet availability benefits regional economic development also with low adoption rates. This implies that, if a few adopters generate such an effect that it is measurable at the aggregate, the Internet must have great spillover effects.

Related to Internet are mobile phones. Jensen (2007) shows that the adoption of mobile phones led to a reduction in price dispersion and an increased consumer and producer welfare. In a related paper, Aker and Mbiti (2010) study how the introduction of mobile phones between 2001 and 2006 affected grain prices in Niger. These papers emphasize the importance of rolling out mobile network infrastructure for improving economic efficiency of markets. More generally, mobile communication offers a major opportunity to advance economic growth in developing countries, for example by providing information about prices, improving the management of supplies, increasing the productive efficiency of firms, reducing transportation costs, and other means (Aker and Mbiti, 2010). Fixed-line Internet, as we analyse, might work through the same channels but accesses international information sources.

Second, our work contributes to the literature on urban and regional development. Starting with Nunn and Puga (2012) who showed that in Africa less fortunate geography has a positive impact on today's economy and Henderson et al. (2012) who indicated that the hinterland grows faster than coastal areas and that primate cities do not grow faster than their hinterland, a strand of literature focuses on the catch-up from secondary to primate cities, with no conclusive results. While many papers show that secondary cities are meaningful to reduce poverty (see e.g., Christiaensen and Todo, 2014; Christiaensen and Kanbur, 2017; Fetzer et al., 2016), Bluhm and Krause (2018) show with an adjustment for top coding in NTLs that primate cities remain the economic centers. We contribute by focusing on even smaller towns and showing that even there economic development is happening.

In section 2, we provide a brief overview of early Internet in SSA. section 3 lays out the empirical strategy and in section 4 the data is described. Results are presented in section 5. section 6 discusses our results in comparison with related research. section 7 concludes.

[5] Moreover, labor market effects (see e.g., Atasoy, 2013; Czernich, 2014; Akerman et al., 2015) and effects on firm productivity (see e.g., Akerman et al., 2015; Grimes et al., 2012; Colombo et al., 2013) both with mixed results were investigated.

2 Background

There are three major components of Internet infrastructure determining the availability and bandwidth of the Internet in a given location. First, international fiber-optic submarine cables (SMCs) connect SSA countries to the global Internet backbone.[6] Second, within-country inter-regional fiber-optic cables form the national backbone. A precondition for internet availability in a location is an access point to the national backbone close by. Finally, individual users in a location are reached via the 'last mile' infrastructure.

2.1 International Backbone: Submarine 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-continental (Kende and Rose, 2015; Chavula et al., 2015). Before the first SMCs landed on SSA shores, the only way to connect to the Internet on the continent was via satellite.[7] While being largely unconstrained by geography and local infrastructure, satellite connection is costly and allows only for very narrow bandwidths. With SMCs, a joint effort of governments, private investors, and/or multinational organizations, an 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 1999 and the early 2000s. These 'first-generation' cables had the capacity to provide Internet at basic speeds.[8] The biggest of them was SAT-3 and started operating in 2001. It featured landing points on the shores of nine SSA countries on the western coast of Africa.[9] These landing points, typically one per country, constitute the starting point for the respective national backbones (cf. subsection 2.2). Until the late 2000s, most SSA countries were connected to the Internet via these 'first-generation' SMCs.[10]

Landlocked countries are only indirectly connected through SMCs. They rely on their neighboring countries which connect them through a national backbone. The rollout of these inter-regional fiber-optic cables is explained next.

[6] We define Sub-Saharan Africa as the mainland of the African continent without the Northern African countries, Algeria, Egypt, Libya, Morocco, Tunisia, and Western Sahara. Moreover, we exclude South Africa as it is economically more developed and therefore less comparable to the other SSA countries.

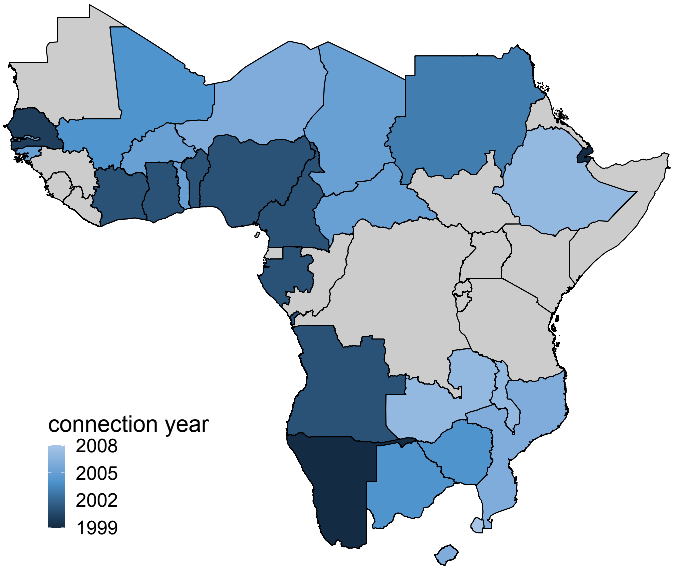
[7] Single-channel and co-axial SMCs for telegraphy and telephony already existed before. The first telegraphy cable ('East coast' cable) started operating as early as 1879.

[8] Hjort and Poulsen (2019) state that SSA users had on average 430 Kbps before the 'second generation' of SMCs arrived. In Benin, for instance, ADSL connections with up to 2 Mbps were possible before the upgrade SMC arrived (Agyeman, 2007)

[9] These countries are: Angola, Benin, Cameroon, Côte d'Ivoire, Gabon, Ghana, Nigeria, Senegal, and South Africa. It started in Sesimbra, Portugal, and Chipiona, Spain, and also passed the Canary Islands in Alta Vista.

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

Figure 1: Internet connection years



Notes: The figure shows SSA with all countries getting an Internet connection before 2008. The color gradient depicts the connection year (darker blue colors indicate earlier initial SMC connection years). Gray indicates countries not connected to the Internet until 2008.

2.2 National Backbone: Inter-Regional Cables

After being routed through an SMC, Internet traffic travels through the national backbone. The national backbone infrastructure consists of inter-regional fiber-optic cables. Therefore, as soon as a new SMC arrives at a landing point of a SSA country, Internet becomes available countrywide 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 two decades, national backbones were continuously improved and expanded in parallel with the installation of SMCs.[11] This backbone 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).[12] This often led to a backbone evolution where the national capital (often a coastal city and located closely to the landing point) was connected first. Then, the backbone spread out to the next largest or politically important cities. Due to their role as nodes in the national backbone networks, we call these cities 'nodal cities'.

[11] Many of these cables were constructed 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. There are no network operators owning backbones in more than one country.

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

Inter-regional cables are almost always constructed along pre-existing infrastructure, e.g., 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 got Internet access as well due to their fortunate location between two nodal cities. Our empirical strategy (cf. section 3) focuses on these incidentally connected towns which get an Internet connection because of their location next to an inter-regional cable.

2.3 Local Transmission: 'Last Mile' Infrastructure

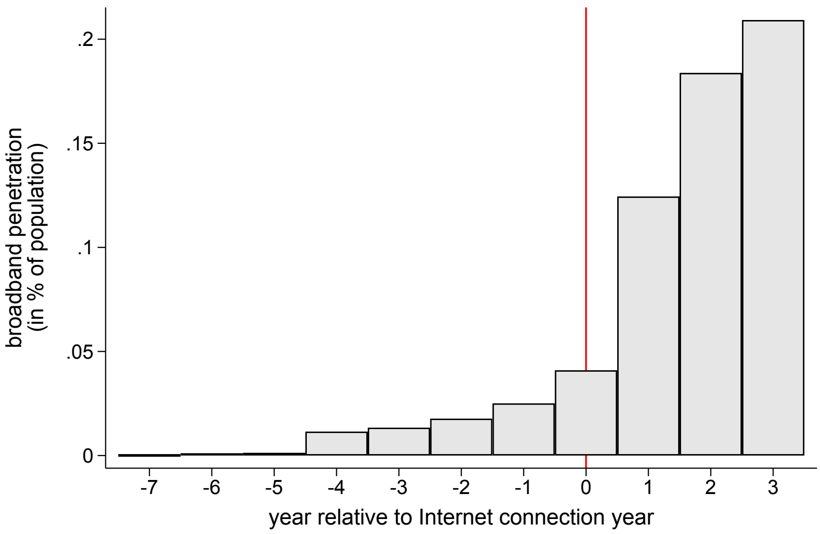
Internet traffic transported by inter-regional cables is accessed at access points. There are several technologies transmitting Internet traffic from these access points to the user. These 'last mile' transmission technologies include fiber cables (FTTH/B), copper cables, and wireless transmission using cellular towers (e.g., mobile or WiMax). Unlike in many developed countries which rely heavily on transmission to the end user via pre-existing telephony cable infrastructure, in SSA countries households are seldomly connected through copper or even fiber cables. Instead, traffic data is exchanged wirelessly. For this technology, no local cable network connecting each user's exact location (firm, household) is needed. 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 access points even in on-route towns, which are typically much smaller than nodal cities.

Figure 2 shows how the usage increases in countries that were served by a 'first generation' SMC. The change in absolute numbers is rather low (0.2 percentage points). However, it is notable that the increase starts when the Internet becomes available. Compared to the year before the Internet was accessed through an SMC, but only via a satellite connection, broadband penetration increased by eight times in only four years. And although broadband penetration is low among the population, anecdotally most users access the Internet through cybercafés. Thus, the share of Internet users might be a lot higher than the penetration in Figure 2 suggests. Moreover, the broadband penetration in firms might be a lot higher. Although data on broadband adoption of SSA firms is not widely available for that time, the *World Bank Enterprise Survey* shows even before the 'second generation' of SMCs landed on SSA shores that 52 percent of all firms used email for communication and 23 percent had an own website.[13]

Countries that were connected later had in theory more time for the backbone network rollout. In Figure 3, we plot the percentage of agglomerations with an access points when the countrywide internet connection was established (from all access points being constructed until today) against the country's connection year. A positive correlation is shown. Regressing these two variable gives an estimator of 0.027, which is statistically significant at the 10 percent level. So, an delay of about four years would lead to 10 percentage point higher share of connected towns.

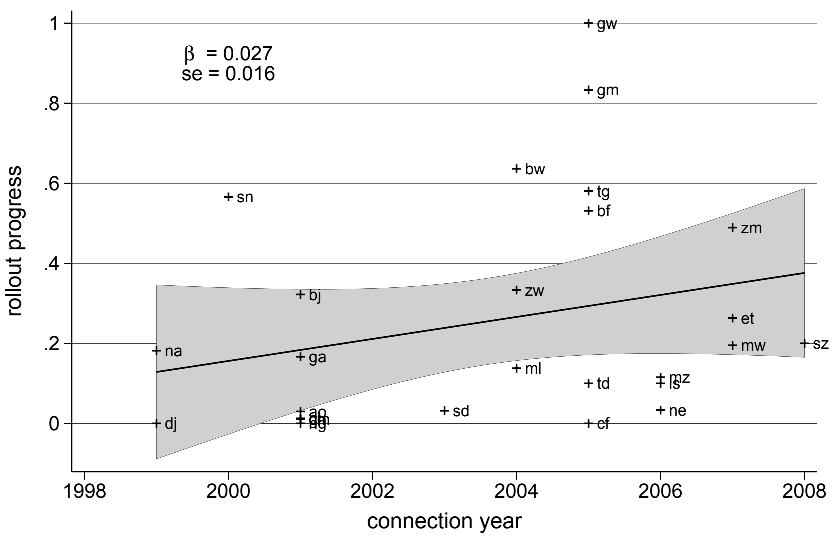
[13] `https://www.enterprisesurveys.org`, accessed on XYZ.

Figure 2: Internet connection and adoption



Notes: Adoption rates are calculated relative to the establishing year of the Internet connection in each country and then aggregated taking the weighted mean. Weights are population size in 2000.

Figure 3: Internet connection and rollout



Notes: Restricted to countries being connected to the Internet before 2009.

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 as Internet access is not randomly assigned and likely driven by commercial interest and/or political and administrative planning.

To address these endogeneity concerns, we leverage a distinct feature of Internet infrastructure evolution in SSA countries. First, we use plausibly exogenous time variation in connections to submarine cables (SMCs), which determine Internet availability countrywide for coastal countries, to investigate the effect of Internet availability. Following Hjort and Poulsen (2019), we argue that the exact timing of SMC arrival is essentially random.[14] The arrival is exogenous, first, because each SMC typically connects many countries. Therefore, coordination difficulties among consortium members might delay the construction.[15] Second, the connection years are highly uncertain due to unforeseen delays in construction. For example, the cable EASSy was delayed by five years due to coordination difficulties among consortium members (Poppe, 2009). Moreover, a country's geographical location within SSA can influence the connection year. First, Eastern and Western SSA countries get independently their respective SMCs. Second, landlocked countries get their connection through the national backbone of their neighboring countries and rely therefore on the construction speed of another country's national backbone. This construction speed again is exogenous for the respective landlocked country.

We estimate the effect of early Internet at basic speeds and exploit the arrival of the 'first generation' of SMCs. When the next generation with higher capacities arrives, starting in 2009, countries immediately get a speed upgrade. Therefore, we estimate on a sample containing only years for which countries did not receive a speed upgrade yet. Due to the staggered timing of the 'second generation' of SMC, this sample is unbalanced. To estimate on a balanced sample, we restrict the estimation to three post-treatment years.

In a difference-in-differences (DiD) design, we compare towns that already have access to the national backbone when the Internet becomes available countrywide to a control group of similar towns getting an access point in later years (first difference) before and after the country's Internet connection (second difference). We exclude nodal cities, i.e., cities close to an access point that are endogenously connected (cf. subsection 2.2): the landing point, the capital, regional capitals, and economic centers (cities with a population of more than 50,000 inhabitants). For robustness, we vary the threshold of 50,000 inhabitants as the definition of an economic center.

All towns in our analysis get connected eventually, mainly because of their favorable location between nodal cities. Hence, towns that are still waiting for an access point today will only serve for robustness. As an

[14] This exogeneity was also exploited by Cariolle (2021).

[15] Consortium investors usually are public and private telecom operators and neighboring and foreign investors (Jensen, 2006).

additional robustness check, we vary the last possible year of connection for the control group, such that treated towns are not compared with very late connected towns. Moreover, towns being connected after the Internet became available countrywide but in the estimation period (up to three years after the treatment year) are excluded as they would contaminate the control group. They do not get the full treatment and would thereby confound our analysis. In a robustness check, we define them as treated with the access point construction year as treatment year.

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

$$y_{c(i)t} = \beta_0 + \beta_1 \left(connection_{ct} \times access_{c(i)} \right) + \beta_2 GSM_{c(i)t} + \mathbf{X}'_{c(i)} \phi_t + \alpha_{c(i)} + \alpha_{ct} + \varepsilon_{c(i)t}, \tag{1}$$

where $y_{c(i)t}$ is economic growth of town i in country c in calendar year t as proxied by nighttime light (NTL) intensity (cf. section 4), where the logarithm is used to estimate changes in the growth rate instead of changes in levels. The dummy variable $connection_{ct}$ indicates if country c has a countrywide Internet connection in calendar year t. The variable $access_{c(i)}$ is one if town i in country c is located within 10 kilometers distance to an access point that was established in the year when the Internet became available countrywide or before. Contrary, the indicator is zero if town i in country c is located within 10 kilometers to an access point that was established in the years afterwards. Thus, the interaction term $connection_{ct} \times access_{c(i)}$ 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 growth of early versus later connected towns. For robustness, we vary the distance including smaller and higher values than 10 kilometers. The control variable $GSM_{c(i)t}$ contains time-varying mobile coverage as the share of a town's area covered by GSM technology. The vector $\mathbf{X}'_{c(i)}$ contains time-invariant controls, such as the distance to the capital, to the submarine cable's landing point, to the country's border, to the coastline, to the next port, to the next river, to the (rail)road network, to the electricity grid, and the town's altitude and the ruggedness of the area. Accordingly, ϕ_t are time-varying coefficients on these controls. We include two types of fixed effects into the model. Time-constant differences across towns are captured by town fixed effects $\alpha_{c(i)}$. Differences across calendar years common to all towns within a country are absorbed by country-year fixed effects α_{ct}. Note that this allows for country-specific time trends, especially country-specific growth rates, and variations in satellite sensor quality over years. Choosing country-year fixed effects implies that we stack individual country estimations. 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 group towns would have evolved similarly in the absence of the 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 the control group. Therefore, we analyze the dynamic impact of Internet availability on local economic activity using an event-study design:

$$y_{c(i)t} = \sum_{j=\underline{T}}^{\overline{T}} \mu_{1j} \big(t_j \times access_{c(i)}\big) + \mu_2 GSM_{c(i)t} + \mathbf{X}'_{c(i)} \psi_t + \delta_{c(i)} + \delta_{ct} + e_{c(i)t}, \tag{2}$$

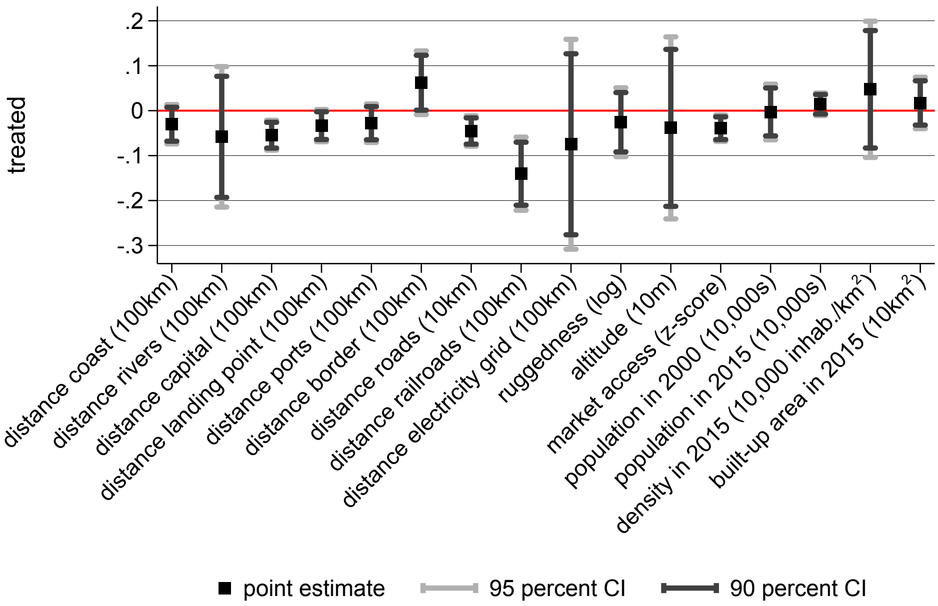
where t_j indicates the year relative to treatment year, i.e., the year when the Internet became available countrywide, starting in relative year $j = \underline{T}$ and ending with relative year $j = \overline{T}$. The treatment year is normalized to $j = 0$. We exclude $j = -1$ as the reference point. Thus, the interaction $t_j \times access_{c(i)}$ 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 μ_{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 of the pre-treatment relative-year dummies from zero, the treatment and control group follow similar trends before the treatment, supporting the parallel-trends assumption.

Balance Test The exogenous shock is at the country level. And in Figure 3 we already showed that the timing of the countrywide internet connection has an influence on countries rollout progress. As the rollout of the fiber network might not be random, we can still test whether our observable time-invariant controls correlate with the defined treatment status of the town, given country fixed effect, in the cross section. Although not being random, if the treatment status cannot be predicted from the controls, this would strengthen our identification. The access point rollout follows existing (transportation) infrastructure. Capital cities are usually close to the landing point (in coastal countries) and are aimed to be connected early. Thus, a negative correlation with distance to capital cities and roads and railroads is not surprising (Figure 4). We will handle this factor in heterogeneity exercises deeply later. Importantly, there is no statistically significant correlation with other observables. Additionally, we check whether the treatment status correlates with today's towns' education, finance, and health infrastructure, again given country fixed effects (Figure A.1). Only for colleges, we find some statistically significant correlation.

For the same variables, we can aggregate them, taking the (weighted) average, at the country level and estimating correlations between these characteristics and the country's connection date. As coastal countries have a natural advantage, we control in all estimations for an indicator which is one if the country is coastal.[16] Figure 5 shows that for most variables we find a point estimate very close to zero and lacking statistical significance. Only for the distance to the electricity grid, we find a large and statistically significant effect. However, its sign is contra-intuitive: countries with a higher average distance of their agglomerations from

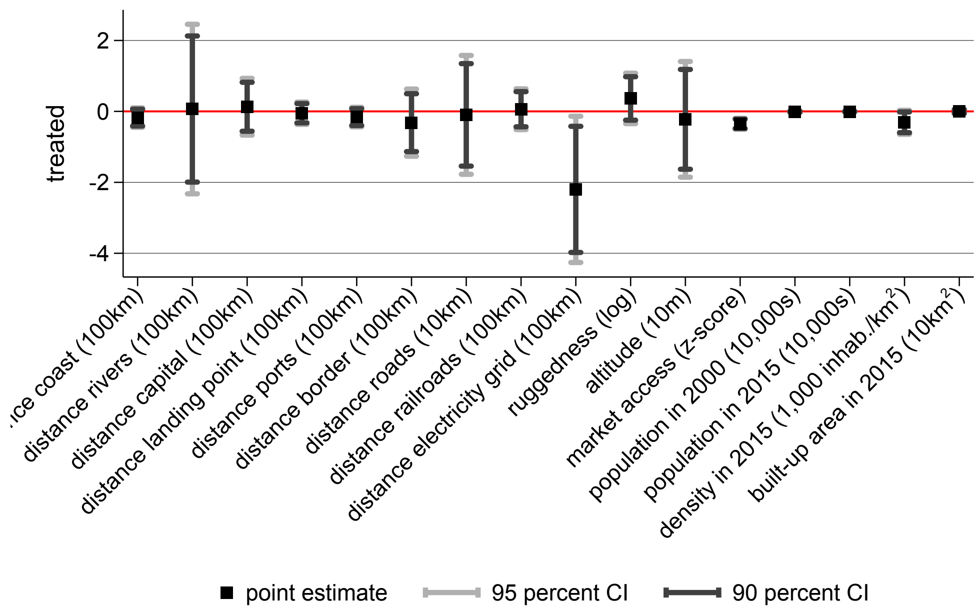
[16]This estimate is not reported. We restrict the sample to contain only countries which were connected before 2009. These countries' first internet connection is at basic speed.

Figure 4: Balance test



Notes: Confidence intervals reported at the 90% and the 95% level.

Figure 5: Balance test



Notes: Confidence intervals reported at the 90% and the 95% level. Sample: Countries with their first internet connection before 2009.

the electricity grid are connected earlier. Market access is also statistically significantly correlated with the connection date. However, with a very small point estimate.

The exogenous shock is at the country level. As the rollout of the fiber network might not be random, we can still test whether our observable time-invariant controls correlate with the defined treatment status of the town, given country fixed effect, in the cross section. Although not being random, if the treatment status cannot be predicted from the controls, this would strengthen our identification. The access point rollout follows existing (transportation) infrastructure. Capital cities are usually close to the landing point (in coastal countries) and are aimed to be connected early. Thus, a negative correlation with distance to capital cities and roads and railroads is not surprising (Figure 4). We will handle this factor in heterogeneity exercises deeply later. Importantly, there is no statistically significant correlation with other observables. Additionally, we check whether the treatment status correlates with today's towns' education, finance, and health infrastructure, again given country fixed effects (Figure A.1). Only for colleges, we find some statistically significant correlation.

4 Data

We analyze the effect of Internet availability on local economic growth in SSA. 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 and construction year of access points to the national fiber-cable backbone. Moreover, we use data on towns' built-up area, merged with characteristics, such as administrative status and population, and infrastructure, such as (rail)roads, mobile coverage, and the electricity grid. Finally, we make use of the countries' connection dates to the submarine cables (SMCs) or via neighboring countries.

4.1 Local Economic Activity: Nighttime Lights and Built-up Areas

We measure economic activity at the town level. To identify town locations and extent, we use the established data from *Africapolis* on built-up areas.[17] This database contains the geographical delineation of 5,616 SSA agglomerations with more than 10,000 inhabitants in 2015. The median size is around 21,000 inhabitants and about 90 percent have less than 100,000 inhabitants. The population is also been made available for earlier years. [18]

Since geographically and chronologically granular data on economic activity in SSA is lacking, especially for the period we investigate, we deploy NTL satellite data. This data measures human-caused NTL emissions in a geographically high resolution and on a yearly basis. The data was collected in the *Defense*

[17] `https://africapolis.org`, accessed on 05.01.2023.

[18] In 2000, which is closer to the first SMC connections, the median population was close to 10,000 inhabitants and about 90 percent of the agglomerations had a population of less than 45,000 inhabitants.

Meteorological Satellite Program (DMSP) *Operational Linescan System* (OLS) between the early 1990s and 2013. The instruments of DMSP-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 square kilometers at the equator). The data is then combined to yearly composite images. We use the harmonization by Li et al. (2020). This procedure excludes noise from aurora, fires, boats, and other temporal lights and inter-calibrates the data globally for each year as well, making it temporally consistent.

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 the city level. At larger geographic resolutions, Bruederle and Hodler (2018) added the relation to household wealth, education, and health for *Demographic and Health Surveys* cluster locations as well as for grid cells of roughly 50×50 kilometers.

4.2 Internet Infrastructure: Backbone Access Points and Submarine Cables

For the treatment year, we use information on SMCs' landing dates on the shores of SSA countries for coastal countries from *Submarine Cable Map*.[19] We geo-coded the landing point to merge it to the respective built-up area. If the connection was established through a neighboring country, we assign the establishment year of a country border access point to the national fiber-cable backbone as the treatment year. The geo-locations of the access points and their respective establishment years come from *Africa Bandwidth Maps*.[20] Figure A.2 shows a map of all access points and their construction year. Table B.1 shows the country-specific connection years for all SSA countries that were connected before 2009. In the last column, the year of the speed upgrade through the next SMC is shown. These SMCs had a lot higher capacities and landed in SSA between 2009 and 2012.

Africa Bandwidth Maps contains the most comprehensive set of access points for Africa. It covers the period starting from 2009 and is updated on a yearly basis. The data is directly sourced from the network operators.[21] As access points existing in 2009 were largely established earlier, we conducted an extensive review of backbone deployment projects for each country. Thereby, we determined the construction years of the access points from 2009 going back to the late 1990s for all SSA countries. Note that it was not always possible to determine the exact year of construction. However, in these cases, it was still possible to determine which access points were constructed until in the year the countrywide Internet connection was established, which is still sufficient for our analysis. This makes it possible to identify which towns already had access to the national fiber-cable backbone when the Internet became available for the first time. We

[19] `https://www.submarinecablemap.com`, accessed on XYZ.

[20] `http://www.africabandwidthmaps.com`, accessed on XYZ.

[21] To date, there are 2,708 access points in SSA countries. About half of them were constructed since 2013. Especially in bigger cities, more than one access point is usually established to account for the limited capacity of each access point. In 2019, for example, although 189 new access points were constructed, only 27 new cities and towns were connected. In total, around 900 cities and towns have an access point close by.

match access points to towns via their geo-location: First, we calculate the distance between the towns' border and the closest access point. Then, we assign a national fiber-cable backbone connection to towns within a distance of less than 10 kilometers.[22]

4.3 Further Data Sources

We use the share of the area a town has mobile coverage as control variable for the rollout of an alternative digital infrastructure.[23] The data is sourced from *Collins Bartholomew*.[24] Though, since the early 2000s the new mobile-phone standard became 3G, none of the countries in our analysis has rolled out 3G. Therefore, mobile coverage in our data refers to GSM (2G) which allows for basic applications (calls and SMS) but not for mobile Internet.

From *OpenStreetMap (OSM)*, we take the definition of nodal cities. Capital cities and region capitals are marked there. Additionally, we take the location of rivers and information on today's financial, health, and educational infrastructure from OSM. For the definition of economic centers, we take the population in the year 2000 from *Africapolis*. This data source also contains information on towns' altitude, and built-up area and population density in 2015. Moreover, we use the locations and the population data to calculate a market access measure.[25] As time-constant measures of infrastructure, we take shapefiles for roads and railroads from *Natural Earth (NE)*. *Africa Infrastructure Country Diagnostic (AICD)* provides a shapefile for the electric grid in the year 2007. We take the location of ports from the *World Port Index* .[26] For the terrain ruggedness, we take he 30 arc-second cells raster data from Nunn and Puga.[27]

We examine changes in industry shares as a mechanism for the Internet growth effect. We aggregate census microdata from *IPUMS-International* to a regional level of second order.[28] For the industry shares, the data contains whether the employment is in agriculture, manufacturing, or services. The data comes usually every ten years. Therefore, we estimate a long difference with one pre-treatment and one post-treatment period.

To exclude ethnic favoritism as a plausible explanation for the effect, we use the "geo-referencing of ethnic groups" (GREG) data set.[29]

[22] We conducted interviews with industry experts to verify this decision. In addition, in a robustness check we vary this distance.
[23] The share is usually either 0 or 1.
[24] `https://www.collinsbartholomew.com/`, accessed on XYZ.
[25] XYZ source.
[26] `https://msi.nga.mil/Publications/WPI`, accessed on XYZ.
[27] `https://diegopuga.org/data/rugged`, accessed on XYZ.
[28] The admin-2 level is below the state level.
[29] `https://icr.ethz.ch/data/greg/`, accessed on XYZ.

4.4 Combining the Data

Our analysis is focused on rather mid-sized towns. These towns might not be precisely measured by the satellites' instruments. In fact, for very small towns we observe that they are not bright enough to reach the instruments' sensitivity threshold in each year. We, therefore, remove towns that do not have positive light intensity in all years. Thus, we reduce measurement error and additionally the sample loses very small towns. If towns are visible in all years, we can additionally be sure that they have stable electricity available. So, we can rule out a potential source which might confound our results.

As light blurs out to adjacent pixels, cities appear bigger in the data than they actually are. By taking the extent of the towns in 2015, we capture some of the blurring as the towns might have been growing after our observation period. However, for some towns, the NTLs still might blur over the extent of the built-up areas. Therefore, we account for blurring by adding a radius of 2 kilometers to the built-up area, such that the growth of light emissions in the extensive margin is properly captured.[30] 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). In our sample, less than 2 percent of pixels are assigned light intensities of 60 or more.

Figure 6 shows for Dassa-Zoumè, Benin, its NTL emissions, built-up area, and infrastructure. A road and a railroad connecting Dassa-Zoumè with its neighbouring cities (red and darker red line) and the access points (red triangles) constructed in 2001 are shown in all panels. Panel (1) shows moreover the NTLs for the year 2004 (three years after the countrywide Internet connection and at the end of the analysis period), where a brighter gray reflects higher NTL intensity. Panel (2) adds Dassa-Zoumè's built-up area from *Africapolis* in a dark blue. It shows that through blurring, the NTLs exceed the built-up boundaries. Therefore, we draw a buffer of 2 kilometers around the built-up area in a lighter blue (shown in Panel (3)). This allows us to take all NTL emissions into account.

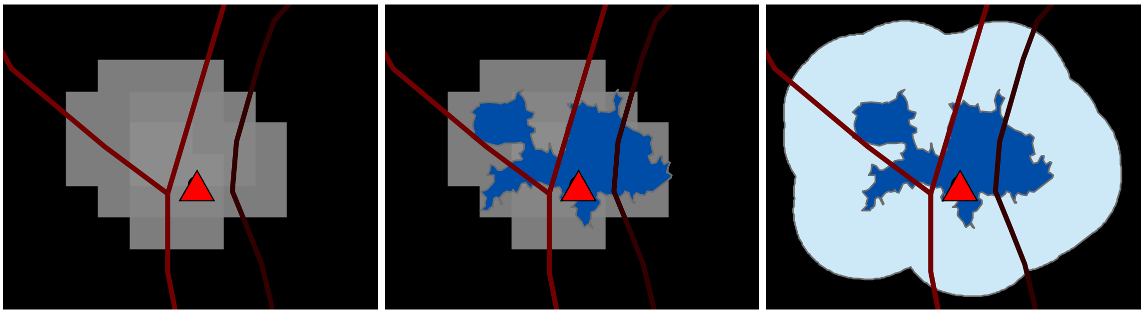
Within each town, we define several outcome measures.[31] Local economic activity is measured by summing the light intensity of all pixels within a town (and the 2 kilometers buffer) in each year. This measure was established by Storeygard (2016) and accounts for both increased light intensity and geographical extension. As alternative measures, we calculate the average light intensity of pixels and the sum of all lit pixels, ignoring light intensity. We interpret the average light intensity as a proxy for density in terms of population or per capita economic activity (intensive margin) and the sum of lit pixels as a proxy for spatial extension of a town (extensive margin). For an example treatment town and for an example control group town, Figure A.3 shows how their respective NTLs have changed from the year of the Internet connection to three years later.

The final estimation sample consists of ten countries, which (i) were connected at basic speeds, (ii) have at least one town in the treatment and in the control group, which is not a nodal city, and (iii) have at least

[30] For robustness, we also show the results for a specification without a buffer as well.

[31] As specified in section 3, we apply the logarithm of each outcome measure in the regression analysis.

Figure 6: Data example: Dassa-Zoumè, Benin (2004)



(a) **(b)** **(c)**

Notes: Panels (1) through (3) show our data for Dassa-Zoumè, Benin, in 2004. Dassa-Zoumè is in the treatment group as one of the incidentally connected towns. Panel (1) shows the access point existing in 2001 (red triangles) and NTLs for the year 2004 (three years after the connection year of Benin). The access point lies within the towns boundaries. The red line represents a major road connecting Dassa-Zoumè with its neighboring cities and the darker red line the railway connection. The black-to-white scale indicates light intensity, with brighter colors reflecting higher light intensities. Panel (2) adds its built-up area from *Africapolis* (shown in darker blue). Finally, Panel (3) shows in blue a 2 kilometers buffer around that built-up area.

three post-treatment years before a high-capacity SMC connects them.[32] While the first restriction is due to the capacity of SMCs, the second restriction depends on the access point rollout within each country. The last restriction is necessary to estimate the effects of Internet availability at basic speeds on a balanced panel. A longer post-treatment period would shrink the sample further. Therefore, we estimate the main specification with three post-treatment years. For robustness, we will relax these restrictions. The first connected country in our sample is Senegal, which was connected in 2000. Therefore, we are less restricted in the pre-treatment period and do not lose any country there. For the estimation sample, Figure A.4 shows the geographical distribution and the location of the treatment and control group towns without the nodal cities. There are four countries in West Africa and Southern Africa, respectively, and two countries in East Africa in our sample. Of the ten countries, five are coastal and five are landlocked.[33]

4.5 Descriptive Statistics

We focus our analysis on mid-sized towns. From 510 agglomerations, for which NTL data is detected in each year, in the ten countries of the estimation sample, 143 were connected to the Internet via an access point before the country was connected via SMC or a neighboring country. Therefore, they are part of our treatment group. Of these agglomerations, 70 are nodal cities. Another 147 towns got an access point in the subsequent years and are therefore in the control group. The remaining 118 agglomerations are still not

[32]These countries are Angola, Benin, Botswana, Ethiopia, Mali, Sudan, Senegal, Togo, Zambia, and Zimbabwe. An overview of the procedure how the estimation sample of countries emerges can be found in Appendix D.

[33]Sudan is a special case as ten towns are in the control group but only one town is in the treatment group. Angola has very similar issues. We account for that by grouping fixed effects for East, Southern, and West African countries for robustness.

connected. Further 102 towns were connected in the three years after the countrywide Internet connection and are therefore not considered as they would confound our control group.

Figure A.5 compares the average size of cities and towns by the year they get an access point, relative to the treatment year. In the early years, until the Internet becomes available countrywide, many nodal cities are connected besides the towns in the treatment group. While connected nodal cities are bigger on average in the early years and decline in their size in subsequent years, towns in the treatment group only have a population of 16,595 inhabitants on average. Control group towns, which are connected in the subsequent years after the observation period of five years, have on average a very close population of 16,314 inhabitants.[34] This population varies for almost all subsequent years between a population of 10,000 and 20,000 inhabitants. Only in one year there is an outlier with an average population of close to 35,000 inhabitants. Especially, when only examining treated and control towns, i.e., excluding nodal cities, there is no clear pattern with respect to population size over time anymore. A decreasing pattern is clearly identified for nodal cities (especially for the first five years). This finding suggests that treated towns are not selected into treatment because of their population size. Moreover, nodal cities are still connected in further years after the arrival of the first internet connection, showing that the rollout continues through other parts of the country. Their size decreases after the first two years as capital cities are usually connected early and are usually a lot bigger than other nodal cities. Nonetheless, the size of later connected nodal cities is still bigger on average than the size of the control group towns.

Table B.2 gives a broad overview of the towns. The statistics of the outcome measure of the light intensity show a value of 463.04 on average one year before the treatment (161.50 at the 25th percentile, 285.00 at the median, and 530.50 at the 75th percentile). For the size of the towns, measured with the NTL data, values are as followed: 43.35 on average one year before the treatment (24 at the 25th percentile, 35 at the median, and 53 at the 75th percentile). On average, including no-lit pixels with a value of zero, towns have values of 7.50 on average one year before the treatment (3.22 at the 25th percentile, 5.25 at the median, and 10.07 at the 75th percentile). Given that the instruments pickup light usually at a threshold of 4, the average values are rather modest. The rather high number of lit pixels corresponds to the condition that towns have to show up in each year in the NTL data. Coming to the other variables, mid-sized towns have a population of around 20,500 inhabitants on average in 2000 (8,500 at the 25th percentile, 16,000 at the median, and 30,000 at the 75th percentile). Mobile coverage is available in about 62 percent of the towns one year before treatment, given that usually the percentage covered is either zero or one. By construction, the maximal distance to the closest access point is with 9.43 kilometers smaller than 10 kilometers. On average, this distance is a lot smaller with 1.26 kilometers. More than half of the towns have an access point even within the built-up area and most cities have it within 2 kilometers (1.21 kilometers at the 75th percentile). The distances to further infrastructure, such as the road network, railroad network, or electricity grid, are usually

[34] Treated and control group towns are not only almost identical in their average population, their population distribution also looks very similar (Figure A.6).

small with median distances of 0 kilometers (3.8 kilometers for the railroad network). Further distances are given for the next port, for coastal countries, as well as to the capital city, to the next regional capital, and geographical measures, such as the coastline or the next river.

5 Results

5.1 Main Effects and Mechanism

We estimate the effect of Internet availability on local economic growth. Particularly, we are interested in the effect of early Internet availability brought by the ‘first generation’ of SMCs. Nodal cities are excluded. We estimate a linear model on a balanced panel by difference-in-differences, where town and country-year fixed effects are included and standard errors are clustered at the town level. We measure economic activity by the logarithm of the sum of NTL intensities as the main outcome. Table 1 shows the main results. Columns (1) and (2) show the effect of Internet availability on light intensity. This effect is then separated into growth on the intensive and extensive margin (Columns 3 and 4). Column (5) investigates population growth.

In line with our expectations, we find an economically and statistically positive effect of the availability of Internet at basic speeds on local economic growth. In our preferred specification (Column (2)), towns which were connected to the Internet in the year of an SMC arrival become 7 percent brighter than towns without Internet access. This finding supports our initial claim that towns which get incidentally connected to the Internet grow faster in comparison to otherwise comparable towns. The mobile coverage control does not turn out to be statistically significant and is smaller in size in comparison to the main effect. It makes the estimation more precise as it controls for differences in another ICT. As it increases slightly the point estimates of the main effect, we are not worried that the main effect is transported through mobile coverage. We will discuss the role of mobile coverage in more detail in **??**.

We translate the effect in terms of light intensity to an approximation of the implied economic effect, by a back-of-the-envelope calculation using the GDP-luminosity elasticity from Henderson et al. (2012). Henderson et al. (2012) show that growth in light intensity serves as a good approximation of economic development at the country level. The elasticity remains robust for a global sample as well as a sample of low and middle-income countries. Storeygard (2016) further shows that the elasticity at the country level is if anything slightly higher for SSA countries and for coastal primates. Moreover, he shows that the elasticity holds at the sub-national level as well. We follow his argumentation that the elasticity in SSA could be lower as countries have lower light intensities on average or that the elasticity could be higher as SSA finds fewer top-coded cities. Both issues are specifically the case in our sample. We, therefore, use the elasticity of $\varepsilon_{GPD,light} = 0.284$ from Henderson et al. (2012), for which the calculation translates the increase in light intensity of 7 percent into about 2 percentage points higher GDP growth.

Table 2 investigates further Column (4) of Table 1. First, Column (4) of Table 1 is repeated. Then, instead of restricting the sample to lit pixels from 1995, we estimate on the brightest pixel from each year. Selecting the top x percent of each year and calculating the mean light intensity on these. Thus, previous unlit pixels do not affect the towns’ mean light intensity. Column (2) takes the top 10 percent. This value is increased

Table 1: The effect of Internet availability on the economic growth of towns

VARIABLES	(1) light intensity	(2) light intensity	(3) light intensity	(4) intensive margin	(5) extensive margin
connection × access	0.124***	0.134***	0.112**	0.0821**	0.0624*
	(0.0419)	(0.0423)	(0.0452)	(0.0316)	(0.0343)
GSM coverage		0.0835*	0.0801	0.0494*	0.0558
		(0.0477)	(0.0500)	(0.0294)	(0.0417)
Observations	2,519	2,519	2,519	2,519	2,519
R-squared	0.936	0.937	0.949	0.923	0.918
#countries	11	11	11	11	11
#towns	229	229	229	229	229
share treated	.454	.454	.454	.454	.454
town FE	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓
controls			✓	✓	✓

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. The measure for the intensive margin is logarithmic mean light intensity and for the extensive margin logarithmic sum of lit pixels (coded as 1 if a pixel is lit), all on the same area. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage is calculated as the percentage share of the town area, also including the 2 kilometers buffer, covered with mobile signal. Further control variables are distance to the capital, to the SMC's landing point, to the country's border, to the coastline, to the next port, to the next river, to the road and railroad network, to the electricity grid, and the town's altitude and the ruggedness of the area. All of these controls are constant over time and estimated through a vector of time-varying (yearly) estimates. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

by steps of 10 up to the top 50 percent upper half) in Column (6). In all specifications, the estimate remains very robust, but increases slightly from Column (2) towards (6). This shows on the one hand that towns are indeed becoming brighter on average. But on the other hand, there is no evidence that the brightest pixels (i.e., the town's business district) are becoming brighter, on average. However, recalling that we are investigating mid-sized towns, there might not be a classical business district.

Table 2: Intensive Margin

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
connection × access	0.0821**	0.0533*	0.0600**	0.0674**	0.0693**	0.0705**
	(0.0316)	(0.0280)	(0.0297)	(0.0311)	(0.0329)	(0.0337)
GSM coverage	0.0494*	0.0388	0.0429	0.0415	0.0425	0.0432
	(0.0294)	(0.0252)	(0.0268)	(0.0288)	(0.0300)	(0.0309)
Observations	2,519	2,519	2,519	2,519	2,519	2,519
R-squared	0.923	0.963	0.959	0.955	0.951	0.949
#countries	11	11	11	11	11	11
#towns	229	229	229	229	229	229
share treated	.454	.454	.454	.454	.454	.454
controls	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓
lights	lit in 1995	top 10 percent	top 20 percent	top 30 percent	top 40 percent	upper half

Notes: Outcome: intensive margin, measured as the logarithmic mean light intensity of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Column (1) repeat the measure of Table 1. Columns (2) through (6) restrict the pixels to those with the highest value in each year. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

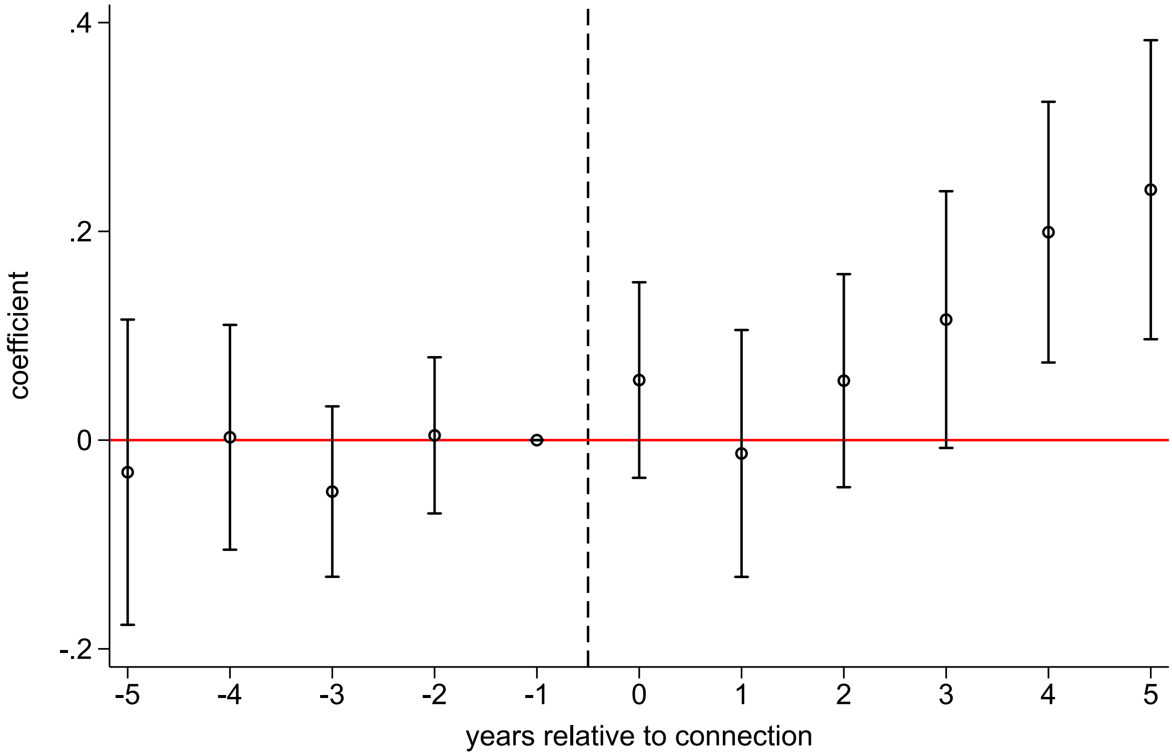
Figure 7 presents event-study coefficients for our preferred specification (Column 3 of Table 1). Before SMC connection, the point estimates are close to zero and statistically insignificant (even with 90% confidence intervals (Figure A.7)). This supports the assumption that treatment and control group towns are not on different growth parts preceding the countrywide internet connection, conditional on fixed effects. From the dynamic perspective, there is no evidence for a potential fading out of the effect. In contrast, in the year of the countrywide Internet connection, the point estimate turns positive but remains statistically insignificant. It again starts to increase in year t+2, increases further in all subsequent periods, and becomes statistically significant at the 5% level in year t+4.[35]

The increase of the effect over time is in line with the expectation that internet adoption takes time and that growth effects develop only some time after Internet adoption. Moreover, it indicates that the effect might not be completely induced by adopting firms but be partly induced by spillovers of the local economy. Finally, it is a strong sign that the effect is not coming from an electricity demand that was satisfied after

[35]Figure A.7 shows that in year t+3 the point estimate turn statistically significant at the 10% level.

giving internet access to the town. If that was the case, the increase in light emission should be found earlier and would not be increasing over time after the treatment year.

Figure 7: Event-study coefficients



Notes: The figure presents event study results based on Equation 2. Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. The event is defined as the year in which the country is connected to the Internet. Included are GSM coverage, all time-invariant controls as stated in the notes of Table 1, and town and country-year fixed effects. The event study is estimated on a balanced panel from which nodal cities as defined in Table 1 are excluded. Confidence intervals are drawn at the 95 percent level using standard errors clustered at the level of the closest access point.

We further investigate the intensive and extensive margin of regional development. Therefore, we take the logged mean luminosity and the logged sum of lit pixels (Table 1, Columns (3) and (4)) as outcome measures. The observed increase in light intensity of towns having Internet available could be explained by growth in productivity, population density, or populated area, among other explanations. While we investigate the effects on population later, the intensive and extensive margin can show whether the town is growing only in size at its border (extensive margin) or whether existing pixels are getting brighter (intensive margin). While towns' light intensity increases by 7 percent, their brightness increases by 5 percent and their size increases by 5 percent as well. The observed increase in light intensity can thus be explained by both an increase in brightness (Column (3)) and in size (Column (4)). However, while the effect on the intensive margin is statistically significant at the 5% level, the effect on the extensive margin is only statistically significant at the 10% level. As brighter lights glow further, the increase in the extensive margin is at least partly an indirect effect of the growth in the intensive margin. This finding is important as it is not clear a priori that the effect manifests beside the towns' border.

Next, we show how the industry composition changes in regions with Internet access as these changes might be a channel through which Internet availability affects local economic growth. We use survey data from IPUMS-International, a collection of census microdata, to calculate the share of jobs in each industry (agriculture, manufacturing, and services).[36] We estimate the effects on growth rates and changes in the industry shares.

The data contains 21 SSA countries. However, only eight of them have more than one year, as for many countries early data is not available. From the remaining countries, seven countries were connected to the Internet in 2008 or earlier and six of them had constructed at least some access points when the countrywide Internet connection was established. Finally, for five countries, both a treatment and control group can be defined. These countries are: Benin, Mali, Malawi, Mozambique, and Zimbabwe. Allowing for surveys close to the connection year, we can estimate also on countries that were connected late (in 2006 and 2007), for which the upgrade induced by 'second generation' SMCs came shortly after the first connection. Malawi and Mozambique only have two post-treatment years and were not included in the estimations so far. The treatment is defined as before: We remove regions that contain a nodal city and define a region as treated if at least one town in that region has early Internet access. For most of the remaining countries, data is available with a frequency of ten years. Only Mali (and Benin) has a difference of eleven years (once). For Benin, there are three survey rounds available. However, as one round was in 2002, only one year after the connection year, we drop this year. Thus, we estimate a long difference with one survey year before the arrival of the countrywide Internet connection and one afterwards.

Figure A.8 shows the changes descriptively. Both the treatment and the control group have relatively high shares of agricultural employment. Unexpectedly, the shares in services are higher than in manufacturing. After the treatment, agricultural shares decline in both groups and manufacturing and services shares increase in both groups. Especially for manufacturing, but also in both other industries, the changes are larger in the treatment group.

The results in Table 3 indicate that Internet availability shifts jobs from agriculture to manufacturing (and slightly to services). In Columns (1) through (4), we estimate growth rates, while we estimate changes in industry shares in Columns (5) through (7). In the final sample 144 towns are contained, of which 26 are treated. The share of jobs in agriculture declines by more than 3 percentage points. In contrast, job shares in manufacturing increase by more than 2 percentage points and job shares in services increase by around 1 percentage point. However, only the effect on manufacturing is statistically significant at

[36] `https://international.ipums.org/international/`, accessed on XYZ.

the 5% level. Comparing these results with the growth rates in Columns (2) through (4), the signs for the agricultural and manufacturing sector are equal. However, all estimates lack statistical significance. Column (1), again underlines that the effect is not driven by migration as the total number of individuals is not statistically greater in the treatment group than in the control group with a point estimate very close to zero. The results are in line with Hjort and Poulsen (2019), who find an increase in employment in Ethiopian manufacturing firms and an increase in net firm entry in services. As manufacturing might emit NTLs differently than agriculture, the estimated higher light intensity might reflect both: economic growth through more manufacturing jobs and a change in the industry structure (independent from the growth). We lack evidence whether the effect stems from newly created firms or growing already existing ones.

Table 3: Employment growth rates and shares by industry

	growth rate		industry share	
	(1)	(2)	(3)	(4)
	total	agriculture	manufacturing	services
access	-0.0215	-0.0328	0.0225**	0.0103
	(0.0569)	(0.0239)	(0.0102)	(0.0184)
GSM coverage	-0.0570	-0.00696	-0.000510	0.00747
	(0.0570)	(0.0172)	(0.00583)	(0.0130)
observations	288	288	288	288
R^2	0.967	0.954	0.893	0.958
#countries	5	5	5	5
#regions	144	144	144	144
share treated	.181	.181	.181	.181
region FE	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓

Notes: Regional industry composition comes from *IPMUS International*. All specifications are estimated on a sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 50,000 inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.2 Robustness

Table 4 presents regression results for robustness. Column (1) repeats our preferred specification from Table 1. First, we test the robustness of our results to alternative assumptions about the variance-covariance matrix by clustering the standard errors at a higher level and by re-estimating Equation 1 with different fixed effects. Then, we show that the results are robust to different choices of removed towns. Finally, we drop the 2 kilometers buffer around the built-up area and test for ethnic favoritism.

Error Correlation within Regions A potential concern is that model errors are spatially correlated within regions. If more than one town is located within 10 kilometers to the access point, an access point would serve more than one town. Therefore, we cluster at the access point level in the main specification. Although, treatment and control group towns might not be closes to one another, as they are certainly assigned to different access points, for some this might still be the case. Moreover, the access point might generate further spillover effects in the towns' surrounding area. To take this into account, we apply a higher level of clustered standard errors for robustness. We re-estimate Equation 1 correcting standard errors for clusters at the level of states. Column (2) of Table 4 presents the estimates with a higher level of clustered standard errors. The standard error of our variable of interest increases only very slightly (from 0.0452 to 0.0499).

Table 4: Robustness

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
connection × access	0.112**	0.112**	0.148***	0.0814**	0.0832**	0.0818***	0.0793**	0.0933**
	(0.0452)	(0.0499)	(0.0412)	(0.0400)	(0.0361)	(0.0304)	(0.0398)	(0.0364)
GSM coverage	0.0801	0.0801	0.157***	0.00346	0.0366	0.0502	-0.0214	0.183**
	(0.0500)	(0.0528)	(0.0446)	(0.0267)	(0.0375)	(0.0403)	(0.0307)	(0.0753)
Observations	2,519	2,519	2,519	4,224	3,619	3,618	2,101	1,793
R-squared	0.949	0.949	0.937	0.938	0.944	0.946	0.982	0.946
#countries	11	11	11	11	11	11	10	
#towns	229	229	229	384	329	329	191	163
share treated	.454	.454	.454	.271	.316	.62	.445	.454
controls	✓	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓		✓	✓	✓	✓	
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓
cluster level	access point	state	access point	access point	access point	access point	access point	access point
#ethnic group-countries								15
ethnic group-country × year FE								✓
no buffer							✓	
general DiD						✓		
all late APs included					✓			
no not treated towns				✓				
year FE			✓					

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. Columns (4) and (5) extend the control group. Column (6) extends the treatment group. Column (7) is estimated on a sample containing pixels within the *Africapolis* built-up area only. Ethnic groups for Column (8) are drawn from *GREG* data. All specifications include town and country-year fixed effects (only Column 3 only includes year fixed effects and Column 8 even ethnic group-country-year fixed effects) and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point (only in Column 2 at the state level) and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Different Fixed Effects As explained above, we apply country-year fixed effects to account for country-specific growth paths in the countries' economies. For robustness, we re-estimate Equation 1 with the classical two-way fixed effects: towns and calendar years. This specification is less demanding in the set of fixed effects. A concern with these fixed effects might be that countries on a higher growth path might construct more access points faster. Therefore, this specification serves as a robustness check and not as the main specification. Nevertheless, the estimate presented in Column (3) of Table 4 even increases. In subsection 5.3 (External Validity) we estimate a further model with the classical two-way fixed effects, which allow for a sample with more countries as the treatment and control group are compared across countries.

Extending the Control Group with Towns Not Having an Access Point Thus far, we tested the robustness of our results to alternative assumptions about the variance-covariance matrix. In the following, we add so far excluded towns (mainly) to the control group. This increases the sample size and shows robustness of the results with respect to other choices who to define the estimation sample. In Column (4) of Table 4, we add to the control group also all towns without an access point. They are comparable to the control group as they cannot access the Internet when it becomes available countrywide. At that time, it might not yet be known which towns will get an access point in the future. While this sample increases to 384 towns, the share of treated towns declines to 27.1 percent. The estimation results remain unaffected.

Access Points During the Post-Observation Period In Column (5) of Table 4, we add those towns to the estimation sample that were connected only within five years after the treatment year. In our main specification, these towns are excluded as they neither belong clearly to the treatment nor the control group and would thus confound our analysis. Now, we add them to the control group and treat them as if they could not access the Internet in the whole period. Thus, we add further 100 towns in the same eleven countries to increase the estimation sample. As all towns by construction are added to the control group, the share of treated towns declines to 0.316. The point estimate of the main effect declines as well, but only very slightly, as these towns were originally been excluded for being potential confounders.

No Buffer In Column (6) of Table 4, we remove the 2 kilometers buffer and estimate on the original *Africapolis* built-up areas, which can be interpreted more as the core of the town. We adjusted the built-up areas in the main specification due to the blurring of the NTL data. When examining the smaller built-up area, we might lose some pixels at the towns' border. These pixels might be of low intensity. Thus, this approach loses some towns, which have only lit pixels outside the built-up area but are still in the buffer for at least one year. This also leads to losing one country (Angola).[37] The main effect, in comparison to Table B.3 (Column 2), is robust (if anything it slightly increases). With this robustness check, we can thus

[37] Estimating on the sample of the main specification without Angola is shown in Table B.3 (Column 2). The sample shrinks to 212 towns. The main effect estimate is with 0.078 (standard error 0.0426) a bit lower than in the main specification.

not only show that our results do not depend on the adjustment of the built-up area but also that local growth does not predominantly happen at the towns' border.

Ethnic Favoritism A further concern could be that certain ethnic groups were favored during the rollout. Though the exogenous shock comes from the countrywide connection year and the parallel trends in the event study do not underpin this concern, a remaining threat could be that certain ethnic groups are also favored in other dimensions, which causes the observed difference in growth over time. Our strategy to overcome this threat is two-fold. First, many countries construct access points for more than one ethnic group before the treatment period (Figure A.9). This indicates that not a specific ethnic group is favored by giving them access to the Internet. For the countries in our analysis, all countries but Angola provided at least two different ethnic groups with access points. And Angola only established very few early access points in total. Therefore, the low number of equipped ethnic groups is not surprising. On the other hand, Ethiopia and Togo provided internet access for even six different ethnic groups very early. Second, we perform our analysis by constructing country-ethnic group entities instead of countries. By estimating (Equation 1) including town fixed effects and country-ethnicity-year fixed effects, treatment and control group towns are compared within an ethnic group. If ethnic favoritism were at play and would drive the found effects, our estimate should vanish as towns with certain ethnic groups should grow, and less importantly get an access point, while towns with other ethnic groups remain on a worse growth path. The results are shown in Column (8) of Table 4. A slightly smaller sample size shows that for most ethnic groups for which access points were constructed in the treatment period, access points were also constructed afterwards. Only in Botswana, this is not the case. In the remaining ten countries, there are thirteen ethnic groups and fifteen country-ethnic group entities in the estimation.[38] The result remains robust, showing that even comparing treatment and control group towns of the same ethnic group, internet availability has a positive effect on local economic activity.

Role of transportation infrastructure We further control the role of transportation infrastructure access. Table 5 re-estimates Equation 1 excluding towns with out access to a major paved road (Column 1), to the railroad network (Column 2), and to either of them (Column 3).[39] Column (4), then, also excludes towns which are in the road between the two most important points of a country along backbone network in the connection year. This most important road usually connects the capital and the next biggest city. Column (1) shows that, that few towns without road access do not affect the results. Column (2) indicates an a lot stronger effect for towns with access to the railroad network. The results in Column (3) then look very similar to those in Column (1). In Column (4), one country is dropped out as the early role out only was along the most important road. For the other countries, the estimate remains. Hence, the rollout giving

[38]Estimating on the sample of the main specification without Botswana is shown in Table B.3 (Column 4). The sample shrinks to 219 towns. The main effect estimate is with 0.117 (standard error 0.0459) almost identical to the main specification.

[39]In Column (2), the sample shrinks to only 91 towns, such that we had to estimate without additional time-invariant controls to not run into singularity issues.

an advantage for incidentally connected towns at the most important inland connection, does not determine the positive effect.

Table 5: Robustness

VARIABLES	(1)	(2)	(3)	(4)
connection × access	0.0938**	0.198**	0.0978**	0.0870*
	(0.0462)	(0.0945)	(0.0452)	(0.0479)
GSM coverage	0.0791	0.0837	0.0842*	0.0290
	(0.0495)	(0.0735)	(0.0485)	(0.0553)
Observations	2,101	1,001	2,167	2,211
R-squared	0.955	0.949	0.956	0.931
#countries	11	11	11	10
#cities	191	91	197	201
share treated	.461	.538	.467	.413
controls	✓		✓	✓
town FE	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓
without	no road access	no railroad access	no transp. infra.	first roads

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Towns without (rail-)road are excluded (Columns 1 through 3) and in Column (4), towns with a particularly good transportation infrastructure, being at a country's most important road, are excluded. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.3 Extension

The robustness checks possibly allow to estimate on a larger sample including more countries. Before, we wanted to show that the estimates remain in different specifications on the same sample. Now, we want to show the same specifications as before but on a larger sample.

Column (1) of Table 6 repeats the main specification. In Column (2), we repeat the robustness check with only year fixed effects. However, now we relax the restriction that each country must contain a treatment and a control group town. The sample increases to 20 countries, containing 353 towns, with a decreased share of treated towns of 31 percent. The estimate increases, even in comparison to Table 4 (Column 3). In Column (3), towns without an access point are included to the control group. Two countries are added that now have at least one control group town (and not only a treated town as before). In comparison to Column (4) of Table 4, the sample increases by eleven towns and the estimate increases slightly. Also when allowing towns with an access point which was constructed shortly after the nationwide connection was established to be included in the control group, the sample increases by two countries. In comparison to Column (5) of Table 4, the sample increases by eight countries and the main effect increases only very slightly. Lastly, when estimating on the sample without the 2 kilometers buffer, an additional country and seven additional towns are included. Again, the main effect increases only very slightly. Overall, the main effect remains when not restricting the sample to the original countries from Table 1.

Table 6: Robustness

VARIABLES	(1)	(2)	(3)	(4)	(5)
connection × access	0.112**	0.227***	0.105***	0.0976***	0.0835**
	(0.0452)	(0.0424)	(0.0368)	(0.0357)	(0.0373)
GSM coverage	0.0801	0.0417	0.00791	0.0316	-0.0160
	(0.0500)	(0.0298)	(0.0260)	(0.0364)	(0.0296)
Observations	2,519	3,883	4,345	3,707	2,178
R-squared	0.949	0.916	0.937	0.944	0.981
#countries	11	20	13	13	11
#towns	229	353	395	337	198
share treated	.454	.309	.268	.315	.455
controls	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓
country × year FE	✓		✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓
year FE		✓			
no buffer					✓
all late APs included				✓	
no not treated towns			✓		

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. Columns (3) and (4) extend the control group. Column (5) is estimated on a sample containing pixels within the *Africapolis* built-up area only. All specifications include town and country-year fixed effects (only Column 3 only includes year fixed effects) and are estimated on a balanced panel, but not restricted to the original set of countries from Table 1. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We are interested in the effect of internet availability at basic speeds on local economic growth. Therefore, we limit our analysis to countries that were connected to international fiber cable with a narrow capacity. In Table 7, we rephrase the question and include also later connected countries: What is the extensive margin of internet availability (i.e, the first internet connection) on local economic growth? Later connected countries already are connected with international fiber cables with a much broader capacity. Column (1) shows that when including these three countries, the effect size remains and becomes statistically significant even at the 1% level.

So far, we defined "early internet" with respect to the initial connection, allowing the international fiber cable could be upgraded to broader capacities in the observation period. This allows for estimating on a larger sample with five post-treatment periods. Restricting the sample to years when only internet at basic speeds was available and re-estimating Equation 1 reduces the sample to six countries (Column 2). However, the estimate doubles approximately. These early connect countries, with a long time between the first connection and the speed upgrade (which started in 2009) are mostly coastal. To get back to the original sample of countries, we reduce the number of post-treatment periods to four and then three. Containing more countries, the estimate decreases close to the original level.

Table 7: Robustness: Speed Upgrade from 2009 Onwards

VARIABLES	(1)	(2)	(3)	(4)
connection × access	0.113***	0.251***	0.130***	0.0932**
	(0.0407)	(0.0578)	(0.0463)	(0.0412)
GSM coverage	0.0581	0.157**	0.114*	0.0949**
	(0.0489)	(0.0632)	(0.0576)	(0.0440)
Observations	3,058	1,507	1,960	2,088
R-squared	0.946	0.952	0.954	0.955
#countries	14	6	9	11
#towns	278	137	196	232
share treated	.511	.416	.434	.418
controls	✓	✓	✓	✓
town FE	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓
years after upgrade included	✓			
connection date restricted		✓	✓	✓
post-treatment periods		5	4	3

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Column (1) has no restriction to includes countries that were connected before 2009. Columns (2) through (4) exclude countries with few post-treatment years before 2009 and vary the number of post-treatment years. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.4 Placebos

The estimated causal effect comes from the interaction of an indicator that depends on the timing of the countrywide internet connection and an indicator that depends on the timing of the intra-country rollout of the backbone network. Our placebo exercise is three-fold: (1) we randomize treatment and control group towns within each country, (2) we add placebos for other non-digital infrastructure, and (3) we randomize country connection years.

Placebo tests with non-digital infrastructure In Table 8, we include additional controls. To save costs, fiber cables are rolled out along the existing (transportation) infrastructure (cf. subsection 2.2). We control for this other infrastructure to rule out that towns closer to this non-digital infrastructure grow faster when the countrywide Internet connection is established, irrespective of whether they are in the treatment or the control group. Similarly, we control for the electricity grid.[40] Unlike for mobile coverage, we do not have time-varying data on other infrastructures. In the main specification, we already control for the distance to these non-digital infrastructures (interacted with year indicators). We now generate indicators similar to the one for the access to the backbone to investigate the role of these non-digital infrastructures further: for each non-digital infrastructure, we construct an indicator that turns one if a town has direct access to this infrastructure or if the infrastructure is within a distance below of 10 kilometers (as we defined treated towns with access points) and then intersect these indicators with the connection indicator with the purpose to construct placebo treatments. We follow these two approaches as non-digital infrastructure might only be effective with direct access. Nevertheless, we construct indicators with the same distance threshold as we applied for internet access for robustness. In Columns (1) and (5), we control for different effects for towns next to a greater (paved) road. These are 78 percent of the sample with direct access and even 93 percent with access within 10 kilometers. In Columns (2) and (6), we control for different effects for towns next to the railroad network. These are 41 percent of the sample with direct access and 55 percent with access within 10 kilometers. In Columns (3) and (7), we control for different effects for towns next to the electricity grid. These are 58 percent of the sample with direct access and 79 percent with access within 10 kilometers. Finally, we include all infrastructure controls jointly (Columns 4 and 8). In each case, the estimate of the main effect remains close to the original level and turns at the 1% level statistically significant. When including all indicators of direct access the estimate even increases slightly. The estimates of the placebos lack statistical significance and are rather small in their economic significance. One exception are roads. When controlling for all direct placebos, the estimate for access to a road turns statistically significant at the 10% level. For the broader access of having a road within 10 kilometers, the estimate is with 0.19 very large and statistically significant at the 5% level. This remains when adding the other two placebos with

[40]For the infrastructure, we cannot be sure that roads and railroads were existing prior to the rollout of the national backbone. However, as we know that the rollout followed this infrastructure we take the existing data for the placebo exercise. Moreover, the electricity grid data we have is from 2007 and thus might assign an electricity grid to towns that only were connected to the electricity grid when they became access to the Internet after the countrywide Internet connection was established. We, therefore, discuss the role of the electricity grid in more detail later.

the same distance threshold. However, only about 7 percent of towns being further away from a road than 10 kilometers remain.

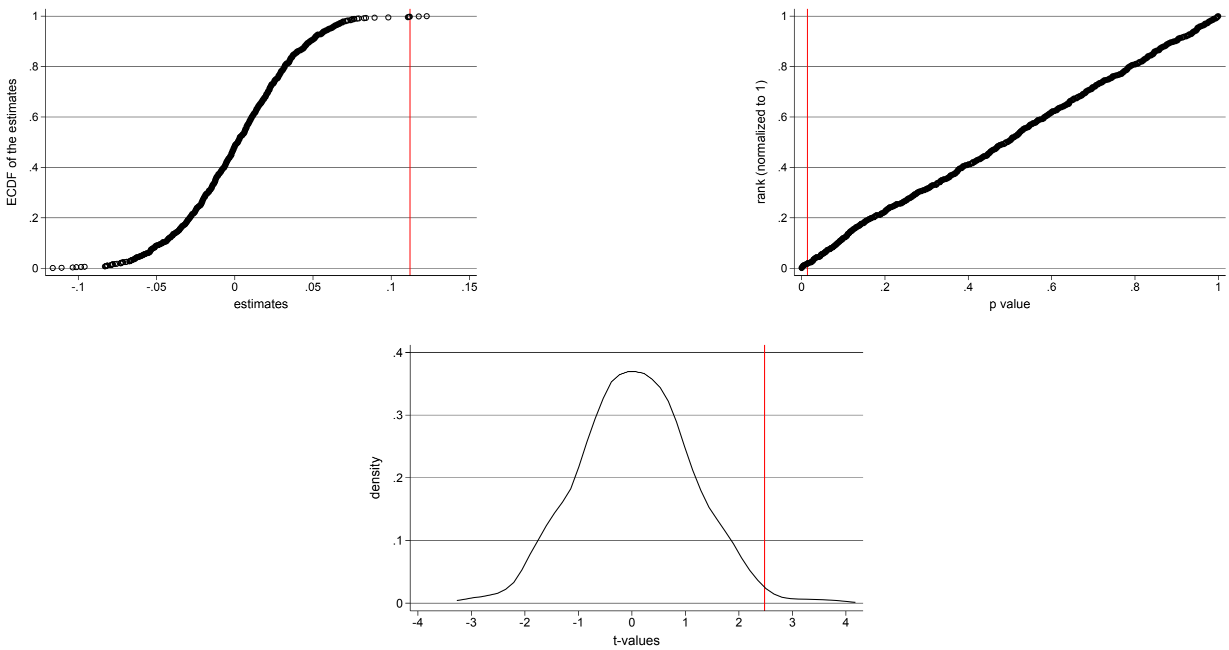
Table 8: Placebo (competing infrastructure)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
connection × access	0.118***	0.118***	0.116***	0.129***	0.111**	0.119**	0.114**	0.121***
	(0.0448)	(0.0444)	(0.0444)	(0.0431)	(0.0448)	(0.0455)	(0.0453)	(0.0455)
connection × road (dummy)	0.0778			0.0988*				
	(0.0632)			(0.0581)				
connection × railroad (dummy)		-0.0215		-0.0453				
		(0.0445)		(0.0413)				
connection × electricity grid (dummy)			0.0409	0.0393				
			(0.0443)	(0.0429)				
connection × road (within 10km, dummy)					0.189**			0.198**
					(0.0949)			(0.0976)
connection × railroad (within 10km, dummy)						-0.0144		-0.0257
						(0.0487)		(0.0524)
connection × electricity grid (within 10km, dummy)							0.00871	0.0160
							(0.0454)	(0.0525)
GSM coverage	0.0869*	0.0772	0.0898*	0.0936**	0.0885*	0.0776	0.0875*	0.0946**
	(0.0483)	(0.0487)	(0.0485)	(0.0455)	(0.0474)	(0.0486)	(0.0491)	(0.0451)
Observations	2,519	2,519	2,519	2,519	2,519	2,519	2,519	2,519
R-squared	0.948	0.948	0.948	0.947	0.948	0.948	0.948	0.947
#countries	11	11	11	11	11	11	11	11
#towns	229	229	229	229	229	229	229	229
share treated	.454	.454	.454	.454	.454	.454	.454	.454
interaction avg.	.78	.41	.58		.93	.55	.79	
controls	✓	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. The simple dummy shows whether a town has direct access to the (rail-)road network or the electricity grid; the other dummy whether this infrastructure is available within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Randomly Assigned Treatment For Figure 8, we constructed a placebo access by randomly assigning the treatment status, sampled from the set of treatments within each country, while maintaining each country's connection date. We do so to conduct a non-parametric permutation test of $\beta_1 = 0$. In the top left, the Empirical CDF of the estimates is depicted, resulting from permuting treatments 1,000 times and re-estimating Equation 1 on the placebo internet access following Chetty, Looney, and Kroft (2009). The vertical red line represents the true estimate. The empirical CDF can show the implied p-value on the y-axis. As the true estimate is close to the very top of the empirical CDF, where only twice a higher estimate was achieved, the implied p-value is smaller than 0.01. In the top right, we plot the p-values of each permutation against its rank. The vertical red line, showing the estimated p-value in the main specification, is very close to the left and only in 18 of 1,000 permutations the p-value was lower than in our preferred specification. Last, in the top, a density plot of the t-values is shown, again with a vertical red line representing the t-value from the main specification. In 18 cases out of 1,000 permutations a higher t-value was estimated. All plots indicate a true p-value around 0.01.

Figure 8: Placebo (randomly assigned treatment status)

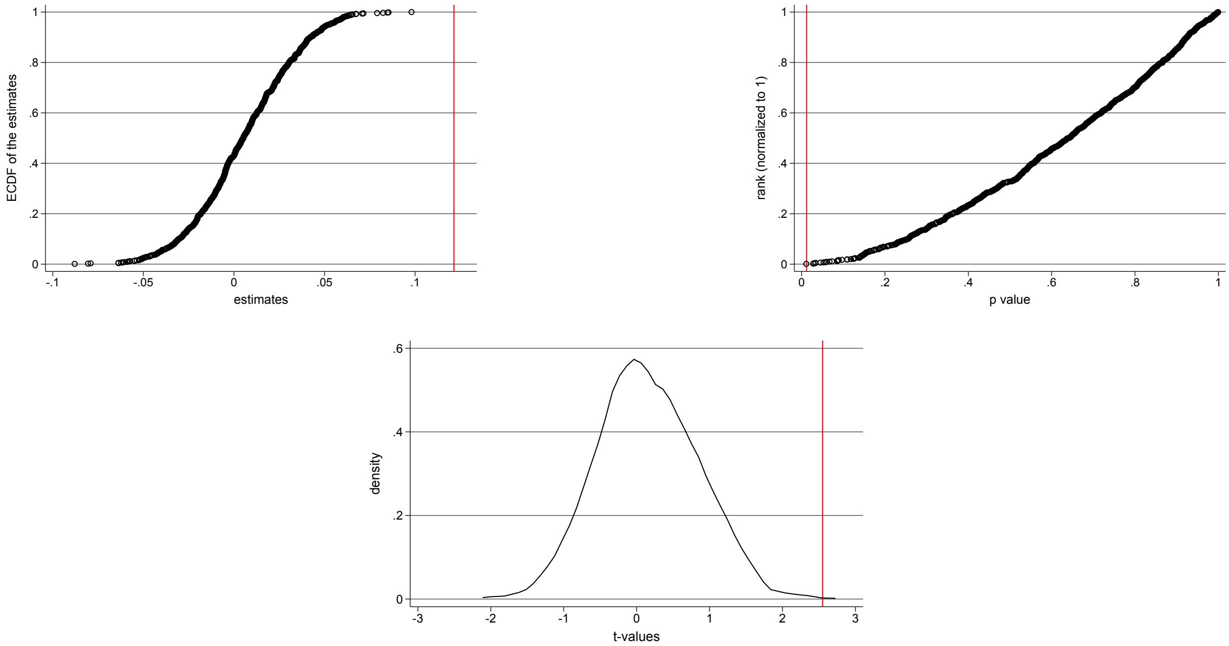


Notes: The figure depicts the distribution of the main effect, estimated with 1000 within country randomly assigned treatment combinations.

Randomly Assigned Connection Years We identify a causal effect under the assumption that the connection year of each country is exogenous. In the event-study plot, it was already shown that the effect only starts after the connection year. As serial correlation might bias the standard errors, we now assign randomly connection years to each country individually (at least one year prior to the actual connection year) as a placebo test. We also shift the rollout of the access points accordingly by the same number of years as the actual connection year was shifted. We re-estimate Equation 1 with these randomly assigned connection years 1,000 times.[41] In only 1 cases out of 1,000 permutations a higher t-value (a lower t-value) was estimated. And all placebo estimates were lower than the estimate from our preferred specification. Thus, all plots indicate a true p-value below 0.01.

[41] Shifting a placebo connection year back in time, we are limited with early observation years. Leaving the estimation window as it is, we do not restrict the towns to have at least one lit pixel in each year from 1995 onwards, but to have at least one lit pixel all years (as for some placebo connection years, the pre-treatment period might start as early as 1992. Thus, for comparison, one has to take Table D.2 (Column 1) and the sample shrinks to 197 towns in the same eleven countries. We also leave out the time-constant controls to avoid singularity matrices.

Figure 9: Placebo (randomly assigned connection year)



Notes: The figure depicts the distribution of the main effect, estimated with 1000 within country randomly assigned connection year combinations.

5.5 Heterogeneity

Heterogeneous effects help to understand the results in two ways. First, they disclose the underlying mechanism behind the average treatment effect. Second, as the access point rollout did not necessarily occurred randomly, they can show how the rollout shaped the effect. To estimate heterogeneous effects, we add a third interaction term to Equation 1:

$$\begin{aligned} y_{c(i)t} =& \gamma_0 + \gamma_1 \left(connection_{ct} \times access_{c(i)} \times x_{c(i)}\right) + \gamma_2 \left(connection_{ct} \times access_{c(i)}\right) + \gamma_3 \left(connection_{ct} \times x_{c(i)}\right) + \\ & \gamma_4 GSM_{c(i)t} + \mathbf{X}'_{c(i)} \tau_t + \lambda_{c(i)} + \lambda_{ct} + \nu_{c(i)t}. \end{aligned} \tag{3}$$

While γ_1 shows how much the main effect changes by a change in $x_{c(i)}$, γ_2 now gives the average main effect. Additionally, we estimate γ_3, showing a potential difference by a change in $x_{c(i)}$ after the connection date.

Table 9 show the regression results for different x. Table B.4 shows all coefficients. First, we investigate the mechanism of trade. Therefore, we insert for x the distance to the next port, to the next river, to the coastline, to the country's border, to the country's capital city, to the next landing point, and a market access measure. Throughout columns (1) towards (8), the main effect remains mostly. However, some interesting heterogeneities can be observed, indicating that trade is an important mechanism. With higher distance to the next port, the main effect diminishes (Column 1). The same holds for the coastline, a broader measure for international trade (Column 3). For the distance to the next river, the negative point estimate lacks statistical significance (Column 2). The same is true for distance to the next border (Column 5). Moreover, Column (4) shows that the effect is driven from coastal countries. Nonetheless, the main effect remains positive (losing economic and statistical significance). Moreover, the distance to a country's capital city shows no statistically significant differences (Column 6). The heterogeneous effect for the landing point is very similar to the effect of the next port (Column 7). Finally, towns that have higher market access profit more from internet access (Column 8). In all cases, after being connected there is no statistically significant differential effect along x. And in most cases, this effect size is lower than the interaction term's estimate.

Table B.5 shows heterogeneity with respect to access to other infrastructure, measured by the distance from a town to the (rail)road network and the electricity grid.To save costs, fiber cables are rolled out along the existing (transportation) infrastructure (cf. subsection 2.2). Due to this nonrandom rollout, distances of these infrastructures was partly correlated with the treatment status of towns. We investigate whether towns that are closer to these nondigital infrastructures grow faster when the countrywide internet connection is established. Column (1) shows a very similar main effect and a positive but statistically insignificant effect for the additional interaction term with distance to the road network. Column (2) shows a slightly higher main effect and a statistically insignificant decline with increasing distance to the railroad network. Last, Column (3) shows a slightly lower main effect. Surprisingly with higher distance to the electricity grid,

Table 9: Heterogeneity

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
connection × access	0.0977**	0.105**	0.0980**	0.0220	0.119***	0.0949**	0.104**	0.132***
	(0.0419)	(0.0442)	(0.0423)	(0.0549)	(0.0451)	(0.0466)	(0.0452)	(0.0426)
connection × access × distance port	-0.116***							
	(0.0416)							
connection × access × distance river		0.0108						
		(0.0397)						
connection × access × distance coastline			-0.120***					
			(0.0429)					
connection × access × coastal country (dummy)				0.269***				
				(0.0855)				
connection × access × distance border					-0.0421			
					(0.0508)			
connection × access × distance capital						-0.0246		
						(0.0541)		
connection × access × distance landing point							-0.0980**	
							(0.0385)	
connection × access × market access								0.172**
								(0.0753)
GSM coverage	0.0827*	0.0767	0.0802*	0.0873*	0.0878*	0.0840*	0.0846*	0.0924*
	(0.0477)	(0.0482)	(0.0479)	(0.0475)	(0.0506)	(0.0479)	(0.0471)	(0.0495)
Observations	2,541	2,541	2,541	2,541	2,541	2,541	2,541	2,541
R-squared	0.949	0.948	0.949	0.949	0.948	0.948	0.948	0.948
#countries	11	11	11	11	11	11	11	11
#towns	231	231	231	231	231	231	231	231
share treated	.459	.459	.459	.459	.459	.459	.459	.459
interaction avg.	0	0	0	.4000000059604645	0	0	0	0
controls	✓	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. Triple-interaction variables are standardized to have mean zero and unitary standard deviation (holds for all variables but the coastal dummy). GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

towns' growth increases faster. We assume that those towns were connected to a better electricity network in the following years and therefore show some increase in light intensity. However, the remaining main effect and the sample restriction to towns with light emission in all years should give enough evidence for the main effect. In all specifications, the main effect remains highly statistically significant.

6 Discussion

Previous estimates of economic growth induced by broadband Internet serve for comparison with our results. We find for cities with Internet access an increase in economic growth of 2 percentage points.

Czernich et al. (2011) investigate GDP growth induced by broadband Internet in OECD countries. They find that the broadband Internet increased GDP per capita by 2.7 to 3.9 percent, implying a .9 to 1.5 annual per capita growth when Internet penetration is increased by 10 percentage points (with penetration ranging between 13.5 percent in Greece and 37.2 percent in Denmark in 2008). Regarding Internet speed and timing, their study is very comparable to ours. They define broadband if a user can surf with at least 256 Kbps. In comparison, Hjort and Poulsen (2019) state that SSA users had on average 430 Kbps before the 'second generation' of SMCs arrived. Most OECD countries introduced broadband Internet between 1999 and 2000 with some late adopters like Greece (2003) and Ireland (2002). In our study, the first countries were connected in 1999 to 2001. However, many countries were connected in the mid-2000s or even later. Two major differences are that we (i) cannot investigate broadband penetration and (ii) compare cities within countries and not across countries. Though, broadband penetration is very low in SSA, it is likely that the very first adopters, mainly firms, have the biggest impact on economic growth. Kolko (2012) investigates broadband Internet expansion in the US and finds, especially in areas with low population density, a positive effect on local economic growth

For SSA, Hjort and Poulsen (2019) estimate a 3.3 percent increase in light activity for the later arrival of fast Internet. First of all, their work differs by the Internet speed available. But most importantly, while we use variation between towns, they use variation within local cells and not across towns. Hence, though in both cases local economic activity is measured, the comparison is different. Finally, the selection of cities and towns differs slightly as we focus on mid-sized towns. All together, it is hard to compare whether the estimates tell something about different speeds or whether they are affected by the named differences. Finally, it cannot be rejected that the effects of the extensive margin, the first connectivity, are still in play when the next generation of SMCs landed. Nevertheless, both studies show that SMCs that brought Internet to SSA at different speeds had both a similar positive effect on local economic growth.

Finally, we want to compare our results to Storeygard (2016) who also estimates local economic growth across cities. Though, not estimating the effects of a digital infrastructure, he is most closely related to our work regarding the outcome measure. Therefore, our estimated effect of Internet availability on a town that is 200 kilometers away from the primate city is equivalent to an oil price shock of 70 US-Dollar.

7 Conclusion

Locations can benefit from the Internet to change to a manufacturing industry if digital infrastructure is in place. We investigate if the availability of Internet at basic speeds fosters economic development in developing countries. In particular, we study the arrival of the first submarine Internet cables in ten Sub-Saharan African countries in the 2000s. To learn about the causal effect of Internet availability on local economic growth, we compare in a difference-in-differences setting economic activity, measured by nighttime light satellite data, of towns connected to the national Internet backbone at the time of countrywide Internet arrival to a control group of similar towns not (yet) connected to the national digital infrastructure but that get an access point later.

We find that the connection of towns to the *World Wide Web*, on average, leads to an increase in light intensity of about 7 percent, relative to similar towns not (yet) connected. This translates into 2 percentage points higher growth in terms of GDP. Moreover, we differentiate the 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 productivity (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 increase their size. Furthermore, our results suggest that this growth is only partly driven by growing populations in connected towns. So, the effect is mainly of an economic development and not a migration effect. Finally, we can show that one mechanism that leads to the growth effects is the change of the industry structure. In regions where Internet access exists, manufacturing has higher growth rates. While the industry shares in employment of manufacturing increase, shares of agriculture decrease.

The rollout of new infrastructure is always expensive. Therefore, policy makers might think of saving money and only rolling out this infrastructure where the effects pay off the costs of the rollout. Our study comes in at this point: We show that even smaller towns that were connected incidentally are growing faster than comparable towns without access points to the Internet. Therefore, first, it is important to account for these smaller towns when evaluating the benefits of an infrastructure. Second, one can derive from our results that the Internet has growth potential not only for economic and political centers but also for smaller towns. Moreover, the effects of the Internet are not bound to a high uptake, but the few adopters generate spillovers. Hence, we recommend to rollout this infrastructure further even when only a low, but positive, uptake is expected. An uptake by some firms might generate external effects for the whole town. Moreover, of course, the Internet might have further effects on educational or political outcomes. Hence, there might be other reasons to connect the whole country which are not targeted in this study, but that could be an interesting direction for further research.

References

Agyeman, O. T. (2007). Survey of ict and education in africa: Benin country report.

Aker, J. C. and I. M. Mbiti (2010). Mobile Phones and Economic Development in Africa. *Journal of Economic Perspective 24*(3), 207–232.

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

Atasoy, H. (2013). The effects of broadband internet expansion on labor market outcomes. *Industrial and Labor Relations Review 66*(2), 315–345.

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.

BenYishay, A., R. Rotberg, J. Wells, Z. Lv, S. Goodman, L. Kovacevic, and D. Runfola (2017). Geocoding afrobarometer rounds 1-6: Methodology & data quality. *AidData. Available online at http://geo. aiddata. org*.

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.

Cariolle, J. (2021). International connectivity and the digital divide in Sub-Saharan Africa. *Information Economics and Policy 55*, 100901.

Chabossou, A. F. (2007). 2007 benin telecommunications sector performance review – a supply side analysis of policy outcomes.

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.

Christiaensen, L. and R. Kanbur (2017). Secondary Towns and Poverty Reduction: Refocusing the Urbanization Agenda. *Annual Review of Resource Economics 9*(November), 405–419.

Christiaensen, L. and Y. Todo (2014). Poverty reduction during the rural-urban transformation - The role of the missing middle. *World Development 63*, 43–58.

Colombo, M. G., A. Croce, and L. Grilli (2013). ICT services and small businesses' productivity gains: An analysis of the adoption of broadband Internet technology. *Information Economics and Policy 25*(3), 171–189.

Czernich, N. (2014). Does broadband internet reduce the unemployment rate? Evidence for Germany. *Information Economics and Policy 29*, 32–45.

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

Dinkelman, T. (2011). The effects of rural electrification on employment: New evidence from South Africa. *American Economic Review 101*(7), 3078–3108.

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review 108*(4-5), 899–934.

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China's national trunk highway system. *Review of Economic Studies 81*(3), 1046–1070.

Fetzer, T., V. Henderson, D. Nigmatulina, and A. Shanghavi (2016). What happens to cities when countries become democratic? (8005077).

Ghani, E., A. G. Goswami, and W. R. Kerr (2016). Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. *Economic Journal 126*(591), 317–357.

Grimes, A., C. Ren, and P. Stevens (2012). The need for speed: Impacts of internet connectivity on firm productivity. *Journal of Productivity Analysis 37*(2), 187–201.

Grogan, L. and A. Sadanand (2013). Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua. *World Development 43*, 252–265.

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.

Hjort, J. and L. Tian (2021). The Economic Impact of Internet Connectivity in Developing Countries . *Working paper*.

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

Jensen, M. (2006). Open access. lowering the costs of international bandwidth in africa. *APC Issue Papers Series October*.

Jensen, R. (2007). The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *The Quarterly Journal of Economics 3*(August), 879–924.

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

Kolko, J. (2012). Broadband and local growth. *Journal of Urban Economics 71*(1), 100–113.

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

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics 94*(1), 20–36.

Poppe, Y. (2009). Subsea cables: Jugular veins for africa's global communications.

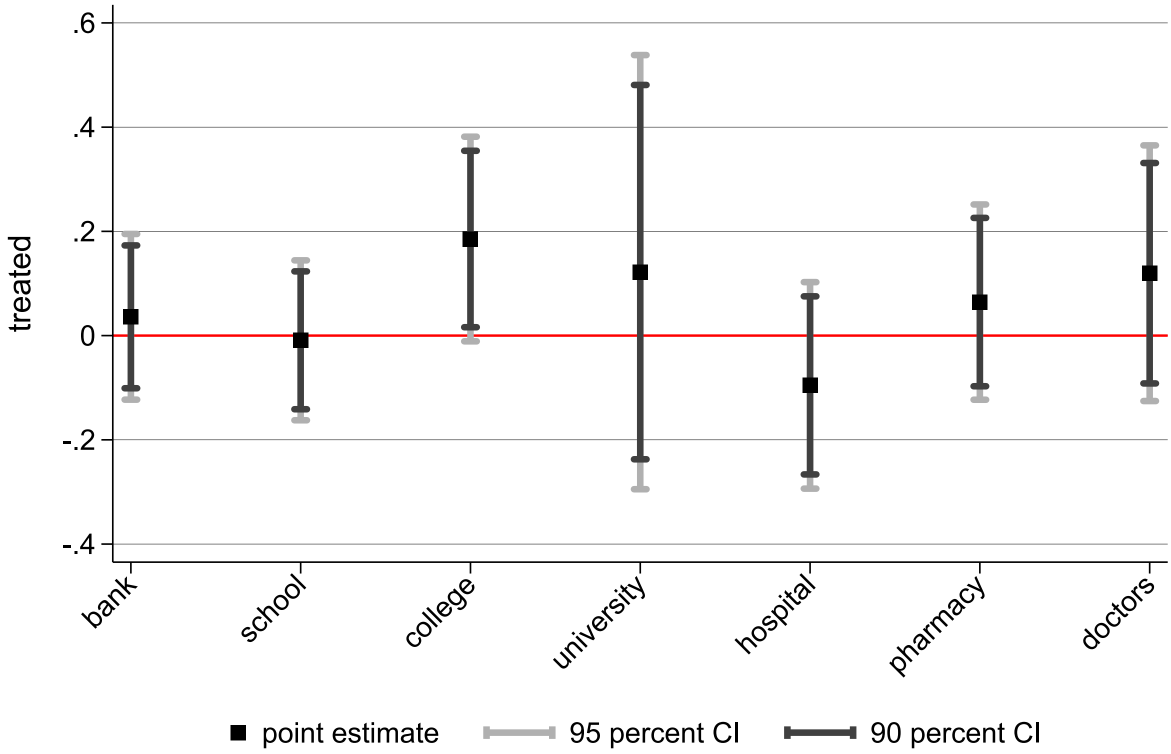
Rud, J. P. (2012). Electricity provision and industrial development: Evidence from India. *Journal of Development Economics 97*(2), 352–367.

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.

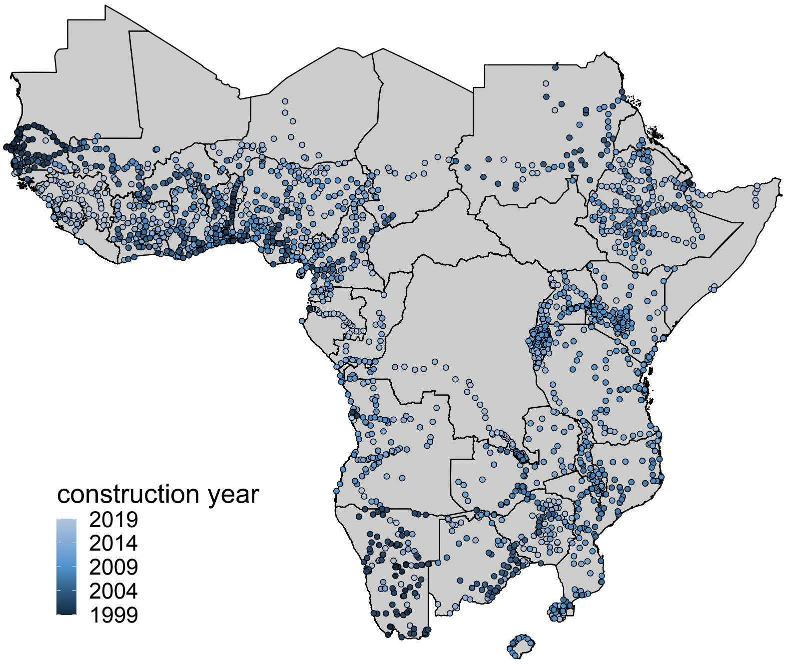
A Figures

Figure A.1: Balance test (OSM)



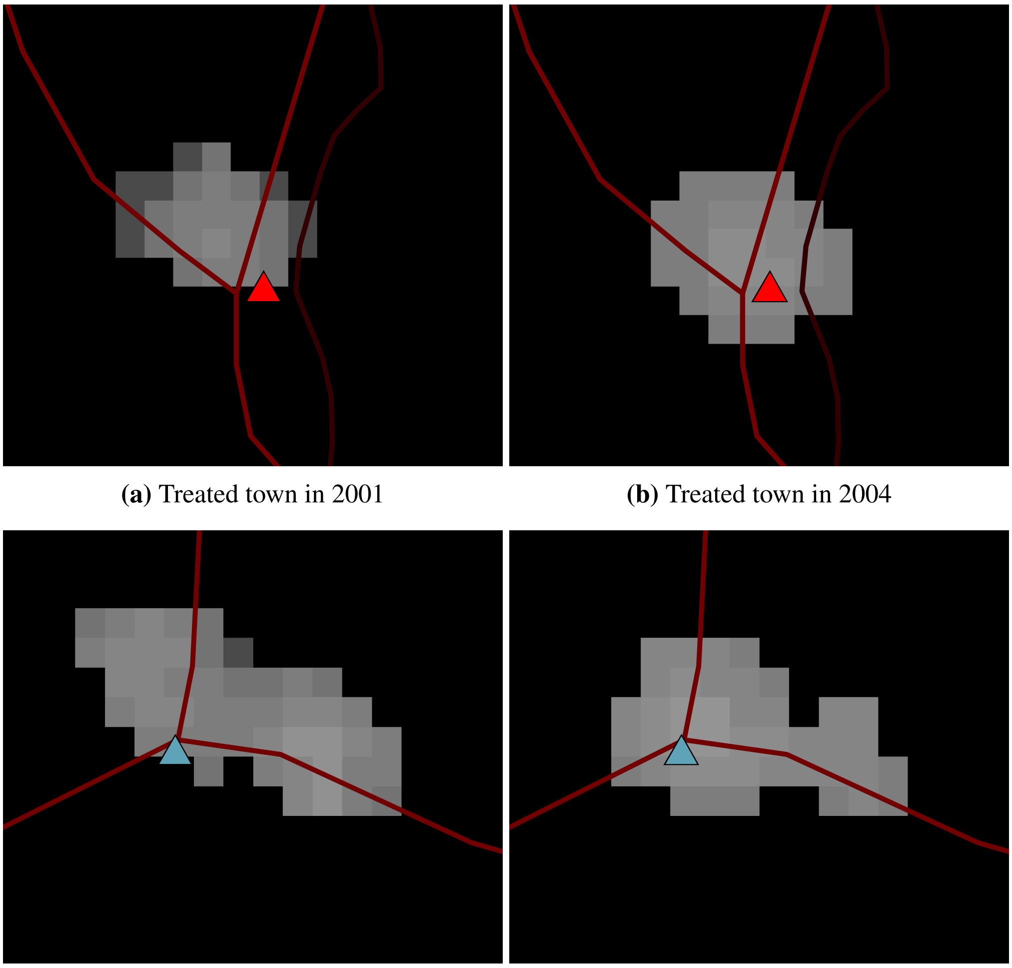
Notes: Confidence intervals reported at the 90% and the 95% level.

Figure A.2: Access points and their construction years



Notes: The figure depicts the location and construction date of all SSA access points. Brighter blue dots correspond to later constructed access points.

Figure A.3: Development of illuminated towns in Benin

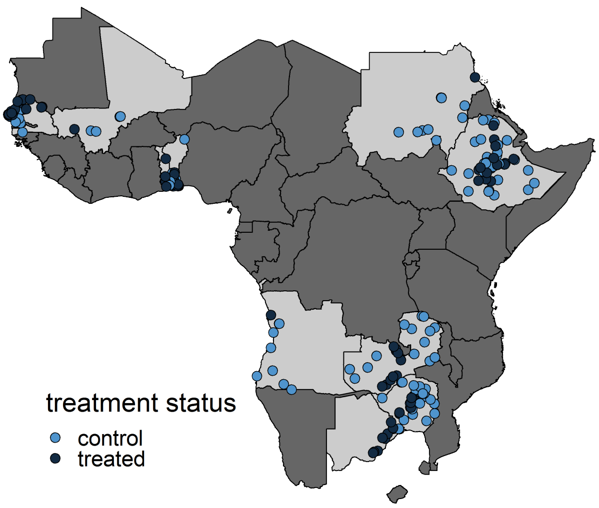


(a) Treated town in 2001 **(b)** Treated town in 2004

(c) Control group town in 2001 **(d)** Control group town in 2004

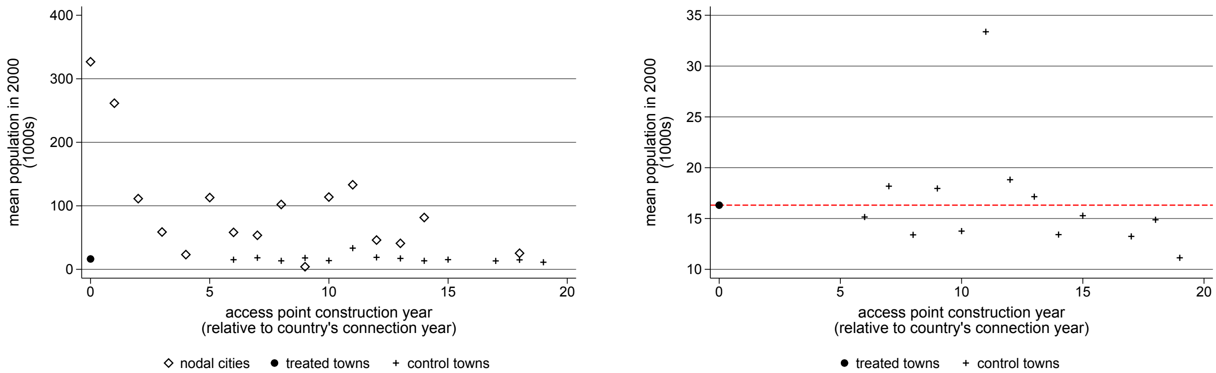
Notes: The panels show a treatment and control group town from Benin, with gray NTLs pixels from 2001 and 2004. Access points are marked with a triangle (red if constructed until 2001 and blue if constructed afterwards). The dark red line represents a major road connecting and the darker red line the railway. The black-to-white scale indicates light intensity, with brighter colors reflecting higher light intensities.

Figure A.4: Countries and towns location in the estimation sample



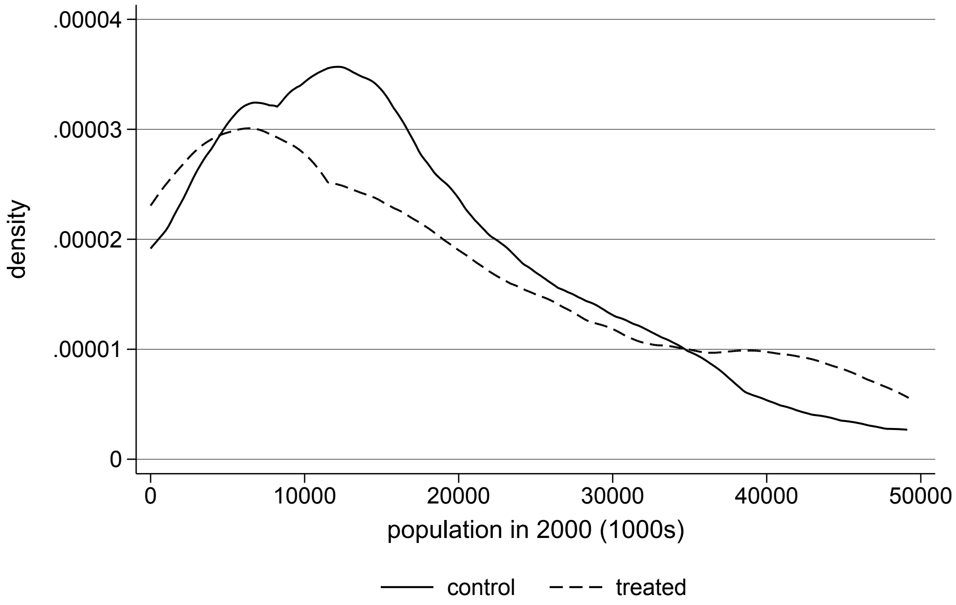
Notes: This figure depicts the countries in our analysis (brighter gray) and for each country the towns in the treatment and control group.

Figure A.5: Population size of connected cities and towns by year (relative to connection year)



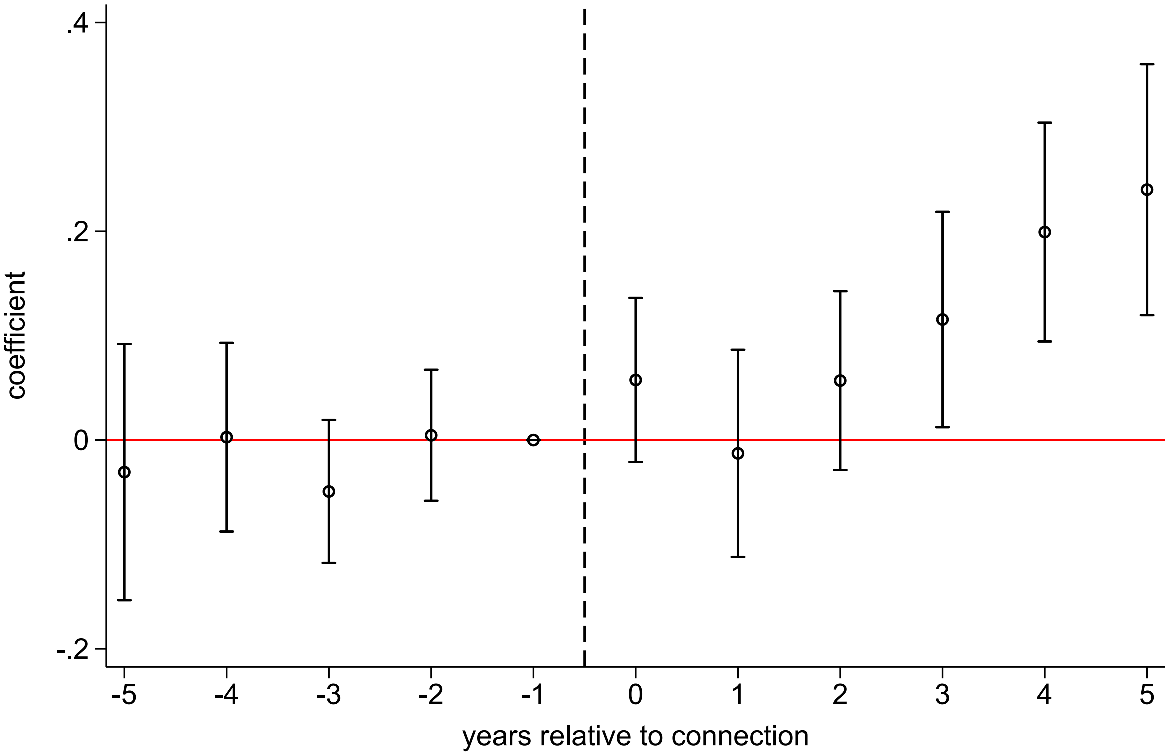
Notes: The figure depicts the average population size of connected cities and towns by year relative to the connection year. On the left, the black dot in the lower left corner represents the treated towns, while the control towns are represented by the plus symbol and the nodal cities by a diamond. For treated towns and nodal cities that were connected in earlier years than the arrival of an SMC are shown in year zero as well for clarity. On the right, the treatment and control group are shown in more detail without nodal cities.

Figure A.6: Distribution of population in 20000 of treated and control group towns



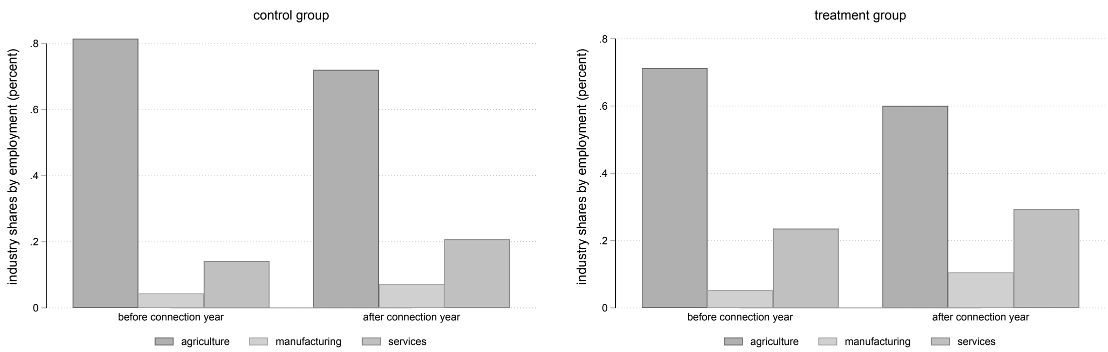
Notes: This figure shows the distribution of population in 20000 of treated and control group towns.

Figure A.7: Event-study coefficients (with 90% level confidence intervals)



Notes: Coefficients for event-study specification of Column (3) of Table 1. Robust standard errors clustered by town. Confidence intervals reported at the 90% level.

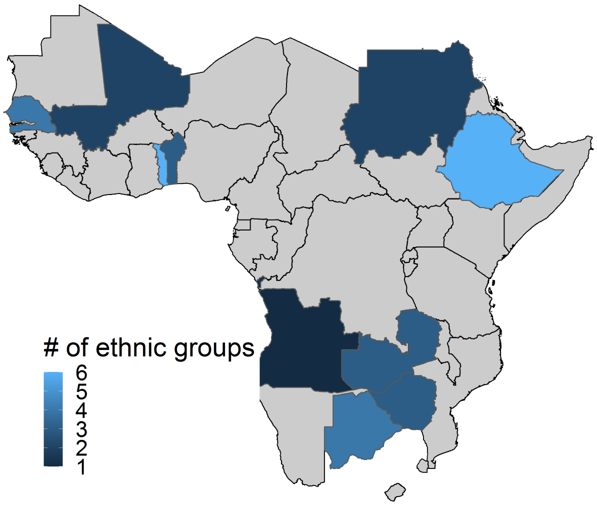
Figure A.8: Comparison of industry shares



(a) **(b)**

Notes: The figures depict the changes in the industry shares before and after the Internet connection for the treatment (1) and control group (2).

Figure A.9: Ethnic groups



Notes: The figure shows for each SSA country in our analysis how many different ethnic groups were provided with at least one access point before the arrival of an SMC. Brighter blue colors indicate more different ethnic groups. Gray indicates countries not included in our analysis.

B Tables

Table B.1: Connection years

Country	Connection year	Connected by	SMC landing point	Upgrade year
Namibia	1999	Neighboring country		2012
Djibouti	1999	Sub-marine cable	Djibouti City	2009
Senegal	2000	Sub-marine cable	Dakar	2010
Angola	2001	Sub-marine cable	Sangano	2012
Benin	2001	Sub-marine cable	Cotonou	2012
Ghana	2001	Sub-marine cable	Accra	2010
Cameroon	2001	Sub-marine cable	Douala	2012
Gabon	2001	Sub-marine cable	Libreville	2012
Nigeria	2001	Sub-marine cable	Lagos	2010
Ivory Coast	2001	Sub-marine cable	Abidjan	2010
Sudan	2003	Sub-marine cable	Port Sudan	2010
Mali	2004	Neighboring country		2010
Botswana	2004	Neighboring country		2009
Zimbabwe	2004	Neighboring country		2011
Burkina Faso	2005	Neighboring country		2010
Togo	2005	Sub-marine cable	Lomé	2012
Gambia	2005	Sub-marine cable	Banjul	2012
Chad	2005	Neighboring country		2012
Central African Republic (CAR)	2005	Neighboring country		2012
Guinea-Bissau	2005	Sub-marine cable	Suro	2012
Mozambique	2006	Sub-marine cable	Maputo	2009
Lesotho	2006	Neighboring country		2010
Niger	2006	Neighboring country		2012
Malawi	2007	Neighboring country		2010
Ethiopia	2007	Neighboring country		2012
Zambia	2007	Neighboring country		2011
Swaziland	2008	Neighboring country		2009

Notes: The table reports the connection years of all SSA countries being connected before 2009. Source: *Submarine Cable Maps* and *Africa Bandwidth Maps*.

Table B.2: Summary statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) p25	(5) p50	(6) p75	(7) max	(8) N
population (in 2015)	36,504.72	27,033.93	10,209.00	17,156.00	28,011.00	46,439.00	205,943.00	229.00
population (in 2000)	15,956.67	13,154.77	0.00	5,398.00	12,772.00	24,239.00	49,217.00	229.00
altitude	874.76	719.08	0.02	60.20	1,016.20	1,372.18	2,816.32	229.00
population density in 2015	4,860.99	4,118.79	710.00	2,639.00	3,982.00	6,029.00	38,637.00	229.00
built-up area in 2015	10.99	12.29	0.35	4.40	7.40	13.47	122.21	229.00
distance to the capital	2.52	2.48	0.02	0.75	1.75	3.59	12.54	229.00
distance to the coastline	3.70	2.89	0.00	0.94	3.84	5.47	11.57	229.00
distance to next river	0.57	0.52	0.00	0.15	0.47	0.91	3.36	229.00
distance to the SMC landing point	5.64	3.85	0.01	1.74	6.02	9.02	14.51	229.00
distance to the road network	0.03	0.12	0.00	0.00	0.00	0.00	1.13	229.00
distance to the railroad network	0.58	0.93	0.00	0.00	0.07	0.85	4.40	229.00
distance to the border	1.20	1.21	0.00	0.20	0.85	1.77	5.02	229.00
distance to next port	4.02	2.90	0.00	1.35	4.29	5.87	11.96	229.00
distance to the electricity grid	0.12	0.32	0.00	0.00	0.00	0.05	2.25	229.00
number of lit pixels (in t-1)	47.31	29.97	3.00	29.00	39.00	59.00	232.00	229.00
light intensity (in t-1)	505.41	601.41	12.00	174.00	308.00	585.00	4,842.00	229.00
average light intensity (in t-1, including zeros)	7.05	5.99	0.29	3.10	4.82	8.91	32.72	229.00
mobile coverage (in t-1, GSM)	0.59	0.48	0.00	0.00	1.00	1.00	1.00	229.00
distance to next AP (in 2020)	1.28	2.58	0.00	0.00	0.00	1.13	9.43	229.00
market access	14804589.90	107781232.19	119.00	1,256.00	4,987.00	12,922.00	988349824.00	229.00
terrain ruggedness (log)	10.60	1.63	0.00	9.80	10.84	11.63	13.36	229.00

Notes: The table reports summary statistics of the estimation sample.

Table B.3: Robustness: Excluding single countries

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
connection × access	0.109**	0.0755*	0.0667	0.114**	0.173***	0.109**	0.123**	0.0926**	0.0891*	0.108**	0.137***	0.0754
	(0.0453)	(0.0428)	(0.0514)	(0.0460)	(0.0517)	(0.0476)	(0.0474)	(0.0450)	(0.0476)	(0.0459)	(0.0448)	(0.0511)
GSM coverage	0.0864*	0.0510	0.0416	0.0898*	0.0920*	0.0949*	0.0847	0.106**	0.0634	0.0893*	0.102*	0.108*
	(0.0494)	(0.0505)	(0.0502)	(0.0496)	(0.0510)	(0.0516)	(0.0514)	(0.0507)	(0.0595)	(0.0505)	(0.0568)	(0.0573)
Observations	2,541	2,354	2,266	2,431	2,090	2,442	2,420	2,453	2,233	2,497	2,277	1,947
R-squared	0.948	0.948	0.956	0.945	0.954	0.949	0.948	0.950	0.950	0.948	0.948	0.941
#countries	11	10	10	10	10	10	10	10	10	10	10	10
#towns	231	214	206	221	190	222	220	223	203	227	207	177
share treated	.459	.491	.451	.439	.463	.464	.455	.471	.409	.454	.449	.508
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
country ✓year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o country	none	ao	bj	bw	et	mw	mz	sd	sn	tg	zm	zw

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.4: Heterogeneity (full interaction terms output)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
connection × access	0.0977**	0.105**	0.0980**	0.0220	0.119***	0.0949**	0.104**	0.132***
	(0.0419)	(0.0442)	(0.0423)	(0.0549)	(0.0451)	(0.0466)	(0.0452)	(0.0426)
connection × access × distance port	-0.116***							
	(0.0416)							
connection × access × distance river		0.0108						
		(0.0397)						
connection × access × distance coastline			-0.120***					
			(0.0429)					
connection × access × coastal country (dummy)				0.269***				
				(0.0855)				
connection × access × distance border					-0.0421			
					(0.0508)			
connection × access × distance capital						-0.0246		
						(0.0541)		
connection × access × distance landing point							-0.0980**	
							(0.0385)	
connection × access × market access								0.172**
								(0.0753)
connection × distance port	-0.00988							
	(0.0535)							
connection × distance river		0.00129						
		(0.0308)						
connection × distance coastline			-0.0232					
			(0.0497)					
connection × distance border					-0.0256			
					(0.0401)			
connection × distance capital						-0.0386		
						(0.0317)		
connection × distance landing point							0.0416	
							(0.0551)	
connection × market access								-0.0128
								(0.0108)
GSM coverage	0.0827*	0.0767	0.0802*	0.0873*	0.0878*	0.0840*	0.0846*	0.0924*
	(0.0477)	(0.0482)	(0.0479)	(0.0475)	(0.0506)	(0.0479)	(0.0471)	(0.0495)
Observations	2,541	2,541	2,541	2,541	2,541	2,541	2,541	2,541
R-squared	0.949	0.948	0.949	0.949	0.948	0.948	0.948	0.948
#countries	11	11	11	11	11	11	11	11
#towns	231	231	231	231	231	231	231	231
share treated	.459	.459	.459	.459	.459	.459	.459	.459
controls	✓	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. Triple-interaction variables are standardized to have mean zero and unitary standard deviation (holds for all variables but the coastal dummy). GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.5: Heterogeneity (Infrastructure)

VARIABLES	(1)	(2)	(3)
connection × access	0.115**	0.107**	0.115***
	(0.0490)	(0.0440)	(0.0436)
connection × access × distance roads	0.0306		
	(0.120)		
connection × access × distance railroads		-0.0224	
		(0.0301)	
connection × access × distance electricity grid			0.0765
			(0.0492)
connection × distance roads	0.0255		
	(0.0325)		
connection × distance railroads		0.0224	
		(0.0554)	
connection × distance electricity grid			-0.000726
			(0.0309)
GSM coverage	0.0859*	0.0880*	0.0803
	(0.0486)	(0.0495)	(0.0500)
Observations	2,541	2,541	2,541
R-squared	0.948	0.948	0.949
#countries	11	11	11
#towns	231	231	231
share treated	.459	.459	.459
interaction avg.	0	0	0
controls	✓	✓	✓
town FE	✓	✓	✓
country × year FE	✓	✓	✓
w/o nodal cities	✓	✓	✓

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. Triple-interaction variables are standardized to have mean zero and unitary standard deviation. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C Example: Case of Benin

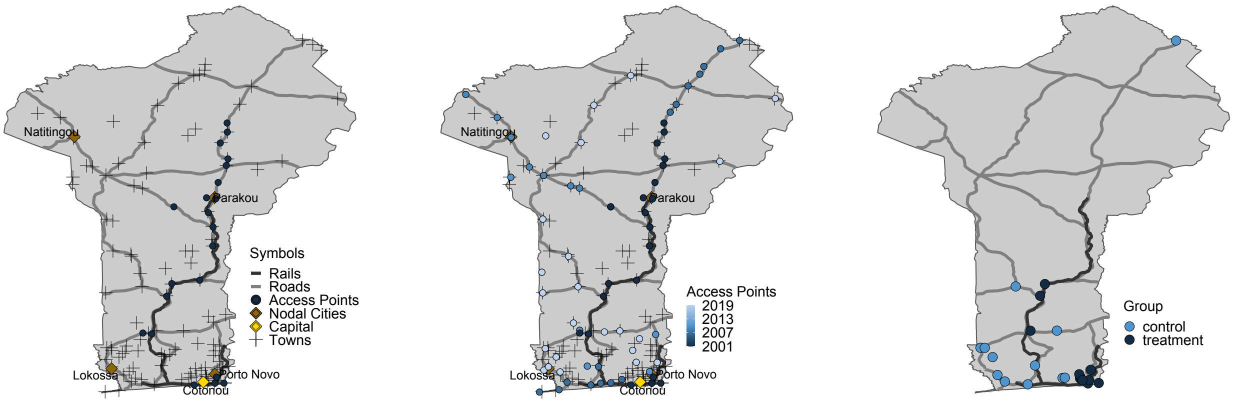
Benin is a good example for how the national backbone was rolled out and how it influenced Internet usage. It is one of the countries that was connected by the SAT-3 SMC, which brought an international connection of 45 Mbps (Chabossou, 2007). The rollout of the national backbone was planned by Benin Telecoms SA, the fixed-line monopolist which manages the gateway to the national Internet, operates as the national carrier, and administers the national domain (.bj). Benin Telecoms SA is state-owned and offers permanent ADSL connections with up to 2 Mbps (Agyeman, 2007).

Infrastructure Rollout Following Chabossou (2007), the SAT-3 SMC landed in Cotonou, Benin's biggest city, the location of the seat of government, and 40 kilometers away from Benin's capital, Porto-Novo. Close by, in Abomey-Calavi Benin's hub is located as well. These cities form with Godomey Benin's largest agglomeration with nearly 2.5 million inhabitants (about a third of Benin's population). From there, first, a connection to Parakou with a 425 kilometers optical fiber cable was constructed in 2001. Parakou is Benin's next largest economic center with more than 150,000 inhabitants in the 2002 census and the capital of the Borgou department. This connection was constructed along Benin's railway line and roads network (Figure C.1) and connected further smaller towns on its way, e.g., Savalou with 30,000 inhabitants. Next, from Parakou connections to the borders to Niger, in the north-east, and Burkina Faso, in the north-west, were constructed along the road network, transforming Benin to a sub-regional digital hub interconnecting Togo, Nigeria, Burkina Faso, and Niger. The first kilometers of the fiber-optic backbone and access points were still constructed until 2001. Consequently, Benin Telecoms SA investment in the telecommunications sector peaked in 2001 with more than 80 billion US-Dollar. The connection towards Burkina Faso and Togo was constructed through Natitingou, the capital of the Atakora department. Again, connecting also further smaller towns, such as Kandi or Djougou, incidentally. Only later on, further rural towns were connected when constructing backbone circles to make the network more reliable, e.g., Nikki, Ségbana, and Banikoara.

Internet Usage All transmission happens via Benin Telecoms SA. They offer data transmission networks to mostly commercial clients (banks, hotels, ministries, etc.) in packets.[42] Having grown exponentially, thousands of cybercafés offer Internet access. While international institutions, major corporations, service providers, and some cybercafés have permanent links, home access remains very limited (Chabossou, 2007). Still, in 2007 only 25 percent of people in Benin's population have used the Internet at least one time. Access is mainly at cybercafés (21 percent) or at the workplace (2.2 percent) while Internet at home remains a luxury. Though, workplace Internet usage is low, it indicates that firms are great adopters of broadband

[42] Network interconnectivity enables new providers to use the incumbent's infrastructure instead of having to invest greatly to build an own one, which incentivizes competitive adaptation. There are, in addition to the former monopolist, which still owns the infrastructure, three licensed providers. However, there are about 50 providers operating without a licence and there is no adequate framework for regulation.

Figure C.1: Rollout in Benin



(a) Access points in 2001 **(b)** All access points **(c)** Treatment/control group

Notes: The figure outlines the rollout of access points in Benin. Besides access points, the maps include the capital city, nodal cities, and all towns. Railroads and roads are included as well. In the left panel, the early rollout with access points being constructed until the arrival of the SMC in 2001 is shown. The middle panel depicts further access points and their respective construction years. The right panel shows the towns of your analysis divided into treatment and control group.

Internet. Among the groups of higher education, Internet usage is also a lot higher. Therefore, we expect local growth through firm's productivity to increase induced by broadband Internet.

D Estimation Sample

We focus on early SMCs bringing Internet connections at basic speeds to SSA in the early 2000s. Therefore, we do not consider countries which were connected after 2008 for the first time, when the next generation of SMCs (which allowed for much higher speeds) landed. This leaves 27 countries, which are listed in Table B.1. Among the first countries that were connected are Djibouti, where an SMC landed in 1999, Namibia, which was connected by a trans-national fiber cable from South Africa in 1999, and Senegal, where an SMC landed in 2000.[43] In 2001, nine more countries were connected by a single SMC, the SAT-3 cable. In the following years until 2008, 17 more countries got an SMC connection or were connected through a neighboring country.

However, not all countries that were connected until 2008 had constructed a national backbone infrastructure before 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 access right after the connection. This reduces the number of countries in our analysis to 23.[44] Moreover, eleven countries established only in nodal cities access points before Internet became available countrywide.[45] Therefore, there are no towns in the treatment group and we cannot estimate on these countries. Finally, we cannot consider Namibia in our analysis because it did not construct further access points after getting the Internet connection. Therefore, we are unable to define a control group. This leaves 12 countries for our analysis.

Due to the staggered arrival of SMCs, this sample represents an unbalanced panel. In our main specification, we take a conservative approach and estimate on a balanced panel. Therefore, we truncate the data to attain a balanced panel. Malawi and Mozambique only have two post-treatment years. They were connected in 2007 and 2006, respectively, and got upgraded by an SMC with more capacity in 2010 and 2009, respectively. Thus, only three years lie between the first connection to the Internet and the Internet capacity upgrade for both countries. Hence, estimating on a balanced sample with three post-treatment years leaves us with a sample of ten countries.

[43] Djibouti and Senegal were connected as single SSA countries through bigger international multi-country SMCs. Djibouti was connected with SeaMeWe-3, which connected Northern and Western Europe with Eastern Asia and Australia. Senegal was connected with Atlantis-2, which went from Spain and Portugal through the Canary Islands to Brazil and Argentina and landed on Sengal's shores on the way.

[44] Central African Republic has not yet constructed a national backbone infrastructure. In Lesotho, the access points were established in 2009 three years after being connected through South Africa. In Djibouti, the first access points were established in 2007, which is eight years after the first SMC connection. Nigeria established its first access points in 2003, which is two years after the arrival of the first SMC.

[45] Guinea-Bissau, Lesotho, and Swaziland established all access points until today only in nodal cities.

D.1 Sample: Satellite

In the main specification, the towns were restricted to have a light emission in all years after 1994 (the earliest year in the sample). In the following, we relax this restriction and either ignore missing light emission or impute a missing. In Table D.1, we allow the sample to have missing light emission in a range from 0 to 6 years (at any point in time). In Columns (1) through (7) there is no other restriction. In Columns (8) through (14), we apply the same restriction to the early years as in the main specification. Thus, Column (8) shows again the main specification (Table 1) Allowing for more missing light emissions increases the sample (by up to more than 60 towns) and also the number of included countries (to up to 15). In general, the results remain very robust, with a slightly smaller and slightly less precisely measured point estimate.

Table D.1: Ignoring missing pixels

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
connection × access	0.0959*	0.109**	0.0932**	0.0734	0.0940**	0.109**	0.0979**	0.0986**	0.0881*	0.0736
	(0.0498)	(0.0460)	(0.0442)	(0.0459)	(0.0473)	(0.0453)	(0.0459)	(0.0474)	(0.0477)	(0.0490)
GSM coverage	0.159***	0.0928**	0.0965**	0.0853*	0.0939**	0.0864*	0.0842*	0.0901*	0.0723	0.0610
	(0.0542)	(0.0461)	(0.0465)	(0.0472)	(0.0469)	(0.0494)	(0.0483)	(0.0471)	(0.0510)	(0.0492)
Observations	2,189	2,536	2,704	2,873	2,997	2,541	2,866	3,000	3,072	3,237
R-squared	0.953	0.948	0.947	0.942	0.936	0.948	0.940	0.937	0.934	0.931
#countries	11	11	12	12	13	11	12	13	13	14
#towns	199	231	247	261	269	231	260	270	276	288
share treated	.467	.446	.453	.444	.439	.459	.438	.433	.428	.438
town FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
#missing values ignored	0	1	2	3	4	0	1	2	3	4
controls						✓	✓	✓	✓	✓
early-year restriction						✓	✓	✓	✓	✓

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. Missing values, in up to four years, are ignored. Columns (1) through (5) are not further restricted, while Columns (6) through (10) are limited towards lit towns in early years as in the main specification. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

When imputing missing values, we only impute values with the two neighboring values are not missing. We impute if a town has up to 4 missing values. Again, Columns (1) through (5) do not have further restrictions, and Columns (6) through (10) show the same restriction for early years as the main specification. Column (6) repeats the preferred specification. Again, more imputed values lead to a larger sample with more towns (to up to 267) and more countries (to up to 13). In general, the results remain very robust, with a slightly smaller and slightly less precisely measured point estimate.

Table D.2: Imputation of missing pixels

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
connection × access	0.0959*	0.0904*	0.0845*	0.0845*	0.0813*	0.109**	0.0920**	0.0944**	0.0922*	0.0922*
	(0.0498)	(0.0493)	(0.0500)	(0.0500)	(0.0489)	(0.0453)	(0.0462)	(0.0475)	(0.0472)	(0.0472)
GSM coverage	0.159***	0.154***	0.160***	0.160***	0.158***	0.0864*	0.0821*	0.0855*	0.0914*	0.0914*
	(0.0542)	(0.0489)	(0.0484)	(0.0484)	(0.0484)	(0.0494)	(0.0493)	(0.0475)	(0.0476)	(0.0476)
Observations	2,189	2,266	2,299	2,299	2,310	2,541	2,860	2,948	2,959	2,959
R-squared	0.953	0.954	0.952	0.952	0.952	0.948	0.941	0.938	0.938	0.938
#countries	11	11	11	11	11	11	12	13	13	13
#towns	199	206	209	209	210	231	260	268	269	269
share treated	.467	.456	.459	.459	.457	.459	.442	.444	.442	.442
town FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
#missing values imputed	0	1	2	3	4	0	1	2	3	4
controls						✓	✓	✓	✓	✓
early-year restriction						✓	✓	✓	✓	✓

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. Missing values between two years with existing values are imputed by the average of the two surrounding years. Columns (1) through (5) are not further restricted, while Columns (6) through (10) are limited towards lit towns in early years as in the main specification. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

E Afrobarometer: Electricity

For already named reasons, the rollout of the electricity grid might be a thread to isolate the effect of Internet availability. We therefore analyze whether households get connected to the electricity grid with survey data from *Afrobarometer* (BenYishay et al., 2017).[46] We take the first four rounds, if they are before the SMC upgrade, from 1999 to 2009 to maximize the number of countries. We take averages for each town and generate weights for the number of surveyed individuals per town. In some countries, only very few towns are visited besides of nodal cities. In Table E.1, we first include all cities (Columns (1) and (2)) and restrict the sample stepwise, dropping first all capitals (Columns (3) and (4)), and in a second step, all nodal cities (Columns (5) and (6)). In odd columns, no weights are used, while in even columns, the number of surveyed individuals per town is used as a weight. All specifications lack statistical significance. In the overall sample (Columns (1) and (2)), the point estimate is very close to zero. It increases slightly for the sample without capital cities (Columns (3) and (4)) and becomes negative for the most restrictive sample (Columns (5) and (6)). We take from these estimations that for some nodal cities (without capital cities) the rollout of the electricity grid advanced at the same time as the Internet was rolled out. However, for smaller towns this cannot be confirmed. We can therefore rule out the concern that the electricity grid confounds our results.

Table E.1: Electricity grid (from *Afrobarometer*)

electricity grid	(1)	(2)	(3)	(4)	(5)	(6)
post x treated	0.000387	-0.0359	0.0411	0.0579	-0.0731	-0.0914
	(0.103)	(0.0688)	(0.114)	(0.0766)	(0.211)	(0.173)
GSM coverage	0.0623	0.0205	0.0580	0.00348	0.107	-0.00385
	(0.111)	(0.0901)	(0.115)	(0.106)	(0.171)	(0.158)
observations	270	270	250	250	102	102
R-squared	0.680	0.806	0.675	0.784	0.720	0.814
#countries	6	6	6	6	4	4
#towns	94	94	88	88	37	37
share treated	.351	.351	.307	.307	.351	.351
town FE	✓	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓	✓
weights		✓		✓		✓
w/o capital+landing point			✓	✓		
w/o nodal cities					✓	✓

Notes: Access to the electricity grid was aggregated at the town/city level and comes from Afrobarometer (rounds 1 to 4). GSM mobile coverage is calculated as the percentage share of town area covered with signal. All specifications are estimated on a sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 50,000 inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

[46] https://afrobarometer.org/

F Further Robustness Checks

Definition of Nodal Cities For the main results, we defined all towns with more than 50,000 inhabitants as nodal cities. However, this threshold is chosen arbitrarily. Therefore, when estimating the specification in Column (3) of Table 1, we vary the population threshold as a further robustness check. Tables F.1 and F.2 show that the estimate remains independently of the chosen population threshold (varying from 20,000 to 1,000,000 inhabitants). In both tables, Column (4) repeats the main specification. While Table F.1 is restricted to the sample of the main specification and therefore containing less countries and towns, Table F.2 is not restricted in that way.

Table F.1: Population threshold for nodal cities (restricted to sample of main specification)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
connection × access	0.0919*	0.0961**	0.0985**	0.109**	0.105**	0.0993**	0.0961**	0.0981**	0.0957**
	(0.0516)	(0.0452)	(0.0448)	(0.0453)	(0.0422)	(0.0423)	(0.0420)	(0.0418)	(0.0410)
GSM coverage	0.0722	0.0686	0.0646	0.0864*	0.0966**	0.0964**	0.103**	0.102**	0.102**
	(0.0570)	(0.0479)	(0.0433)	(0.0494)	(0.0482)	(0.0481)	(0.0472)	(0.0486)	(0.0483)
Observations	1,683	2,112	2,387	2,541	2,706	2,739	2,827	2,871	2,904
R-squared	0.944	0.939	0.947	0.948	0.953	0.956	0.960	0.963	0.967
#countries	9	11	11	11	11	11	11	11	11
#towns	153	192	217	231	246	249	257	261	264
share treated	.458	.448	.442	.459	.467	.474	.482	.49	.496
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓	✓
population threshold	20000	30000	40000	50000	75000	100000	250000	500000	1000000

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. The population threshold relevant for being included in the estimation sample is varied. Nodal cities in this specification include the landing point and the capital city, regional capitals. All specifications include town and country-year fixed effects and are estimated on a balanced panel. The sample is restricted to the countries of the main specification. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Definition of Internet Access For the main results, Internet access was defined for towns with an access point to the national backbone within 10 kilometers as within this distance internet should be accessible. However, this threshold is not sharp with respect to internet access. Therefore, when estimating the specification in Column (3) of Table 1, we vary this distance from 0 to 30 kilometers in steps of 5 kilometers as a further robustness check. For very low distances, internet access might be higher over the whole town's area. For very high distances, internet access can still be provided with an additional fiber-cable rollout. This rollout is not in place in all towns and cannot be observed with our data. Tables F.3 and F.4 show that the estimate remains independently of the chosen distance to the backbone to define internet access. In both tables, Column (3) repeats the main specification. While Table F.3 is restricted to the sample of the main specification and therefore containing less countries and towns, Table F.4 is not restricted in that way. Moreover, one should note that the distance to the access point influences the sample. When allowing for

Table F.2: Population threshold for nodal cities

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
connection × access	0.0919*	0.0961**	0.0985**	0.109**	0.105**	0.0993**	0.0961**	0.0989**	0.0964**
	(0.0516)	(0.0452)	(0.0448)	(0.0453)	(0.0422)	(0.0423)	(0.0420)	(0.0416)	(0.0408)
GSM coverage	0.0722	0.0686	0.0646	0.0864*	0.0966**	0.0964**	0.103**	0.103**	0.103**
	(0.0570)	(0.0479)	(0.0433)	(0.0494)	(0.0482)	(0.0481)	(0.0472)	(0.0487)	(0.0484)
Observations	1,683	2,112	2,387	2,541	2,706	2,739	2,827	2,893	2,926
R-squared	0.944	0.939	0.947	0.948	0.953	0.956	0.960	0.963	0.968
#countries	9	11	11	11	11	11	11	12	12
#towns	153	192	217	231	246	249	257	263	266
share treated	.458	.448	.442	.459	.467	.474	.482	.49	.496
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓	✓	✓
population threshold	20000	30000	40000	50000	75000	100000	250000	500000	1000000

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. The population threshold relevant for being included in the estimation sample is varied. Nodal cities in this specification include the landing point and the capital city, regional capitals. All specifications include town and country-year fixed effects and are estimated on a balanced panel. The sample is *not* restricted to the countries of the main specification. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

higher distances to define internet access, the control group shrinks. It is important that the treatment group contains only towns that can use the Internet, while in the control group Internet should not be accessibly. First, for low distances, the latter might not hold anymore. Hence, the ATT compares a treatment group with a control group which contains actually treated towns. Second, for high distances, the treatment group might contain some towns without actual internet access. At the same time, the control group shrinks in size as only very few towns remain that did not have an access point in a certain higher distance at the beginning. These towns might also be less developed and less growing because of their unfortunate location.

Definition of Control Group A further concern might be that towns being connected through an access point which was constructed many years after the first internet connection are not comparable to the treated towns which were connected through an access point constructed before the first internet connection. However, Table F.5 shows that when restricting the year when control towns were connected does not have a strong impact on the estimate. In contrast to the a priori concern, economic and statistical significance increase when only including towns that were connected shortly after a countrywide internet connection was established to the control group.

Mobile Coverage Lags When controlling for mobile coverage, we therefore control for the difference in having a different ICT infrastructure available. As it might take some time for an infrastructure to affect economic outcomes as we have seen for Internet availability, we also include different lags for mobile coverage instead of current mobile coverage. Table F.6 shows that the main effect remains robust in all lag

Table F.3: Distance threshold to the backbone for nodal cities (restricted to sample of main specification)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
connection × access	0.0975**	0.109**	0.0820*	0.0861*	0.0816*	0.0975**	0.0967**
	(0.0460)	(0.0453)	(0.0452)	(0.0452)	(0.0453)	(0.0456)	(0.0462)
GSM coverage	0.0699*	0.0864*	0.0628	0.0682	0.0486	0.0487	0.0407
	(0.0411)	(0.0494)	(0.0491)	(0.0479)	(0.0480)	(0.0467)	(0.0446)
Observations	2,387	2,541	2,662	2,673	2,838	2,904	3,025
R-squared	0.950	0.948	0.949	0.949	0.947	0.948	0.945
#countries	11	11	11	11	11	11	11
#towns	217	231	242	243	258	264	275
share treated	.419	.459	.488	.506	.523	.58	.6
controls	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓
backbone distance	7.5km	10km	12.5km	15km	20km	25km	30km

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within a certain distance. This distance is varied in these specifications. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. The sample is restricted to the countries of the main specification. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table F.4: Distance threshold to the backbone for nodal cities

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
connection × access	0.0975**	0.109**	0.0820*	0.0861*	0.0732	0.0730*	0.0800*
	(0.0460)	(0.0453)	(0.0452)	(0.0452)	(0.0452)	(0.0424)	(0.0462)
GSM coverage	0.0699*	0.0864*	0.0628	0.0682	-0.0407	-0.000472	0.00761
	(0.0411)	(0.0494)	(0.0491)	(0.0479)	(0.0389)	(0.0289)	(0.0283)
Observations	2,387	2,541	2,662	2,673	3,608	4,499	4,587
R-squared	0.950	0.948	0.949	0.949	0.934	0.930	0.928
#countries	11	11	11	11	13	15	15
#towns	217	231	242	243	328	409	417
share treated	.419	.459	.488	.506	.418	.386	.408
controls	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓
backbone distance	7.5km	10km	12.5km	15km	20km	25km	30km

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within a certain distance. This distance is varied in these specifications. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. The sample is *not* restricted to the countries of the main specification. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table F.5: Robustness (connected control towns)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
connection $\times$ access	0.109**	0.0879*	0.150**	0.146**	0.109**	0.0613	0.122**
	(0.0453)	(0.0503)	(0.0577)	(0.0647)	(0.0453)	(0.0487)	(0.0558)
GSM coverage	0.0864*	0.0968*	0.0515	0.0747	0.0864*	0.0208	0.0960
	(0.0494)	(0.0566)	(0.0805)	(0.0906)	(0.0494)	(0.0623)	(0.0914)
Observations	2,541	2,101	1,496	1,177	2,541	2,079	1,320
R-squared	0.948	0.948	0.956	0.960	0.948	0.953	0.956
#countries	11	9	8	6	11	10	8
#towns	231	191	136	107	231	189	120
share treated	.459	.492	.522	.467	.459	.439	.592
controls	✓	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓	✓
country $\times$ year FE	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓
backbone before	2020	2018	2016	2014			
backbone border					20	14	8

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. Lately constructed access points are removed from the control group: In Columns (1) through (4) by the construction year, in Columns (5) through (7) by relative years since country's connection year. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

specifications. The strongest effect of mobile coverage on economic growth is estimated with a lag of one year. Afterwards, the point estimate shrinks and also loses statistical significance.

Table F.6: Robustness (mobile coverage lags)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
connection × access	0.109**	0.112**	0.104**	0.108**	0.105**	0.102**
	(0.0453)	(0.0443)	(0.0444)	(0.0438)	(0.0438)	(0.0439)
GSM coverage	0.0864*					
	(0.0494)					
GSM coverage (lag 1)		0.100**				
		(0.0393)				
GSM coverage (lag 2)			0.0212			
			(0.0406)			
GSM coverage (lag 3)				0.0546		
				(0.0400)		
GSM coverage (lag 4)					0.0350	
					(0.0404)	
GSM coverage (lag 5)						0.0282
						(0.0364)
Observations	2,541	2,541	2,541	2,541	2,541	2,541
R-squared	0.948	0.949	0.948	0.948	0.948	0.948
#countries	11	11	11	11	11	11
#towns	231	231	231	231	231	231
share treated	.459	.459	.459	.459	.459	.459
controls	✓	✓	✓	✓	✓	✓
town FE	✓	✓	✓	✓	✓	✓
country × year FE	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓

Notes: Outcome: light intensity, measured as the logarithmic sum of light intensities of DMSP-OLS pixels within the town area, coming from *Africapolis*, and a 2 kilometers buffer. Countrywide internet availability is shown by the connection dummy, while access refers to the town level and having an access point within 10 kilometers. GSM coverage and all control variables are defined in the notes of Table 1. Nodal cities include the landing point and the capital city, regional capitals, and cities with more than 50,000 inhabitants. All specifications include town and country-year fixed effects and are estimated on a balanced panel. Robust standard errors clustered at the level of the closest access point and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.