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

Moritz Goldbeck[a], Valentin Lindlacher[b,*]

[a]*Email:* goldbeck@ifo.de. *LMU Munich and ifo Institute for Economic Research, Munich (Germany)*
[b]*Email:* lindlacher@ifo.de. *LMU Munich and ifo Institute for Economic Research, Munich (Germany)*

March 15, 2022

Abstract

Does Internet availability at basic speeds foster local economic growth in developing countries? We analyze 220 towns in 10 Sub-Saharan African (SSA) countries in the early 2000s and measure local economic growth of towns using nighttime light satellite data. In a difference-in-differences setting, we exploit quasi-random variation in Internet availability induced by sub-marine cable arrivals. Our findings show that basic Internet availability leads to about two percentage points higher annual economic growth of SSA towns in the years after connection compared to a control group of similar but later connected towns. This result seems to be driven mainly by per capita productivity growth and only to a small extent by migration into connected towns. Moreover, Internet availability is accompanied by a shift from agriculture to manufacturing in regional employment shares.

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

JEL-Codes: O33, O18, R11

1. Introduction

In the last decades, the provision of digital infrastructure in many countries enabled widespread access and adoption of modern information and communication technologies (ICT), most prominently the Internet. Evidence shows positive effects of broadband Internet availability on individual-level economic performance (Akerman et al., 2015) and country-level economic growth (Czernich et al., 2011) for developed countries. Hopes are high that Internet access 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 to bring the Internet to everyone.

* Corresponding author. Poschingerstr. 5, 81679 Munich, Germany. We thank Vojtech Bartos, Mathias Bühler, Oliver Falck, Thomas Fackler, Jonas Hjort, Anna Kerkhof, Markus Ludwig, Niklas Potrafke, Helmut Rainer and Maria Waldinger for valuable comments and suggestions. Further, we thank Nicolas Göller for excellent research assistance. We are thankful for valuable comments at the 14th RGS Doctoral Conference in Economics, the third International Workshop "Market Studies and Spatial Economics", 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), and the Doctoral Workshop L'Association Française de Recherche en Economie Numerique (AFREN), especially, to the discussants Kathrin Wernsdorf, Marcelo Sant'Anna, and Tobias Korn. Finally, we thank the participants of seminars at ifo Institute and at Economic Research South Africa. All errors are our own.

However, the provision there is more complex and costly due to lacking legacy infrastructure, i.e., fixed-line telephony networks. Hence, SSA countries invested more than 28 billion US-Dollar into their national Internet backbone to date (Hamilton Research, 2020).[1] Despite these enormous investments in Africa's digital infrastructure, a growth effect of Internet in SSA is less obvious than it seems. Low population density apart from a few 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. It is thus crucial to investigate how Internet availability affects regional economic development in developing countries.

In this chapter, we examine if there is a causal effect of Internet availability on local economic growth in SSA even at basic speeds. We focus on the initial introduction of Internet in SSA through the 'first generation' of Internet-enabled 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 as well. Moreover, this 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 towns in 10 SSA countries which get international Internet connection before the Internet speed upgrade and which rolled out a 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 areas of SSA cities and towns from *Africapolis*. Second, we use data on the rollout of national backbones to measure Internet availability of individual cities and towns. The data comes from Hamilton Research (2020) and comprises the geo-location of all Internet access points in SSA. Because data on the establishment year of access points only starts in 2009, we backdate the establishment year of access points to their actual construction year via an extensive review of national backbone deployment projects for each SSA country. This enables us to study the 'first

[1] Facebook recently announced an effort to build a new Internet-enabled sub-marine cable (SMC) to Africa for one billion US-Dollar (Bloomberg, 2020). And China plans to invest more than 60 billion US-Dollar in Africa's digital infrastructure as part of its Belt-and-Road initiative (Invesco, 2019).

generation' of SMC arrivals, which introduced 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 countrywide Internet connections 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 towns that are located between nodal cities (political and economic centers) and are therefore incidentally connected. 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 rollout of the national backbone, which makes Internet available through access points. We assign the treatment status to towns that were connected to the national backbone when the Internet became available countrywide, while the control group consists of towns which get an 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 countrywide 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. The event-study results are robust to accounting for heterogeneous effects in the staggered timing of the treatment using recently proposed estimators by Roth and Sant'Anna (2021*a*), Callaway and Sant'Anna (2021), and Sun and Abraham (2020) and using less restrictive calendar-year fixed effects instead of country-year fixed effects.

We find that connection to the Internet through an access point leads on average to a 7 percent increase in NTL intensity of SSA towns in the first four years since countrywide 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 2 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 that towns with Internet access are becoming both brighter and larger. This provides suggestive evidence that towns with Internet access become more productive. We do not find effects on population growth and therefore reject the mechanism of migration. This strengthens our hypnotises of an effect of economic development. Moreover, we find a shift in regional industry shares. In connected regions, manufacturing employment shares increase by around 2 percentage points in comparison to regions getting connected later. These shares are mainly gained from decreasing agricultural employment shares in these regions, though this coefficient lacks statistical significance. This suggests that the increase in economic activity is at least partly a result of a changing industry structure induced by Internet availability.

To ensure that our results are driven by Internet availability, we control for the rollout of mobile coverage (GSM). At the time we are investigating, 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 existing. Additionally, we perform placebo tests with access to other potentially confounding infrastructure, such as roads, railroads, and the electricity grid. Moreover, we perform placebo exercises on the timing of countrywide Internet connections. We use 1,000 simulations with placebo country-connection years prior to the countries' actual connection year to show that the effect is only present for the actual connection years. For the robustness of our results, we test 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 ethnic favoritism. Finally, we extend the sample by relaxing some 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 employment in SSA between 2009 and 2012, causal evidence on the economic impact of Internet availability in developing countries is surprisingly rare. This is the first study investigating the causal effect of the introduction of (fixed-line) Internet availability at basic speeds on overall local economic growth in developing countries, which is measured by NTL satellite data. Furthermore, we study early Internet effects in a rural developing country 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.[2] Most closely

[2] 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.

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.[3]

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.[4] 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

[3] 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).

[4] 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.

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.

2. Background

There are three major components of Internet infrastructure determining the availability and bandwidth of Internet in a given location. First, international fiber-optic sub-marine cables (SMCs) connect SSA countries to the global Internet backbone.[5] Second, within-country inter-regional fiber-optic cables form the national backbone. 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: Sub-Marine Cables

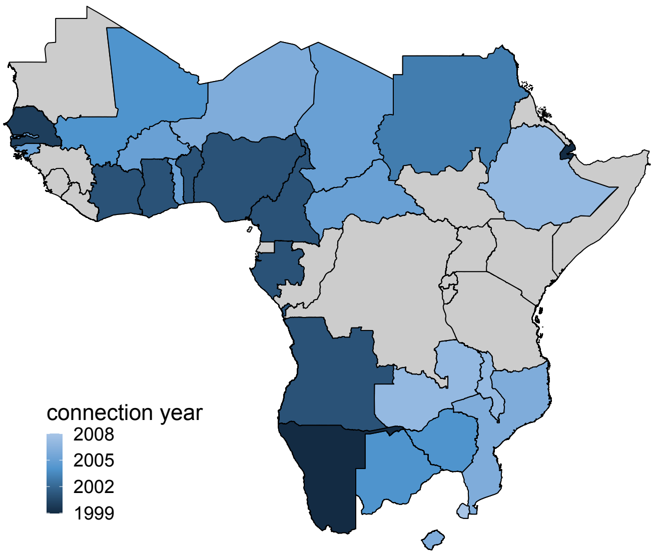
Since the vast majority of web pages and applications is hosted on servers located in North America or Europe, almost all African Internet traffic is routed inter-continentally (Kende and Rose, 2015; Chavula et al., 2015). Before the first SMCs landed on SSA shores, the only way to connect to the Internet on the continent was via satellite.[6] 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

[5] 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.

[6] 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.

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.

Internet at basic speeds.[7] 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.[8] These landing points, typically one per country, constitute the starting point for the respective national backbones (cf. Section 2.2). Until the late 2000s, most SSA countries were connected to the Internet via these 'first-generation' SMCs.[9]

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.

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

7 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)

8 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.

9 The 'second-generation' of SMC landed very similarly between 2009 and 2012.

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.[10] 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).[11] 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'.

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

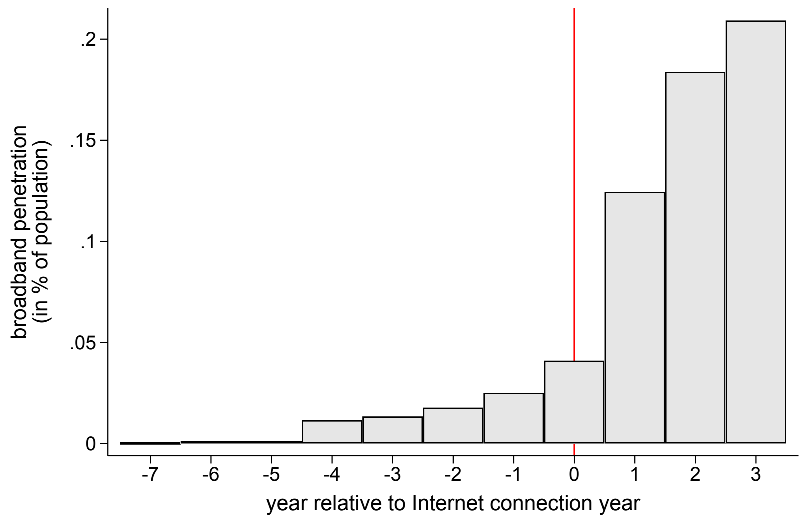
Internet traffic transported by inter-regional cables is accessed at access points. There are several technologies transmitting Internet traffic from these access points to the user. These 'last mile' transmission technologies include fiber cables (FTTH/B), copper cables, and wireless transmission using 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.

[10] 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.

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

Figure 2 shows how the usage increases in countries that were served by a 'first generation' SMC. Though, the change in absolute numbers is rather low (.2 percentage points), one should notice that the increase starts when the Internet becomes available. Compared to the year before Internet was accessed through an SMC, but only via a satellite connection, broadband penetration increased by eight times in only four years. 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 does not exist on a wide scale 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.[12]

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.

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

[12] https://www.enterprisesurveys.org

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 sub-marine 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.[13] The arrival is exogenous, first, because each SMC typically connects many countries. Therefore, coordination difficulties among consortium members might delay the construction.[14] 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 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). The definition of towns in the treatment and control group is depicted in Figure A.1. We exclude nodal cities, i.e., cities close to an access point that are endogenously connected (cf. Section 2.2): the landing point, the capital, regional capitals, and economic centers (cities with a population of more than 100,000 inhabitants). For robustness, we vary the threshold of 100,000 inhabitants as the definition of an economic center.

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

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

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

will only serve for robustness. As an 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

$$y_{c(i)t} = \beta_0 + \beta_1 \big(connection_{ct} \times access_{c(i)} \big) + \beta_2 GSM_{c(i)t} + \alpha_{c(i)} + \delta_{ct} + \epsilon_{c(i)t}, \tag{1}$$

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. 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 δ_{ct}. Note that this allows for country-specific time trends, especially country-specific growth rates, and variations in satellite sensor quality over years. 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:

$$y_{c(i)t} = \beta_0 + \sum_{j=\underline{T}}^{\overline{T}} \beta_{1j}\big(t_j \times access_{c(i)}\big) + \beta_2 GSM_{c(i)t} + \alpha_{c(i)} + \delta_{ct} + \epsilon_{c(i)t}, \tag{2}$$

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 β_{1j} inform about the dynamic effect of Internet availability. Thereby, each coefficient captures relative-year-specific treatment effects. We expect to see no effect before the treatment. Thus, if we cannot distinguish the estimates of the coefficients 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.

As a number of recent contributions have pointed out, two-way fixed effect (TWFE) event-study (or DiD) approaches, similar to the specification in Equation (2), may still yield biased estimates when treatment effects vary over time (see e.g., Athey and Imbens, 2021; de Chaisemartin and D'Haultfœuille, 2020; Borusyak et al., 2021; Goodman-Bacon, 2021; Sun and Abraham, 2020). The main reason is that the TWFE estimator uses already-treated towns as control group for newly-treated towns, causing a violation of the parallel-trends assumption in the presence of treatment effect dynamics. However, this is precluded when applying more rigorous country-year fixed effects as we do. If country-year fixed effects are applied, there is only one treatment (the year the international Internet connection was established). Nonetheless, we can relax the fixed effects and use the classical TWFE model with town and calendar-year fixed effects. To account for the resulting threat to identification, as described above, we perform alternative approaches proposed by Callaway and Sant'Anna (2021), Roth and Sant'Anna (2021*a*), and Sun and Abraham (2020) for robustness. For instance, Callaway and Sant'Anna (2021) suggest a two-step estimation strategy by first estimating 'group-time average treatment effects', where groups are defined by when towns are first treated, before aggregating the treatment effects by relative time using a propensity-score weighting method.

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 sub-marine 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.[15] This database contains the geographical delineation of 5,811 SSA agglomerations with more than 10,000 inhabitants in 2015. The median size is around 20,000 inhabitants and about 90 percent have less than 100,000 inhabitants. The population is also been made available for earlier years by *Gridded Population of the World (GPW)*.[16]

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 Meteorological Satellite Program* (DMSP) *Operational Linescan System* (OLS) between the early 1990s and 2013. The instruments of DMPS-OLS satellites measure light intensity on an integer scale from 0 to 63 with pixels covering 30 arc-second grid cells (an area of .86 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 Sub-Marine 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*.[17] 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

[15] https://africapolis.org
[16] https://sedac.ciesin.columbia.edu/data/collection/gpw-v4
[17] https://www.submarinecablemap.com

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*.[18] 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.[19] 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 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.[20]

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.[21] The data is sourced from *Collins Bartholomew*.[22] 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. For the definition of economic centers we take the

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

[19] 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.

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

[21] The share is usually either 0 or 1.

[22] https://www.collinsbartholomew.com/

population in the year 2000 from *Africapolis*. 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 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.[23] 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.

4.4. Combining the Data

Our analysis is focused on rather small 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 which 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.

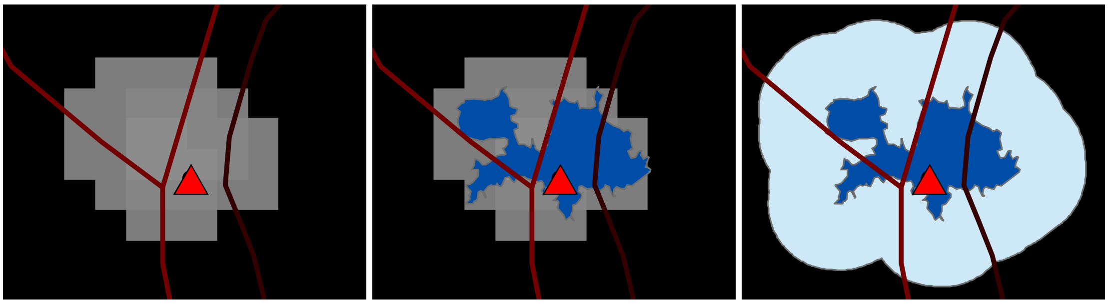
As light blurs out to adjacent pixels, cities appear bigger in the data than they actually are. By taking the extent of the towns in 2015, we capture some of the blurring as the towns might have been growing after our observation period. However, for some towns, the NTLs still might blur over the extent 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.[24] 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 3 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.

[23] The admin-2 level is below the state level.

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

Figure 3: Data example: Dassa-Zoumè, Benin (2004)



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 neighbouring 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 kilometer buffer around that built-up area.

Within each town, we define several outcome measures.[25] Local economic activity is measured by summing the light intensity of all pixels within a town (and the 2 kilometer 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 three post-treatment years before a high-capacity SMC connects them.[26] 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

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

[26] 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 Section Appendix D.

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.[27]

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 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 around 25,000 inhabitants on average. Control towns, which are connected in the subsequent years after the observation period of three years, vary for all subsequent years between a population of 10,000 and 35,000 inhabitants as well. Especially, when only examining treated and control towns, i.e., excluding nodal cities, there is no clear (decreasing) pattern with respect to population size over time anymore in the first ten years. Such a pattern is clearly identified for nodal cities. Only towns which get access rather later are smaller on average. We account for this timing in a robustness check. 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 to 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.

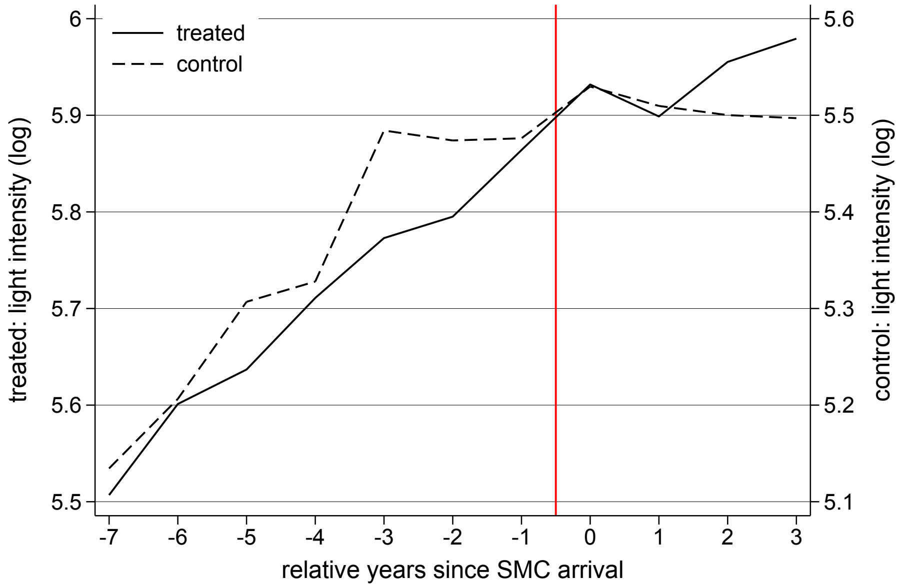
[27] Sudan is a special case as ten towns are in the control group but only one town is in the treatment group. Angola has very similar issues. We account for that by grouping fixed effects for East, Southern, and West African countries for robustness.

Nonetheless, the size of later connected nodal cities is still bigger on average than the size of the control 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 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.

Before presenting the estimation results, we show the development of cities and towns over time. We use the main outcome measure, log light intensity, averaged over the treatment and control group separately but do not include any fixed effects or controls. Figure 4 shows that in the early pre-treatment years both groups grew on average with similar rates. While the gap between treatment and control group towns is almost equal in t-7 and t-1, the control group grew slightly faster in the earlier year and the treatment group in the later years. At the end of the observation period, the gap between these towns grew by about .1 on the logarithmic scale. While for the control group there is a stagnation, for the treatment group the growth is also slowed down. This pattern holds for almost all types of cities. Figure A.6 shows the growth rates of nodal cities and towns. For all groups but the economic centers, annual-growth rates declined after the Internet connection, while the country's GDP growth accelerated in the same period (Figure A.7). Thus, the slowdown in NTL growth is probably an inconsistency in the satellites. As Figure A.6 shows, before the Internet connection all town types had a similar annual growth rate. However, when Internet became available, treated towns showed an a lot larger growth rate than control group towns.

Figure 4: Time trends of treatment and control group



Notes: The figure depicts the average growth of the towns in the treatment and control group over a period of eleven years (seven before and three after the treatment year). The measurement is the logarithm of light intensity.

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 Section Appendix F.

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

	(1) light intensity	(2) light intensity	(3) intensive margin	(4) extensive margin	(5) population
post x treated	0.0633*	0.0703**	0.0513**	0.0516*	0.0116
	(0.0344)	(0.0349)	(0.0231)	(0.0282)	(0.0183)
GSM coverage		0.0486	0.0477**	0.0281	0.00699
		(0.0342)	(0.0240)	(0.0263)	(0.0119)
observations	2,420	2,420	2,420	2,420	2,420
R-squared	0.943	0.943	0.947	0.924	0.999
#countries	10	10	10	10	10
#towns	220	220	220	220	220
share treated	.445	.445	.445	.445	.445
town FE	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓

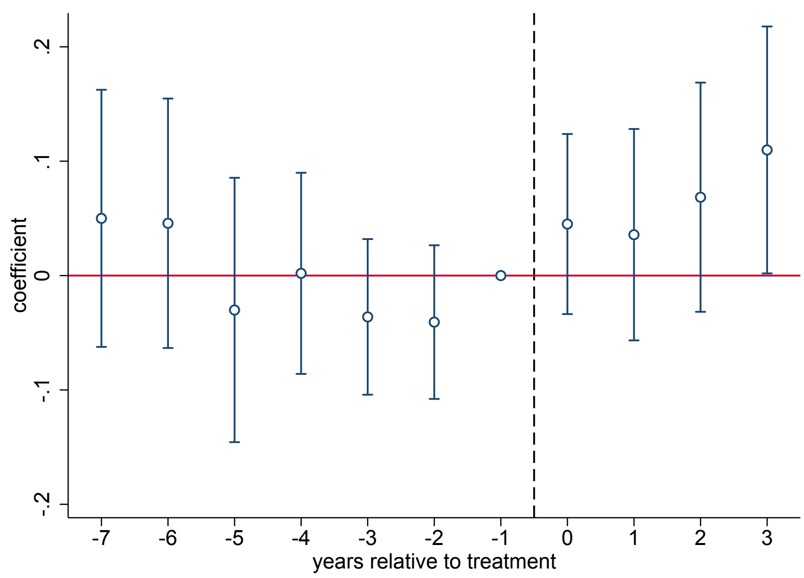
Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer. The measure for the intensive margin is logarithmic mean light intensity and for the extensive margin logarithmic sum of lit pixels (coded as 1 if a pixel is lit). Population is measured as inhabitants per square kilometer within town area (including the 2 kilometer buffer) from *Gridded Population of the World* (GPW). GSM mobile coverage is calculated as the percentage share of town area covered with signal. Nodal cities include landing point and capital cities, regional capitals, and cities with more than 100,000 inhabitants. All specifications include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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 $\epsilon_{GPD,light} = 0.284$ 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.

Figure 5 presents event-study coefficients for our preferred specification (Column (2) of Table 1). Before SMC connection, the point estimates are close to zero and statistically insignificant. This supports the assumption that treatment and control group towns are not of 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 until year t+3. In the years after the countrywide Internet connection, the point estimates are between .05 and .1 and have a slight tendency to increase to nearly .15 in year t+3, when the effect gets statistically significant at the 5% level.

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 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 5: Event-study coefficients



Notes: Coefficients for event-study specification of Column (2) of Table 1. Robust standard errors clustered by town. Confidence intervals reported at the 95% level.

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.

Investigating population as an outcome, the estimate is close to zero and lacks statistical significance. Therefore, the effect of Internet availability is mainly inducing economic development and not population growth or migration. This suggests that the towns' increase in light intensity reflects primarily an increase in their productivity. As we use an interpolation for population from *GPW*, Table B.3 shows different specifications regarding the included periods and years for robustness. The effect of Internet connection on population lacks statistical significance and the point estimate is close to zero in all specifications.

In a further robustness test, we add population as a (bad) control instead of using it as an outcome as before. In Table B.4, the regressions from Table 1 are repeated in odd columns and population is added in even columns. By doing so, we control for which part of the effect is driven by economic growth and which by growth in population. The main effect on light intensity remains robust and statistically significant (Column (2)), strengthening our claim that the effect of Internet availability is mainly inducing economic development and not population growth. Population has an elasticity of around .4 on light intensity. The coefficient of the population control is statistically significant at the 10% level (Column (2)). The main effect on the intensive and extensive margin remains robust as well when additionally controlling for population. However, the population estimate differs in both specifications. For the intensive margin, the population estimate is smaller and statistically insignificant (Column (4)). For the extensive margin, the population estimate is positive and statistically significant (Column (6)). As population is not correlated with brightness, towns are not getting populated more densely. Again, strengthening our claim of an effect on towns' productivity. We take from the correlation between population and size that if there is any small increase in population, it is induced by a few settlers at the towns' borders.

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).[28] 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 2 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 the 5% level. Comparing these results with the growth rates in Columns (2) through (4), the signs for

[28] https://international.ipums.org/international/

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 2: Employment growth rates and shares by industry

	growth rate				industry share		
	(1) total	(2) agriculture	(3) manufacturing	(4) services	(5) agriculture	(6) manufacturing	(7) services
post x treated	-0.0215	-0.105	0.131	-0.0672	-0.0328	0.0225**	0.0103
	(0.0569)	(0.0908)	(0.104)	(0.0970)	(0.0239)	(0.0102)	(0.0184)
GSM coverage	-0.0570	-0.0824	-0.180*	-0.0877	-0.00696	-0.000510	0.00747
	(0.0570)	(0.0739)	(0.0968)	(0.0878)	(0.0172)	(0.00583)	(0.0130)
observations	288	288	288	288	288	288	288
R-squared	0.967	0.964	0.934	0.951	0.954	0.893	0.958
#countries	5	5	5	5	5	5	5
#regions	144	144	144	144	144	144	144
share treated	.181	.181	.181	.181	.181	.181	.181
region FE	✓	✓	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓	✓

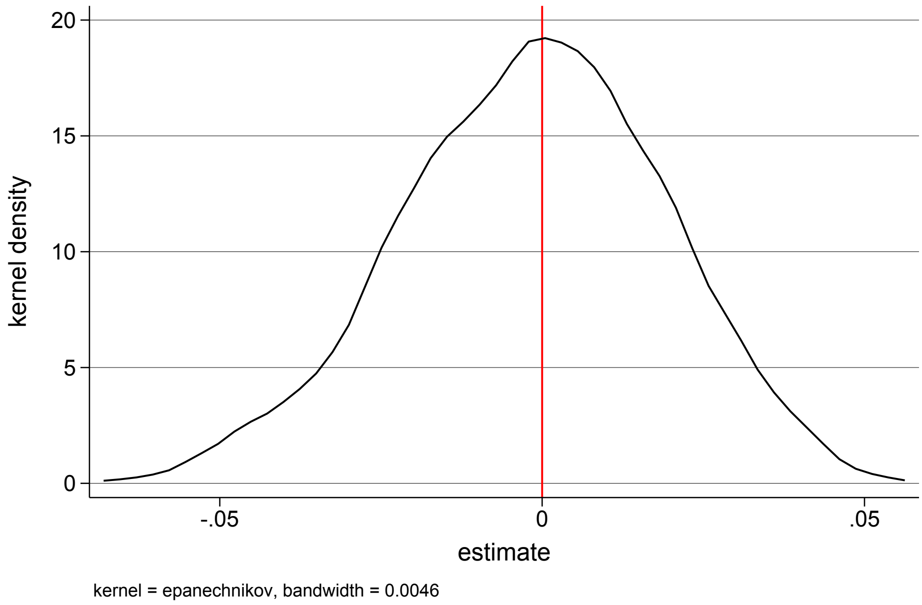
Notes: Regional industry composition comes from *IPMUS International.* All specifications are estimated on a sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.2. Placebos

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 a placebo test, we now assign randomly connection years to each country individually (prior to the actual connection year). 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 1000 times and plot the distribution of the main effect. Figure 6 shows a very symmetric distribution with the peak being very close to zero. The main effect only remains with the actual connection years and diminishes with randomly assigned ones. This indicates that the connection years are exogenous.

Placebo Tests with Nondigital Infrastructure. In Table 3, we include additional controls. To save costs, fiber cables are rolled out along the existing (transportation) infrastructure

Figure 6: Placebo (randomly assigned country-connection years)



Notes: The figure depicts the distribution of the main effect, estimated with 1000 random country-connection year combinations.

(cf. Section 2.2). We control for this other infrastructure to rule out that towns closer to this nondigital 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.[29] Unlike for mobile coverage, we do not have time-varying data on other infrastructures. We therefore generate indicators for whether the distance is below 10 kilometers (as we defined treated towns with access points) and then intersect these indicators with the post dummy for the time after the countrywide Internet connection is established and construct placebo treatments. Column (1) repeats our preferred specification (Table 1, Column (2)). In Column (2), we control for different effects for towns next to a greater (paved) road. In Column (3), we control for different effects for towns next to the railroad network. In Column (4), we control for different effects for towns next to the electricity grid. Finally, we include all infrastructure controls jointly (Column (5)). In each case, the estimate of the main effect remains close to .07 and is statistically significant. The estimates of all placebos lack statistical insignificance and are rather small in their economic significance. In the case of roads, the main estimate looses

[29] 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 which 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 grind in more detail later.

slightly precision and decreases by half a percent to .065. Nevertheless, controlling for access to the railroad network, the electricity grid, or all infrastructures jointly, the estimate is again statistically significant at the 5% level and slightly higher than in the baseline (7.3 to 9.4 percent).

Table 3: Placebo (competing infrastructure)

light intensity (log)	(1)	(2)	(3)	(4)	(5)
post x treated	0.0703**	0.0647*	0.0942**	0.0733**	0.0903**
	(0.0349)	(0.0349)	(0.0376)	(0.0356)	(0.0375)
GSM coverage	0.0486	0.0477	0.0463	0.0480	0.0450
	(0.0342)	(0.0341)	(0.0340)	(0.0343)	(0.0340)
post x road network (dummy)		0.0928			0.104
		(0.0768)			(0.0757)
post x railroad network (dummy)			-0.0624		-0.0657
			(0.0385)		(0.0406)
post x electricity grid (dummy)				-0.0227	-0.00718
				(0.0391)	(0.0405)
observations	2,420	2,420	2,420	2,420	2,420
R-squared	0.943	0.943	0.943	0.943	0.943
#countries	10	10	10	10	10
#towns	220	220	220	220	220
share treated	.445	.445	.445	.445	.445
town FE	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer as in Table 1. 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 100,000 inhabitants and include town and country-year fixed effects. Railroad and road networks are source from *Natural Earth*. The electricity grid is sourced from *Africa Infrastructure Country Diagnostic*. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Electricity Grid. 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).[30] 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 B.5, 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)),

[30] https://afrobarometer.org/

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.

5.3. Robustness

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 (2) of Table 4 remains robust. However, the estimate of the control variable GSM mobile coverage turns statistically highly significant. Figure A.9 shows the event-study plot, which is very similar in comparison to the preferred specification. Though, the pre-trends have higher (statistically insignificant) point estimates. In Subsection 5.4, External Validity, we estimate further models with the classical two-way fixed effects, which allow for a bigger sample, containing more countries as the treatment and control group are compared across countries. Moreover, we apply novel DiD/event-study estimators to account for the staggered timing of the treatment (Roth and Sant'Anna, 2021*a*; Callaway and Sant'Anna, 2021; Sun and Abraham, 2020).

Error Correlation within Regions. Another potential concern is that model errors are correlated within regions. This might be the case when the effects of the access point might generate further spillover effects in the towns' surrounding area. In our preferred specification, we cluster the standard errors at the town level as the treatment, the access point construction, is occurring there. If more than one town is located within 10 kilometers to the access point, an access point would serve more than one town. 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 (3) 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 .0349 to .0358).

Linear Time Trends. In Column (4) of Table 4, we also test the robustness of our results against the inclusion of linear time trends on the town level. To account for possible differential trends among towns, we re-estimate Equation (1) including a linear town-specific yearly trend. This is the most demanding specification. The estimate increases

Table 4: Robustness

light intensity (log)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
post x treated	0.0703**	0.0580*	0.0703*	0.102**	0.0625*	0.0600*	0.0797***	0.121***	0.0720*
	(0.0349)	(0.0338)	(0.0358)	(0.0449)	(0.0336)	(0.0323)	(0.0265)	(0.0392)	(0.0374)
GSM coverage	0.0486	0.0908***	0.0486	0.0193	0.0233	-0.00762	0.0193	0.0611	0.0482
	(0.0342)	(0.0320)	(0.0330)	(0.0332)	(0.0310)	(0.0308)	(0.0249)	(0.0380)	(0.0408)
observations	2,420	2,420	2,420	2,420	2,827	3,839	2,343	2,860	1,804
R-squared	0.943	0.927	0.943	0.962	0.951	0.927	0.978	0.937	0.939
#countries	10	10	10	10	10	10	10	10	9
#ethnic group-countries									13
#towns	220	220	220	220	257	349	213	220	164
share treated	.445	.445	.445	.445	.525	.281	.502	.445	.445
town FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
country x year FE	✓		✓	✓	✓	✓	✓	✓	
year FE		✓							
ethnic group-country x year FE									✓
w/o capital+landing point	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o population >100k	✓	✓	✓	✓	✓	✓	✓	✓	✓
cluster			state-level						
linear time trends				town-level					
late APs					✓				
no not treated towns						✓			
no buffer							✓		

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area. 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 100,000 inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

slightly to .10 and is also statistically significant at the 5% level. Thus, we can show that after the countrywide Internet connection is established, the connected towns grow in a nonlinear manner.

Access Points During the Post-Observation Period. Thus far, we tested the robustness of our results to alternative assumptions about the variance-covariance matrix. In Column (5) of Table 4, we add these towns to the estimation sample that were connected only within three 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. Including them, we allow the treatment status to switch from 0 to 1 in the year the access point is constructed (the post dummy switches in the same year as well). With this approach, we can add them without confounding our analysis. Within our ten countries in the analysis, we can add further 37 towns to increase the estimation sample. We estimate Equation (1) as in our preferred specification with town and country-year fixed effects and cluster standard error at the town level. As the estimate remains, including slightly later connected towns does not affect the results.

Extending the Control Group. In Column (6) 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 countrywide available. At that time, it might not yet be known which towns will get an access point in the future. While this sample increases to

349 towns, the share of treated towns declines to 28.1 percent. The estimation results remain unaffected.

No Buffer. In Column (7) of Table 4, we remove the 2 kilometer buffer and estimate on the original *Africapolis* built-up areas. We adjusted the built-up areas because of the blurring of the NTL data. When examining the smaller built-up area, we might lose some pixels at the towns' border. However, these pixels might be of low intensity. With this robustness check, we can thus 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. The estimate not only remains, but turns statistically significant at the 1% level.

Extending the Post-Treatment Period. Originally, we limited the sample to the years before the 'second generation' of SMCs arrived, such that we can estimate the effect of broadband Internet at basic speeds. The event-study estimates in Figure 5 show that the effect on local economic growth increased on a yearly basis. Therefore, we show for robustness how this effect evolves in the two subsequent years (Column (8) of Table 4). One should note, however, that the effects in the last periods might be driven by fast Internet induced by the new SMCs. Six countries still have Internet at basic speeds available at the end of the new sample. The event study estimates are shown in Figure A.10 and indicate that the growth rate increases further to more than 20 percent.

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 cause 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.11). 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 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 constructing country-ethnic group entities instead of countries. By estimate 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 (7) of Table 4. A hardly 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 nine countries, there are eleven ethnic groups and thirteen country-ethnic group entities in the estimation. 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.

5.4. *External Validity*

Applying the classical two-way fixed effect model of Column (2) in Table 4, we re-estimate Equation (1) on a broader sample. Comparing treatment and control group towns in the whole sample through the calendar year fixed effect, we allow also for countries containing either only control or treated towns. Table 5 shows the results with a stepwise reduction of nodal cities. In Column (1), the most basic specification, all nodal cities are included. Then, we eliminate nodal cities stepwise until we reach our preferred specification, where the remaining towns are comparable: In Column (2), we remove the city of the landing point and the national capital for each country. In Column (3), we additionally remove all regional capitals. In Column (4), we further remove all cities of more than 100,000 inhabitants. Finally, we add GSM mobile coverage as a control variable to account for other telecommunication technologies (Column (5)). With this approach, we estimate on 491 (including nodal cities) to 352 towns in 19 to 17 countries. In one country, only the capital (and/or the landing point) is connected, in another country only economic centers with more than 100,000 inhabitants are connected. It can easily be seen that this sample contains more control towns as most countries added to this sample started the rollout rather late. A concern might be a correlation between countries' development and timing of its construction of access points. However, estimating on a sample containing more countries shows that our estimates also reach a high external validity.

The estimate of the main effect increases slightly from Columns (1) through (5). This was already the case in Table 1 in Column (2) when mobile coverage was added as a control. The estimate is slightly higher, reaching an effect size of 9 percent and reaches significance at the 1% level in our preferred specification in Column (5). In comparison to Column (2) in Table 4, where we applied the same fixed effects on the main results sample, the estimate of the main effect increases by nearly .04 and is statistically significant at a higher level. Hence, in the other SSA countries, towns might have developed worse than the control group towns in the original sample.

The event-study graph shows parallel trends before the treatment (Figure A.12). Especially in the four years prior to the treatment, the point estimates are close to zero. Only the estimate in year t-5 is statistically significant at the 5% level and has a negative sign. This is not the case in the main specification in Figure 5, where all pre-treatment

Table 5: External validity

light intensity (log)	(1)	(2)	(3)	(4)	(5)
post x treated	0.0620**	0.0712**	0.0833**	0.0888**	0.0938***
	(0.0263)	(0.0282)	(0.0325)	(0.0345)	(0.0343)
GSM coverage					0.0415*
					(0.0237)
observations	5,401	5,170	4,048	3,872	3,872
R-squared	0.963	0.947	0.936	0.916	0.916
#countries	19	18	18	17	17
#cities	491	470	368	352	352
share treated	.334	.309	.307	.287	.287
town FE	✓	✓	✓	✓	✓
year FE	✓	✓	✓	✓	✓
w/o capital+landing point		✓	✓	✓	✓
w/o regional capitals			✓	✓	✓
w/o population >100k				✓	✓

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer as in Table 1. GSM mobile coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and all specifications include town and year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

coefficients are statistically insignificant. After the treatment, both event-study estimates are very comparable. In both cases, the point estimate increases especially in year t+2 and even more in year t+3 and becomes statistically significant at the 5% level only in the last observation period.

One concern in this setting is that the classical two-way fixed effect estimator does not account for the staggered timing of the treatment and thus the potential heterogeneous effects of the individual treatment. Recent literature developed estimators for this setting (Roth and Sant'Anna, 2021*a*; Callaway and Sant'Anna, 2021; Sun and Abraham, 2020). Figure A.13 shows event-study estimates for the respective estimators.[31] All three approaches show very similar results. Again, it is shown that the results hold when accounting for heterogeneous outcomes in staggered treatment timing by the propensity score weighting method and potential comparing of treated and not-yet-treated observations. In contrast to Figure A.12, only the point estimate of year t+2 is statistically significant and the point estimate in year t+3 also declines slightly in comparison to the year before. However, the point estimate in year t+2 is bigger than in Figure A.12. Another differences to Figure A.12 is that in only one estimation the point estimate in year t-5 is negative and statistically significant at the 5% level (in the approach by Sun and Abraham (2020)). In the other two cases, this point estimate lacks statistical significance.

[31] We used the staggered R-package by Roth and Sant'Anna (2021*b*).

5.5. *Heterogeneity: Coastal Countries*

Storeygard (2016) investigates coastal countries due to the design of estimating on the distance to a primate city with a harbor. Hjort and Poulsen (2019) analyze coastal countries as they exploit the landing of SMCs and do not consider Internet connections of landlocked countries. So far, we have used the additional information about the connection year through a neighboring country we have on landlocked countries for estimating on a bigger sample. Nonetheless, one might have concerns about the validity of the exogeneity assumption of the timing of the connection year. Therefore, we reduce the sample one more time to estimate only on coastal countries for which the ground work for the identification is already established by Hjort and Poulsen (2019).

A priori, it is not clear whether the Internet has different effects for coastal and landlocked countries and if it does, which geographic location profits more from an Internet connection.[32] On the one hand, one could argue that more developed countries might have higher growth rates as some development has to be existing for the Internet to have an economic effect. On the other hand, less developed countries might have higher growth potential and the Internet could work as a substitute for worse nondigital infrastructure. In this case, countries could be leapfrogging and overtake more developed countries. However, coastal countries might profit more from international trade as the Internet lowers information costs, e.g., for international prices.

Table 6 shows results for coastal countries only. Column (2) presents our preferred specification, while Column (1) shows results without the mobile coverage control. We estimate on the five coastal countries (Angola, Benin, Sudan, Senegal, and Togo), which contain 75 towns in either the treatment or control group of which slightly more than half are in the treatment group. The estimate increases slightly and the standard errors decrease only very slightly from Column (1) through (2). The effect size is higher than in the whole sample (Table 1), indicating that coastal countries profit more from Internet access than landlocked countries. An advantage of investigating coastal countries separately is that they were connected earlier. Therefore, it is possible to analyze them with a longer post-treatment period without allowing the upgrade SMCs to confound the results. In Column (3), we re-estimate Equation (1) on a sample with five post-treatment years. The main effect increases to .27 and is statistically significant at the 1% level. The event-study plot, Figure A.14, shows that the effect size increases further in the fourth and especially the fifth year after the treatment. Parallel trends in the pre-treatment periods are again present. In contrast

[32] Though, coastal countries are not necessarily more developed, in terms of their GPD per capita, than landlocked countries as Botswana has by far the highest GDP per capita of the ten countries in our analysis.

Table 6: Heterogeneity (coastal countries)

light intensity (log)	(1)	(2)	(3)	(4)	(5)	(6)
post x treated	0.161*	0.165**	0.271***	0.218**	0.214**	0.214**
	(0.0820)	(0.0815)	(0.0884)	(0.0979)	(0.0925)	(0.0918)
GSM coverage		0.0797	0.102	0.0878	0.0873	0.0887
		(0.0622)	(0.0687)	(0.0642)	(0.0639)	(0.0625)
post x distance to next port				0.309		
				(0.227)		
post x treated x distance to next port				-0.406		
				(0.497)		
post x distance to coastline					0.352	
					(0.241)	
post x treated x distance to coastline					-0.482	
					(0.504)	
post x distance to landing point						0.136
						(0.216)
post x treated x distance to landing point						-0.358
						(0.398)
observations	825	825	975	825	825	825
R-squared	0.902	0.903	0.898	0.903	0.904	0.903
#countries	5	5	5	5	5	5
#towns	75	75	75	75	75	75
share treated	.507	.507	.507	.507	.507	.507
town FE	✓	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓	✓
w/o nodal cities	✓	✓	✓	✓	✓	✓

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer as in Table 1. 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 100,000 inhabitants and include town and country-year fixed effects. Ports are source from *OSM*. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

to the main specification with all countries (Figure 5), there is a significant increase in the treatment year, before the estimate declines slightly and becomes insignificant again.

A further advantage of coastal countries is that international trade, approximated through the towns' proximity to ports, can be investigated as a potential mechanism. We therefore include an additional interaction term, distance to the next port, in Column (4). The main effect increases slightly and remains statistical significance at the 5% level, while the triple interaction lacks statistical significance and has a negative point estimate. Later means that proximity to a port does not have a significant positive effect on local economic growth and that the mechanism of trade does not play an important role. The additional interaction of the post dummy and the distance to the next port lacks statistical significance as well. Hence, towns closer to a port do not grow faster after the countrywide Internet connection than towns further away from a port. In Columns (5) and (6), we repeat the triple interaction with the distance to the coastline and to the landing point. Results are similar to the triple interaction with the distance to the next port. The distance to the landing point accounts for both the distance the fiber-cable rollout has to cover and for the distance to the country's most important port. The distance to the coastline might reflect general growth potential and economic development of the coastal regions in comparison to the hinterland (Henderson et al., 2012) and also account for smaller ports, which might not be included in *OSM*.

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) estimates 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-size 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 sub-marine 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 Mbiti, I. M. (2010), 'Mobile Phones and Economic Development in Africa', *Journal of Economic Perspective* **24**(3), 207–232.

Akerman, A., Gaarder, I. and Mogstad, M. (2015), '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.

Athey, S. and Imbens, G. W. (2021), 'Design-based analysis in Difference-In-Differences settings with staggered adoption', *Journal of Econometrics* .

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

BenYishay, A., Rotberg, R., Wells, J., Lv, Z., Goodman, S., Kovacevic, L. and Runfola, D. (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 Krause, M. (2018), 'Top lights - bright cities and their contribution to economic development', *CESifo Working Paper* **No. 7411**.

Borusyak, K., Jaravel, X. and Spiess, J. (2021), 'Revisiting Event Study Designs: Robust and Efficienct Estimation', *Work in Progress* pp. 1–48.

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

Callaway, B. and Sant'Anna, P. H. (2021), 'Difference-in-differences with multiple time periods', *Journal of Econometrics* **225**(2), 200–230. Themed Issue: Treatment Effect 1.

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., Feamster, N., Bagula, A. and Suleman, H. (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', Springer International Publishing, Cham, pp. 64–73.

Chen, X. and Nordhaus, W. D. (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 Kanbur, R. (2017), 'Secondary Towns and Poverty Reduction: Refocusing the Urbanization Agenda', *Annual Review of Resource Economics* **9**(November), 405–419.

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

Colombo, M. G., Croce, A. and Grilli, L. (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.

Conley, T. G. (1999), 'GMM estimation with cross sectional dependence', *Journal of Econometrics* **92**(1), 1–45.

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

Czernich, N., Falck, O., Kretschmer, T. and Woessmann, L. (2011), 'Broadband Infrastructure and Economic Growth', *Economic Journal* **121**(552), 505–532.

de Chaisemartin, C. and D'Haultfœuille, X. (2020), 'Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects', *American Economic Review* **110**(9), 2964–2996.

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

Ghani, E., Goswami, A. G. and Kerr, W. R. (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.

Goodman-Bacon, A. (2021), 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics* **225**(2), 254–277. Themed Issue: Treatment Effect 1.

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

Grogan, L. and Sadanand, A. (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., Storeygard, A. and Weil, D. N. (2012), 'Measuring economic growth from outer space', *American Economic Review* **102**(2), 994–1028.

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

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

Hjort, J. and Tian, L. (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 Rose, K. (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., Zhou, Y., Zhao, M. and Zhao, X. (2020), 'A harmonized global nighttime light dataset 1992–2018', *Scientific Data* **7**(1), 1–9.

Nunn, N. and Puga, D. (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'.

Roth, J. and Sant'Anna, P. H. (2021*a*), 'Efficient estimation for staggered rollout designs', *arXiv preprint arXiv:2102.01291* .

Roth, J. and Sant'Anna, P. H. C. (2021*b*), 'Staggered: R module for implementation of DiD estimators'.

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

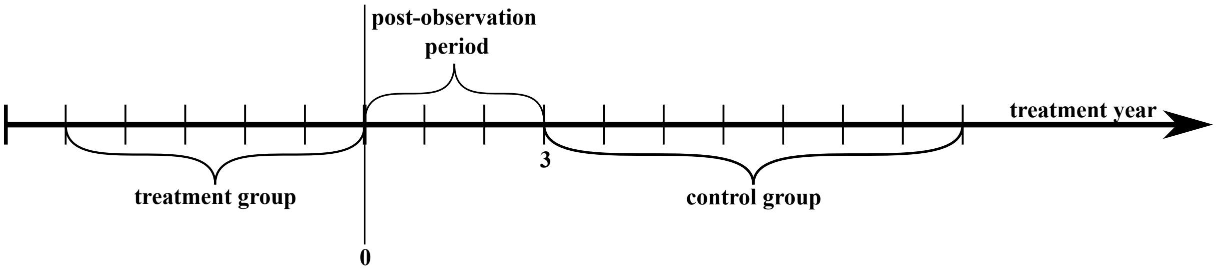
Storeygard, A. (2016), 'Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa', *Review of Economic Studies* **83**(3), 1263–1295.

Sun, L. and Abraham, S. (2020), 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics* .

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

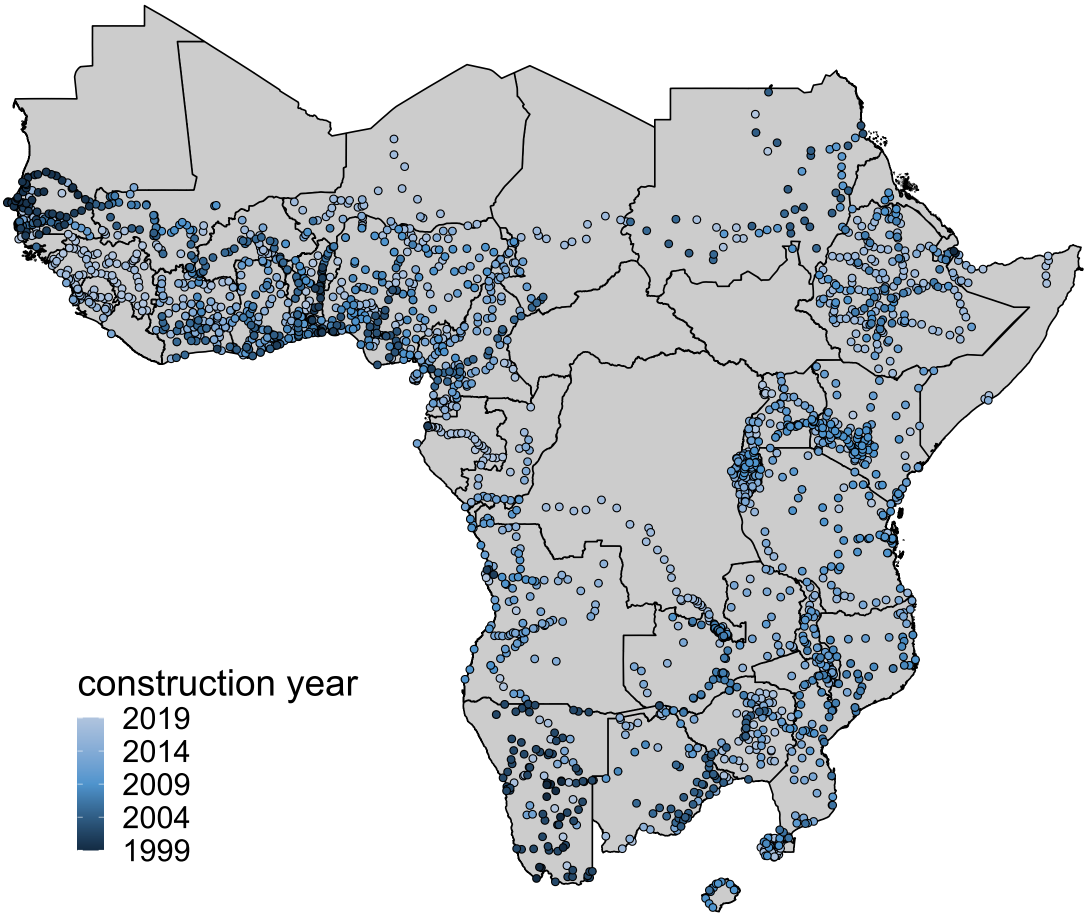
Appendix A. Additional Figures

Figure A.1: Timeline identification



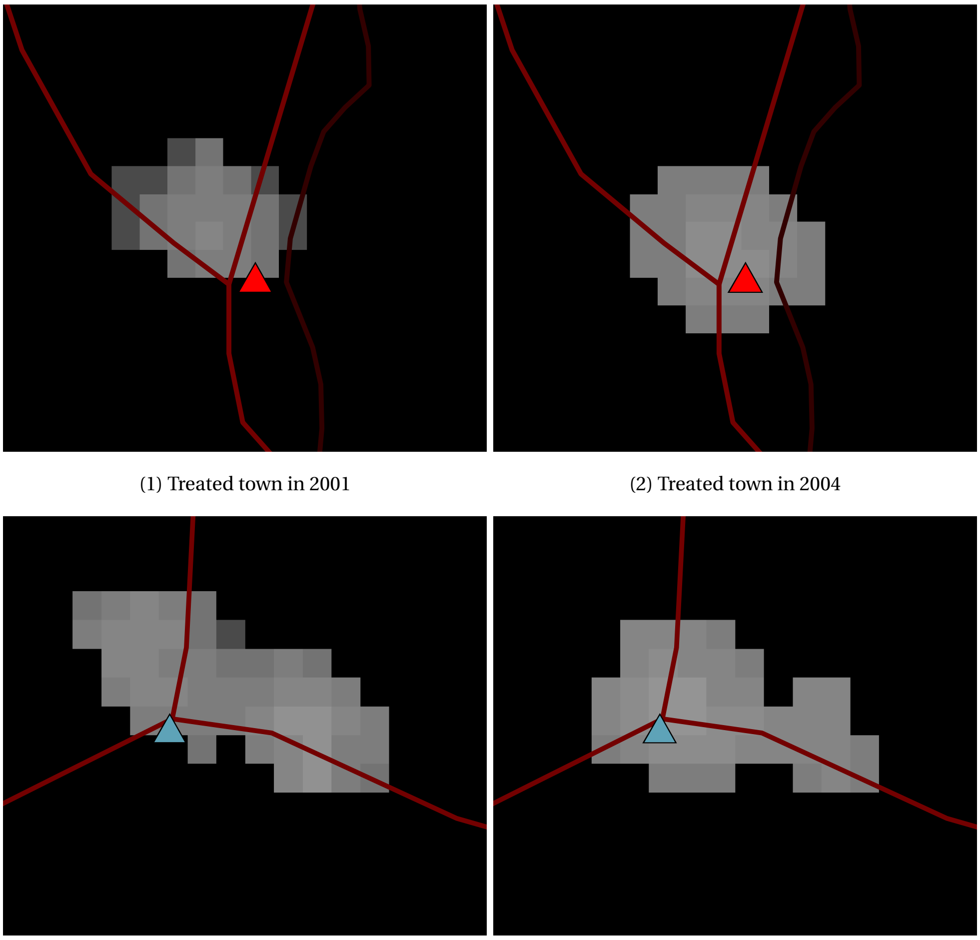
Notes: The figure gives an overview, of the timeline of our identification strategy. Treated are all towns having already an access point to the national infrastructure, when the Internet becomes available countrywide. As we investigate the effect of Internet availability for three years after the treatment, the control group contains all towns that get an access point in any year after theses three years.

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



(1) Treated town in 2001

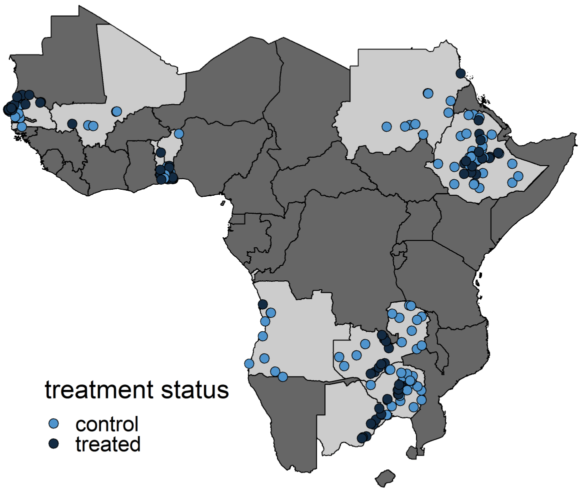
(2) Treated town in 2004

(3) Control group town in 2001

(4) Control group town in 2004

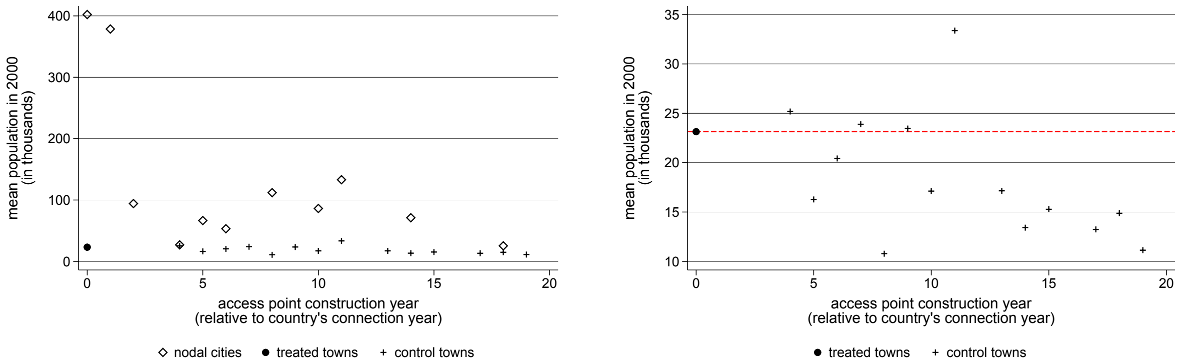
Notes: The panels show a treatment and control group town from Benin, with gray NTLs pixels from 2001 and 2004. Access points are marked with a triangle (red if constructed until 2001 and blue if constructed afterwards). The dark red line represents a major road connecting and the darker red line the 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



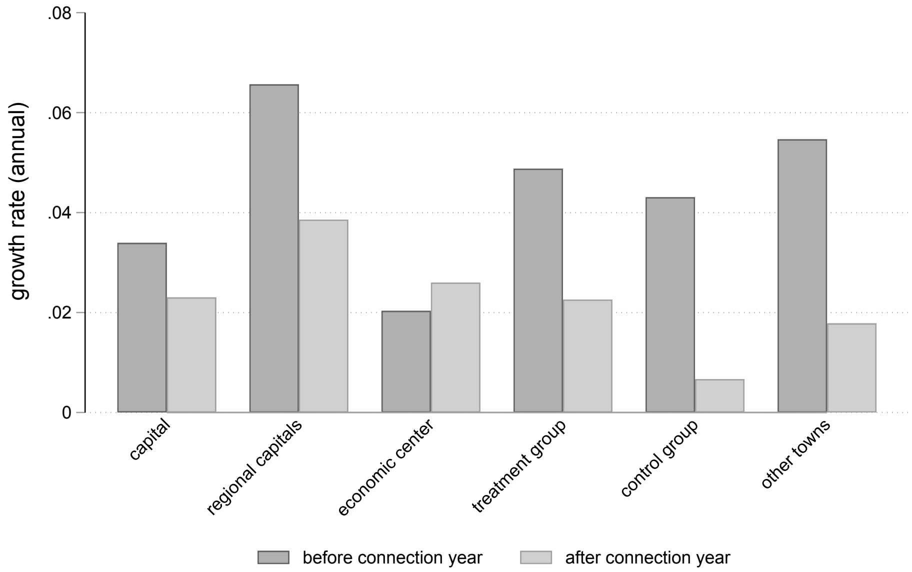
Notes: This figure depicts the countries in our analysis (brighter gray) and for each country the towns in the treatment and control group.

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



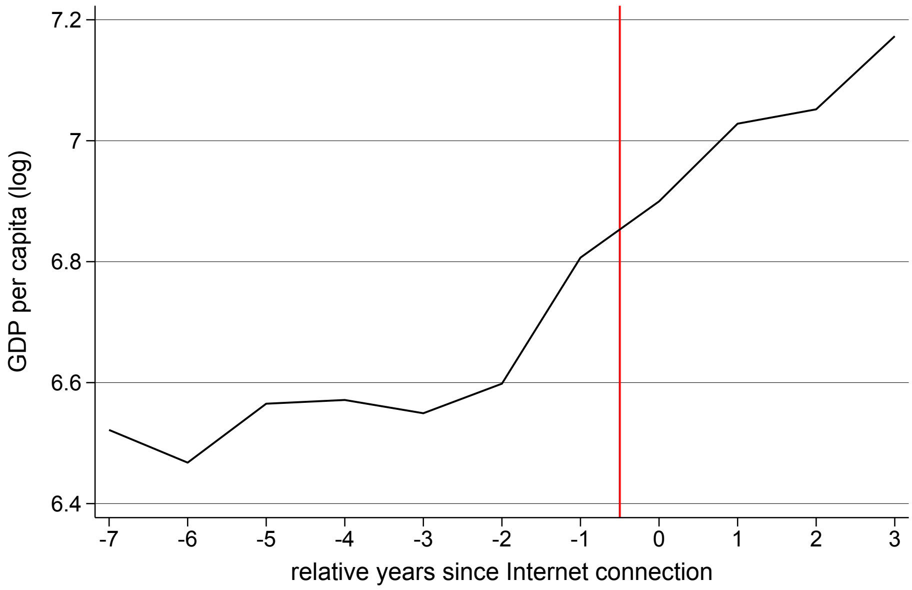
Notes: The figure depicts the average population size of connected cities and towns by year relative to the connection year. On the left, the black dot in the lower left corner represents the treated towns, while the control towns are represented by the plus symbol and the nodal cities by a diamond. 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: Comparison of city types (annual growth rates)



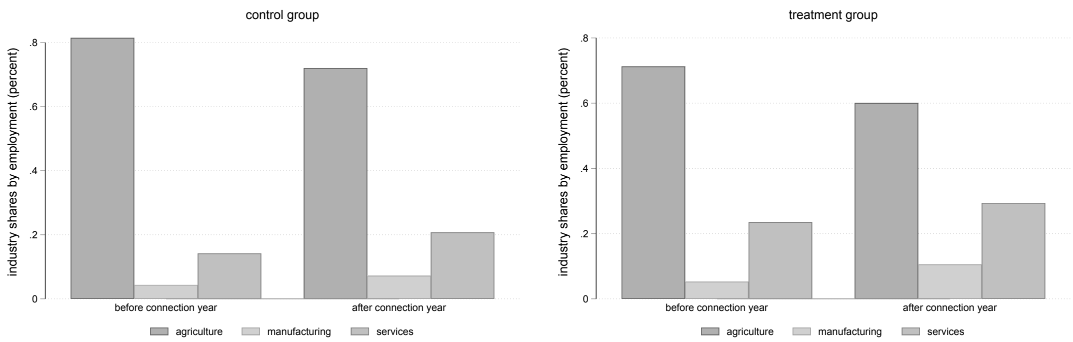
Notes: This figure shows the annual growth rates of the different city types measured by NTL intensity.

Figure A.7: Trend of GDP per capita growth



Notes: This figure shows the national GDP per capita growth before and after the Internet connection was established.

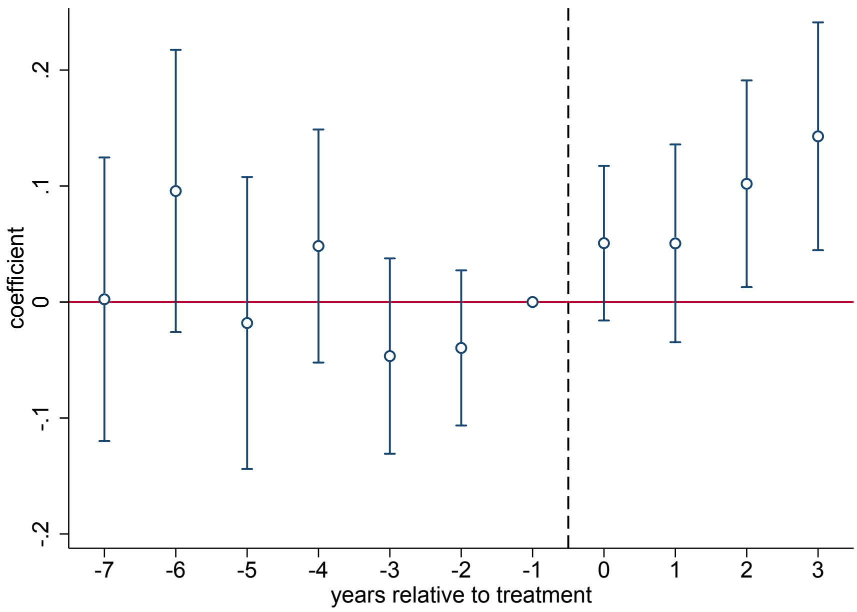
Figure A.8: Comparison of industry shares



(1) (2)

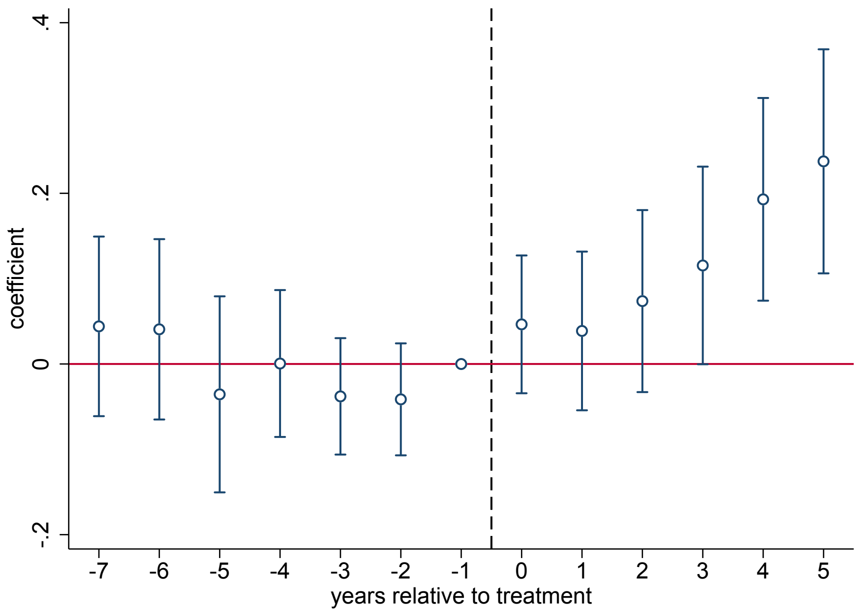
Notes: The figures depict the changes in the industry shares before and after the Internet connection for the treatment (1) and control group (2).

Figure A.9: Robustness event study (classical fixed effects)



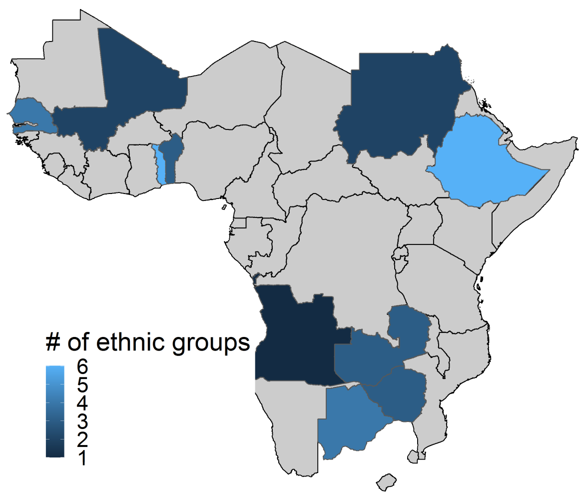
Notes: For external validity, more relaxed fixed effects are applied and therefore more countries are included. Robust standard errors clustered by town. Confidence intervals reported at the 95% level.

Figure A.10: Event-study coefficients with longer post-treatment period



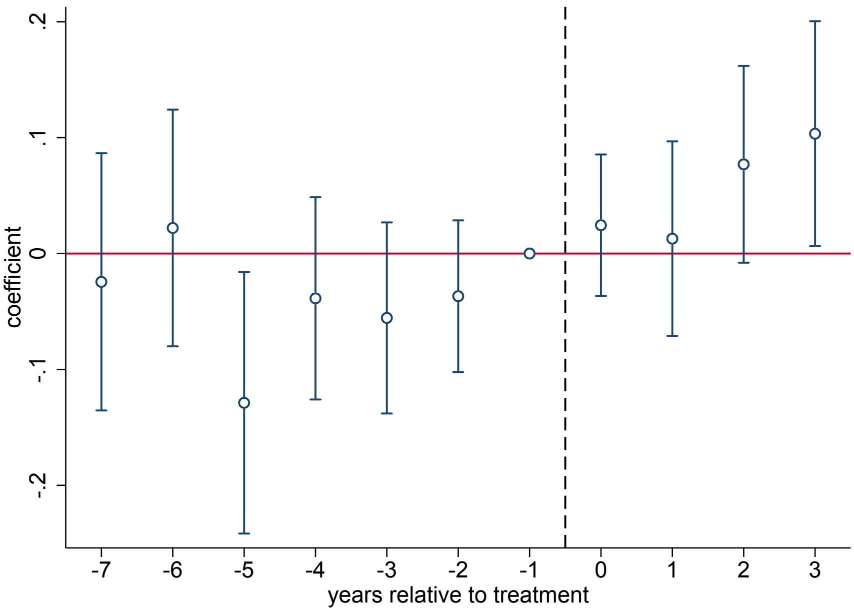
Notes: Coefficients for event study specification with five post-treatment years. Robust standard errors clustered by town. Confidence intervals reported at the 95% level.

Figure A.11: Ethnic groups



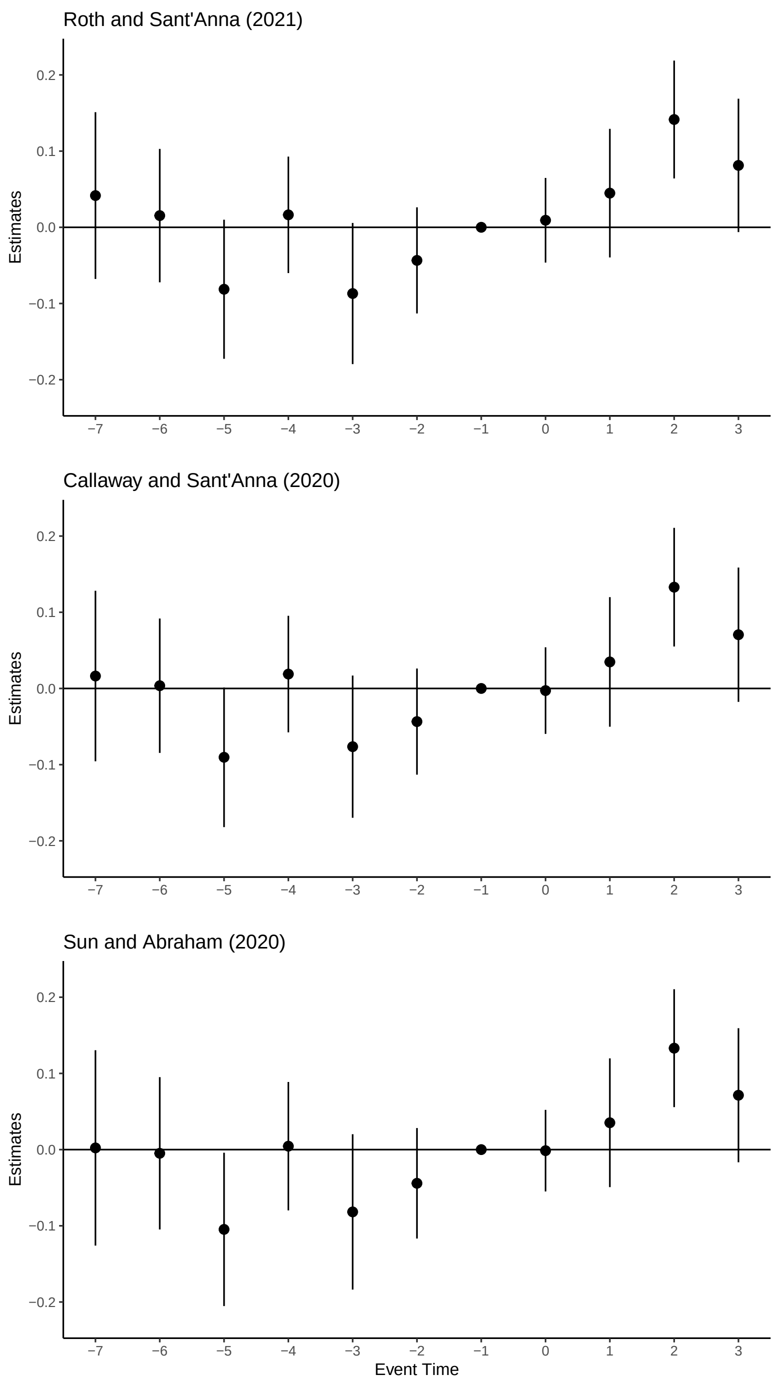
Notes: The figure shows for each SSA country in our analysis how many different ethnic groups were provided with at least one access point before the arrival of an SMC. Brighter blue colors indicate more different ethnic groups. Gray indicates countries not included in our analysis.

Figure A.12: Robustness event study (external validity)



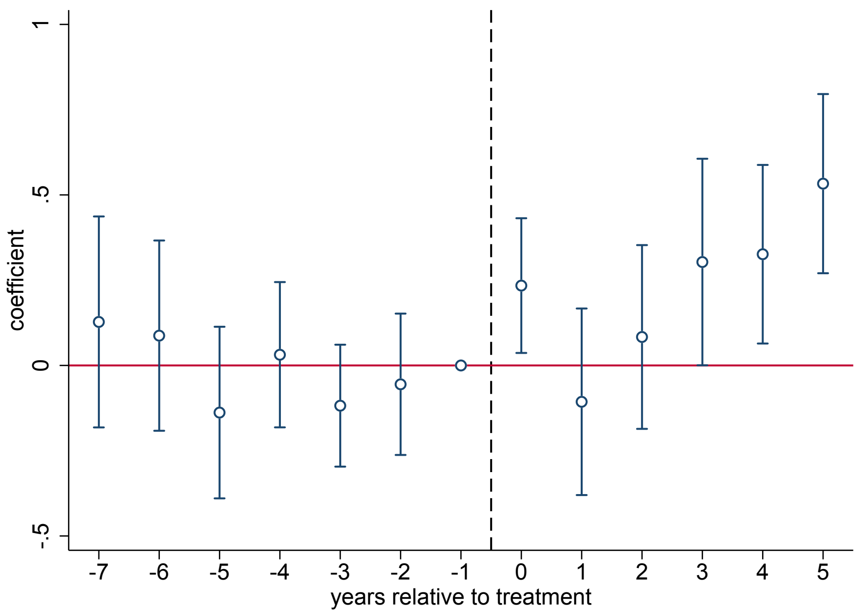
Notes: For external validity, less restrictive fixed effects are applied and therefore more countries are included. Robust standard errors clustered by town. Confidence intervals reported at the 95% level.

Figure A.13: Robustness event study (external validity)



Notes: The figure presents the event study results estimate with the suggested procedures by Roth and Sant'Anna (2021*a*), Callaway and Sant'Anna (2021), and Sun and Abraham (2020). Control variables are not included. Robust standard errors clustered by town. Confidence intervals reported at the 95% level. Results are estimated with the staggered R-package by Roth and Sant'Anna (2021*b*).

Figure A.14: Heterogeneity event study (coastal countries)



Notes: Estimated on a sample including only coastal SSA countries. Robust standard errors clustered by town. Confidence intervals reported at the 95% level.

Appendix B. Additional 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	20,581.39	17,933.61	0.00	8,501.50	16,019.00	30,114.00	82,602.00	220.00
distance to any regional capital	85.45	80.54	1.67	26.98	66.57	129.70	407.28	220.00
distance to the capital	231.65	203.73	1.67	75.81	170.73	355.42	987.20	220.00
distance to the coastline	426.58	307.37	0.00	154.25	427.69	632.76	1,175.48	220.00
distance to next river	56.84	56.65	0.00	15.16	43.99	86.89	411.27	220.00
distance to next port	195.34	272.67	8.23	28.40	74.31	177.52	1,207.12	75.00
distance to the road network	2.58	12.07	0.00	0.00	0.00	0.00	112.57	220.00
distance to the railroad network	57.26	96.44	0.00	0.00	3.80	82.08	440.13	220.00
distance to the electricity grid	13.44	40.58	0.00	0.00	0.00	3.80	350.51	220.00
number of lit pixels	43.35	33.26	1.00	24.00	35.00	53.00	288.00	220.00
summed light intensity	463.04	529.12	21.00	161.50	285.00	530.50	4,026.00	220.00
average light intensity	7.50	5.97	0.26	3.22	5.25	10.07	29.38	220.00
GSM coverage	0.62	0.47	0.00	0.00	1.00	1.00	1.00	220.00
distance to next AP in 2019	1.26	2.52	0.00	0.00	0.00	1.21	9.43	220.00

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

Table B.3: Robustness (population)

population (ln, gpw)	(1)	(2)	(3)	(4)	(5)
post x treated	0.0116	-0.00283	0.0218	0.0124	0.0102
	(0.0183)	(0.00805)	(0.0374)	(0.0277)	(0.0191)
GSM coverage	0.00699	-0.0120	0.0120	0.00104	0.00545
	(0.0119)	(0.00904)	(0.0261)	(0.0287)	(0.0234)
Observations	2,420	1,765	830	610	440
R-squared	0.999	1.000	0.997	0.999	1.000
#countries	10	10	10	10	10
#towns	220	220	220	220	220
share treated	.445	.445	.445	.445	.445
town FE	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓
w/o capital+landing point	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓
w/o population >100k	✓	✓	✓	✓	✓
years	-7 to 3	2000-connection year+3	1995;2000;2005;2010	2000;2005;2010	one pre + one post

Notes: Population measured as the logarithmic sum of of inhabitants per square kilometer within town area and a 2 kilometer buffer from *Gridded Population of the World* (GPW). Columns show different periods. 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 100,000 inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.4: Robustness (population control)

	combined		intensive		extensive	
	(1)	(2)	(3)	(4)	(5)	(6)
post x treated	0.0703**	0.0661**	0.0513**	0.0503**	0.0516*	0.0473*
	(0.0349)	(0.0318)	(0.0231)	(0.0229)	(0.0282)	(0.0244)
GSM coverage	0.0486	0.0461	0.0477**	0.0471*	0.0281	0.0255
	(0.0342)	(0.0336)	(0.0240)	(0.0241)	(0.0263)	(0.0255)
population (ln, gpw)		0.359*		0.0793		0.375**
		(0.190)		(0.113)		(0.148)
observations	2,420	2,420	2,420	2,420	2,420	2,420
R-squared	0.943	0.944	0.947	0.947	0.924	0.925
#countries	10	10	10	10	10	10
#towns	220	220	220	220	220	220
share treated	.445	.445	.445	.445	.445	.445
town FE	✓	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓	✓
w/o capital+landing point	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓
w/o population >100k	✓	✓	✓	✓	✓	✓

Notes: Light intensity in the combined models is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer as in Table 1. The measure for the intensive margin is logarithmic mean light intensity and for the extensive margin logarithmic sum of pixels (coded as 1 if a pixel is lit). GSM mobile coverage is calculated as the percentage share of town area covered with signal. Population is measured as inhabitants per square kilometer within town area and a 2 kilometer buffer from *Gridded Population of the World* (GPW). All specifications are estimated on a sample restricted by excluding landing-point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.5: 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					✓	✓

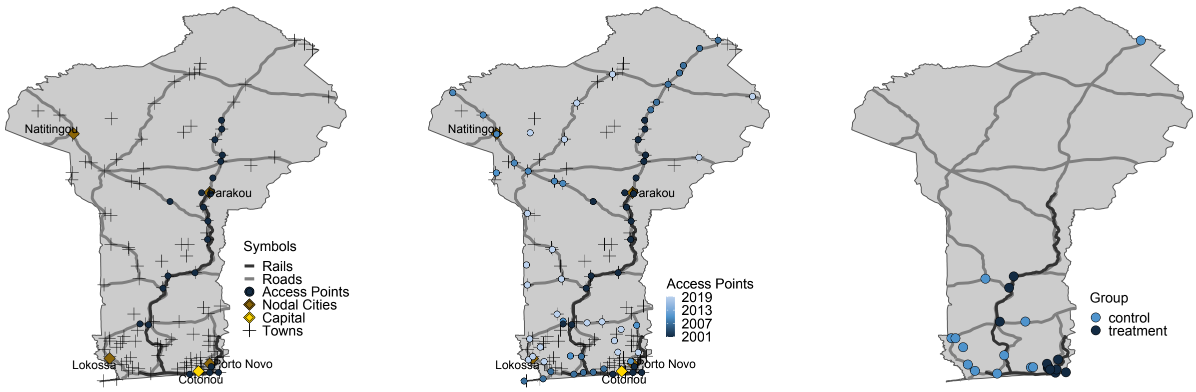
Notes: Access to the electricity grid was aggregated at the town/city level and comes from Afrobarometer (rounds 1 to 4). GSM mobile coverage is calculated as the percentage share of town area covered with signal. All specifications are estimated on a sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix 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 Appendix 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.

Figure Appendix C.1: Rollout in Benin



(1) Access points in 2001 (2) All access points (3) 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 Usage. All transmission happens via Benin Telecoms SA. They offer data transmission networks to mostly commercial clients (banks, hotels, ministries, etc.) in packets.[33] 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 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.

[33] 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.

Appendix 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.[34] 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.[35] Moreover, eleven countries established only in nodal cities access points before Internet became available countrywide.[36] 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.

[34] 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.

[35] 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.

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

Appendix E. Further Robustness Checks

Spatial Correlation. To account for potential spillover effects, we cluster standard errors at the state level for robustness. However, it might be the case that spatial correlation between the towns' location requires correction of the standard errors. Following Conley (1999) we re-estimate Equation (1) correcting standard errors for spatial correlation. Results show statistical significance at the 1% level (Table Appendix E.1).

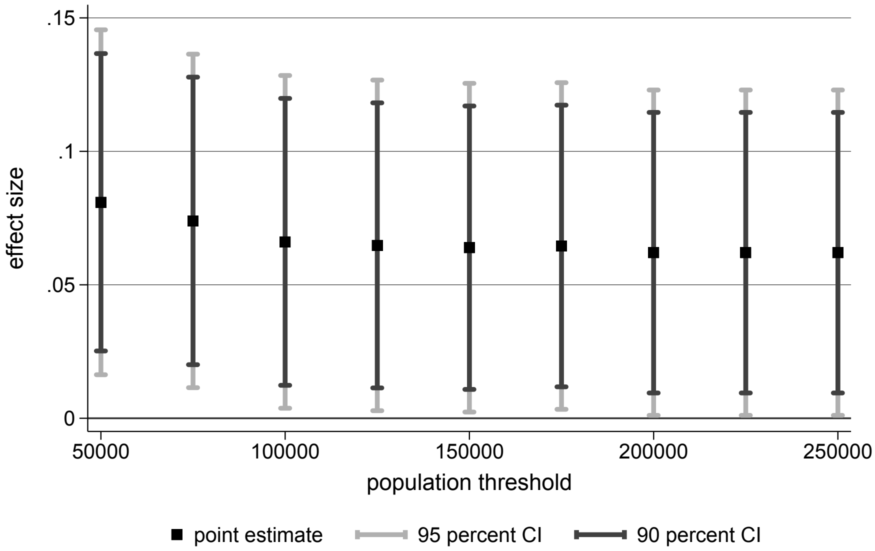
Table Appendix E.1: Spatial correlation

light intensity (log)	(1)	(2)	(3)	(4)	(5)
post x treated	0.0462***	0.0532***	0.0591***	0.0633***	0.0703***
	(0.0179)	(0.0190)	(0.0215)	(0.0226)	(0.0232)
GSM coverage					0.0486**
					(0.0236)
observations	3,190	3,069	2,563	2,420	2,420
R-squared	0.002	0.003	0.003	0.004	0.006
town FE	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓
w/o capital+landing point		✓	✓	✓	✓
w/o regional capitals			✓	✓	✓
w/o population >100k				✓	✓

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer. GSM mobile coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing-point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and all specifications include town and country-year fixed effects. Conley standard errors to account for spatial correlation reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Definition of Nodal Cities. For the main results, we defined all towns with more than 100,000 inhabitants as nodal cities. However, this threshold is chosen arbitrarily. Therefore, when estimating the specification in Column (2) of Table 1, we vary the population threshold as a further robustness check. Figure Appendix E.1 shows that the estimate remains independently of the chosen population threshold. There is a small tendency of a declining effect from a threshold of 50,000 over 75,000 to 100,000 inhabitants.

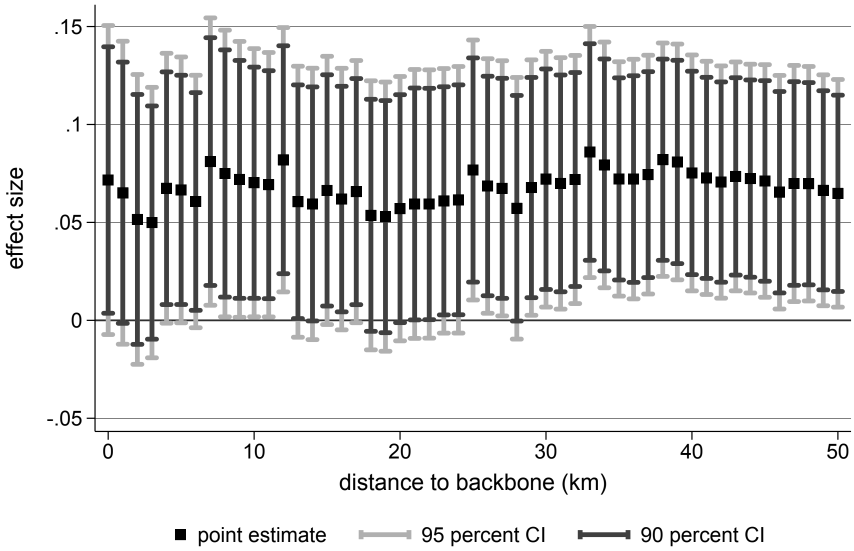
Figure Appendix E.1: Robustness (nodal cities)



Notes: Variation of population thresholds are shown. Coefficients for the specification in Column (2) of Table 1. Robust standard errors clustered by town.

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 (2) of Table 1, we vary this distance from 0 to 50 kilometers as a further robustness check in Figure Appendix E.2. 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. Moreover, one should note that the distance to the access point influences the sample. When allowing for higher distances to define Internet access, the control group shrinks. This means that also some countries drop out of the sample if only treated towns remain. The composition of treatment and control group is not constant as well. 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. Hence, though the treatment group is contaminated in this case, the ATT might find high effects.

Figure Appendix E.2: Robustness (Internet access)



Notes: Variation of the distance to the next access point to the national backbone are shown. Coefficients for the specification of Column (2) of Table 1. Robust standard errors clustered by town.

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 Appendix E.2 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. The last column repeats the main effect estimate.

Table Appendix E.2: Robustness (connected control towns)

light intensity (log)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
post x treated	0.277***	0.267***	0.115*	0.121**	0.111**	0.110**	0.107**	0.0980**	0.0700*	0.0625*	0.0703**
	(0.0999)	(0.1000)	(0.0604)	(0.0538)	(0.0480)	(0.0471)	(0.0457)	(0.0458)	(0.0401)	(0.0378)	(0.0349)
GSM coverage	0.0758	0.0790	0.0947*	0.102**	0.0696	0.0688	0.0687	0.0682	0.0592	0.0520	0.0486
	(0.0642)	(0.0598)	(0.0528)	(0.0480)	(0.0428)	(0.0427)	(0.0422)	(0.0421)	(0.0398)	(0.0382)	(0.0342)
observations	1,265	1,364	1,573	1,650	1,793	1,804	1,837	1,848	2,123	2,233	2,420
R-squared	0.954	0.951	0.950	0.950	0.949	0.949	0.949	0.949	0.944	0.944	0.943
#countries	10	10	10	10	10	10	10	10	10	10	10
#cities	115	124	143	150	163	164	167	168	193	203	220
share treated	.852	.79	.685	.653	.601	.598	.587	.583	.508	.483	.445
town FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o capital+landing point	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o population >100k	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
backbone border	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area. GSM mobile coverage is calculated as the percentage share of town area covered with signal. Sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Longer Post-Treatment Trends and Shorter Pre-Treatment Trends. For robustness, we show that the results do not depend on the chosen window around the treatment year. Table Appendix E.3 shows in Column (2) results for a longer post-treatment period. This reduces the sample size to six countries which were connected that early for the first time that they have at least five post-treatment years before a speed-upgrade SMC arrived. Column (3) reduces the pre-treatment period by two years to five pre-treatment years. The number of countries and towns remains as the data on NTL goes a lot further back in time than the connection year of the first country. In both cases, the point estimate and its level of statistical significance increase in comparison to the baseline specification of Column (1). Column (2) shows that growth rates increase further even five years after the treatment. Column (3) indicates that there is no divergence in the years prior to the treatment. In fact, treated and control group towns have a marginal tendency to converge before the treatment and diverge strongly after the treatment.

Unbalanced Panel. Estimating on a balanced panel has the advantage of not depending on the sample composition close to the period boundaries. Relaxing this restriction, however, allows to estimate on a bigger sample and therefore shows that the results have high external validity. Table Appendix E.3 repeats the estimations from before on an unbalanced sample. Therefore, in the baseline specification, the sample increases by two countries (Column (4)). In Column (5), the sample shrinks only by one country, instead of four as in Column (2) in the balanced sample. In Column (6), the sample size again remains at the higher level. The estimates are only slightly lower in comparison to the balanced sample. Again, it can be observed that the main estimate increases from Column (4) to Column (5), when a longer post-treatment period is applied. The same holds for the comparison of Columns (4) and (6).

Table Appendix E.3: Robustness (estimation window and unbalanced panel)

light intensity (log)	(1)	(2)	(3)	(4)	(5)	(6)
post x treated	0.0703**	0.156***	0.0904***	0.0615*	0.0890**	0.0791**
	(0.0349)	(0.0555)	(0.0343)	(0.0333)	(0.0400)	(0.0327)
GSM coverage	0.0486	0.0743	0.0579	0.0472	0.0544	0.0509
	(0.0342)	(0.0484)	(0.0367)	(0.0314)	(0.0381)	(0.0335)
observations	2,420	1,729	1,980	2,690	2,739	2,196
R-squared	0.943	0.926	0.948	0.945	0.939	0.949
#countries	10	6	10	12	11	12
#towns	220	133	220	247	222	247
share treated	.445	.436	.445	.417	.459	.417
town FE	✓	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓	✓
w/o capital+landing point	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓
w/o population >100k	✓	✓	✓	✓	✓	✓
balanced panel	✓	✓	✓			
pre-treatment years	7	7	5	7	7	5
post-treatment years	3	5	3	3	5	3

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer. GSM mobile coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Parts of SSA as Fixed Effects. Table Appendix E.4 show the results when re-estimating Equation (1) with part-year fixed effects instead of country-year fixed effects. The parts are East, Southern, West, and Central Africa. This specification allows for more countries, as it is not necessary for a single country to have a treatment and control group. On the other hand, a comparison is made within the parts of SSA, such that the growth path of different parts is considered. Again, nodal cities are removed in the stepwise procedure. As before, the estimate increases column by column. It is statistically significant at the 1% level in all specifications. Moreover, the point estimate has a higher level. In our refereed specification (Column (5)), it is .125.

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 Appendix E.5 shows that mobile coverage induces economic growth with a lag of one year. All other lags remain insignificant. However, in all lag specifications, the main effect is robust.

Table Appendix E.4: Robustness (parts of SSA as fixed effects)

light intensity (log)	(1)	(2)	(3)	(4)	(5)
post x treated	0.0904***	0.0997***	0.118***	0.122***	0.125***
	(0.0266)	(0.0283)	(0.0317)	(0.0337)	(0.0337)
GSM coverage					0.0320
					(0.0265)
observations	4,895	4,697	3,718	3,553	3,553
R-squared	0.965	0.951	0.942	0.923	0.923
#countries	16	15	15	14	14
#towns	445	427	338	323	323
share treated	.364	.34	.334	.313	.313
town FE	✓	✓	✓	✓	✓
part x year FE	✓	✓	✓	✓	✓
w/o capital+landing point		✓	✓	✓	✓
w/o regional capitals			✓	✓	✓
w/o population >100k				✓	✓

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer. GSM mobile coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing-point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and all specifications include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table Appendix E.5: Robustness (mobile coverage lags)

light intensity (log)	(1)	(2)	(3)	(4)	(5)	(6)
post x treated	0.0703**	0.0707**	0.0647*	0.0646*	0.0646*	0.0610*
	(0.0349)	(0.0349)	(0.0346)	(0.0343)	(0.0343)	(0.0342)
GSM coverage	0.0486					
	(0.0342)					
GSM coverage (lag 1)		0.0734**				
		(0.0359)				
GSM coverage (lag 2)			0.0235			
			(0.0353)			
GSM coverage (lag 3)				0.0491		
				(0.0335)		
GSM coverage (lag 4)					0.0484	
					(0.0346)	
GSM coverage (lag 5)						0.0384
						(0.0362)
observations	2,420	2,420	2,420	2,420	2,420	2,420
R-squared	0.943	0.943	0.943	0.943	0.943	0.943
#countries	10	10	10	10	10	10
#cities	220	220	220	220	220	220
share treated	.445	.445	.445	.445	.445	.445
City FE	✓	✓	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓	✓	✓
w/o capital+landing point	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓
w/o population >100k	✓	✓	✓	✓	✓	✓

Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 kilometer buffer. Mobile coverage as share of built-up area with the most basic technology (GSM). All specifications are estimated on a sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix F. Mobile Coverage

In our preferred specification, we control for mobile coverage. When controlling for mobile coverage, the main estimate increases slightly and is estimated more precisely. Next, we re-estimate Equation (1) with mobile coverage as outcome variable and follow the stepwise removal of nodal cities. In Table Appendix F.1, Internet access is negatively associated with mobile coverage in all specifications. This means that control group towns catch up to the treatment group with respect to the coverage of the mobile network. This effect is irrespective of controlling for population (Column (5)) and light intensity (Column (6)), which are even jointly not significant as control variables (Column (7)).

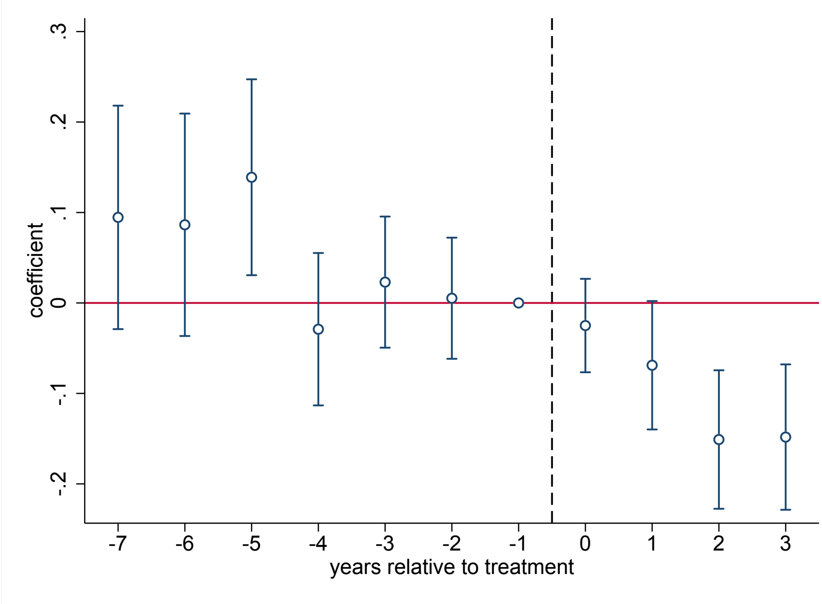
Table Appendix F.1: Mobile coverage

mobile coverage	(1)	(2)	(3)	(4)	(5)	(6)	(7)
post x treated	-0.123***	-0.134***	-0.155***	-0.142***	-0.144***	-0.145***	-0.146***
	(0.0333)	(0.0342)	(0.0364)	(0.0374)	(0.0368)	(0.0374)	(0.0369)
population (ln, gpw)					0.141		0.127
					(0.196)		(0.193)
light intensity						0.0413	0.0393
						(0.0294)	(0.0291)
observations	3,190	3,069	2,563	2,420	2,420	2,420	2,420
R-squared	0.816	0.816	0.816	0.817	0.817	0.817	0.817
#countries	10	10	10	10	10	10	10
#towns	290	279	233	220	220	220	220
share treated	.493	.473	.468	.445	.445	.445	.445
town FE	✓	✓	✓	✓	✓	✓	✓
country x year FE	✓	✓	✓	✓	✓	✓	✓
w/o capital+landing point		✓	✓	✓	✓	✓	✓
w/o regional capitals			✓	✓	✓	✓	✓
w/o population >100k				✓	✓	✓	✓

Notes: Mobile coverage as share of built-up area with the most basic technology (GSM). All specifications are estimated on a sample restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

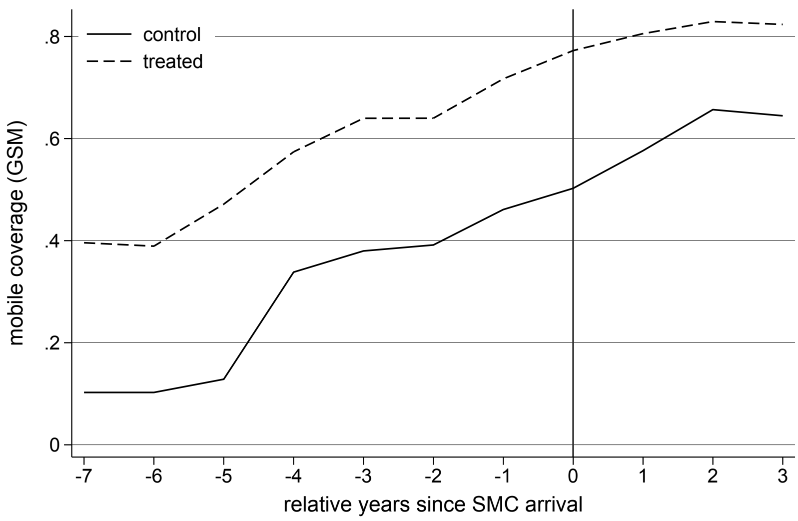
Figure Appendix F.1 indicates that in the early years before the Internet connections, the treatment group had a stronger rollout of the mobile network. While in the years right before the connection, the rollout speed was similar between the treatment and control group, after the connection, the control group caught up with respect to mobile coverage. This interpretation is in line with Figure Appendix F.2 which plots the mobile coverage separated between the treatment group and the control group towns without any fixed effects. While for treated towns the rollout stops slightly above 80 percent coverage shortly after the Internet connection, control group towns continue with the rollout in a linear manner and therefore catch up in the years after the Internet connection was established.

Figure Appendix F.1: Event-study coefficients for mobile coverage



Notes: Coefficients for event study specification of Column (4) of Table Appendix F.1. Robust standard errors clustered by town. Confidence intervals reported at the 95% level.

Figure Appendix F.2: Trends for mobile coverage



Notes: The figure depicts the average growth mobile coverage of the towns in the treatment and control group over a period of eleven years (seven before and three after the treatment year). The measurement is the towns' area covered by GSM technology.