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

Moritz Goldbeck† Valentin Lindlacher‡

Job Market Paper
August 27, 2021

Abstract

We ask if internet availability, even at basic speeds, contributes to overall local economic growth. Analyzing 10 Sub-Saharan African (SSA) countries in the early 2000s, we measure local economic growth of towns using nighttime light data. In a difference-in-differences setting, we exploit quasi-random variation in internet availability induced by sub-marine cable arrivals. Our findings suggest that internet availability at basic speeds leads to about two percentage points higher economic growth of SSA towns in the years after connection compared to a control group of similar towns connected only later.

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

JEL-Codes: O33, O18, R11

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

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

‡Corresponding author. Poschingerstr. 5, 81679 Munich (Germany). University of Munich & ifo Institute, lindlacher@ifo.de.

1 Introduction

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

Lacking legacy infrastructure (i.e., fixed-line telephony networks) to build on makes the provision more complex and costly. Despite these enormous investments, 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 lower adoption rates (World Bank, 2016). On the other hand, the potential of Internet seems particularly high in SSA since alternative ICT like fixed-line telephony is largely absent. It is thus crucial to investigate the effect of internet availability on regional economic development in a developing-country context and especially besides capitals and economic centers.

In this paper, we ask if there is a causal effect of internet availability on overall local economic growth in SSA even at basic speeds. We focus on the initial introduction of Internet in SSA since the early 2000s through the first wave of Internet-enabled sub-marine cables. To investigate if potential individual-level effects matter for the economic development of entire localities, we conduct our analysis at the town level. We observe changes in the spatial expansion of towns (extensive margin) and in their density of economic activity (intensive margin) and interpret these components as pointing more towards population or productivity growth, respectively. Furthermore, we corroborate this suggestive evidence on mechanisms by directly looking at changes in population and in the industry composition.

With the notable exception of Hjort and Poulsen (2019), who find sizable positive individual-level effects of a broadband Internet speed upgrade on employment in SSA around 2010, causal evidence on the economic impact of internet availability in developing countries is surprisingly rare.[1] This is, to the best of our knowledge, the first study investigating the causal effect of internet availability at basic speeds on a measure of overall local economic growth in developing countries: nighttime light satellite data. This measure

[1] See Hjort and Tian (2021) for a comprehensive overview of the current state of literature on the effects of internet connectivity in developing countries.

allows to capture the evolution of towns in 10 SSA countries which get international Internet connection in the 2000s. 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.

To identify the causal effect of internet availability on local economic growth, we exploit quasi-random variation in the timing of internet availability induced by the arrival of the first sub-marine Internet cables (SMCs) in SSA in the early 2000s. This approach was established by Hjort and Poulsen (2019), who use an Internet speed upgrade in SSA induced by SMCs with higher capacities around 2010. In a difference-in-differences setting, we make use of the rollout of access points to the national Internet backbone network to define treatment and control group towns: First, we focus on incidentally connected towns which are located between nodal cities (political and economic centers). These towns are relatively small and are primarily connected due to their fortunate location. Second, we assign treatment status to those towns that already were connected to the national backbone via an access point when the SMC arrives. And our control group consists of similar towns which get internet connection via an access point only later. In a two-way fixed effects model with town and country-year fixed effects, we then compare the growth of towns with Internet access at the time when the first SMCs arrived to a control group of similar towns getting access only later.

We tap two main data sources. First, for local economic growth in SSA towns, we use nighttime light intensity captured by satellites, a well-established proxy introduced by Henderson et al. (2011) and validated by Storeygard (2016) on the city level. To get the local town-level measure, we assign nighttime luminosity to individual agglomerations by linking lit pixels to built-up areas of SSA cities and towns from *Africapolis*. Second, town-level internet availability is determined using data on the location of access points (APs) to the national Internet backbone. Because existing data on the location of APs in SSA only starts in 2009 (Hamilton Research, 2020), we backdate the establishment year of APs to their actual construction year via an extensive review of network deployment projects for each SSA country.

We find that connection to the Internet on average leads to a 7 percent increase in light intensity of SSA towns in the years after connection compared to a control group of similar towns not (yet) connected. Applying the established light-to-GDP elasticity from Henderson et al. (2012) this translates into about 2 percentage points higher economic growth. 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 for to hold, 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), Callaway and Sant'Anna (2020), and Sun and Abraham (2020).

We then differentiate between growth in the number of lit pixels, indicating a spatial expansion of towns (extensive margin), and growth in brightness, which is associated with a higher density of economic activity in the towns (intensive margin). We find that towns with Internet access are becoming both brighter and larger. This provides suggestive evidence that cities with Internet access grow at the intensive margin as well as the extensive margin, i.e. geographically. Controlling for population yields a statistically and economically significant positive coefficient and reduces the main effect only slightly, suggesting that the observed growth is partly driven by migration. Our findings therefore point to the increased local economic growth being attributable to XX to an increase in per capita economic activity and to XX to migration into connected towns. We also find a shift in regional industry shares. In connected regions, the agricultural employment share decreases more rapidly (by around 5 percentage points) and manufacturing and service employment shares increase equally (by around 2.5 percentage points) in comparison to regions getting connected later. This suggests that the increase in economic activity is at least partly a result of a changing industry structure induced by the arrival of the Internet.

To ensure that our results are in fact driven by internet availability induced by SMC arrival, we control for the rollout of mobile-phone coverage and perform placebo tests with other potentially confounding infrastructure, such as roads, railroads, and the electricity grid. Moreover, we test the robustness of our results to alternative assumptions about the variance-covariance matrix. Finally, we extend the sample by relaxing some assumptions to assure the external validity of our results.

The paper proceeds as follows. The related literature is discussed in Section 2. In Section 3, we provide a brief overview of the early Internet in Sub-Saharan Africa and give an example of its rollout in Benin in Section 4. Section 5 lays out the empirical strategy and in Section 6 the data and spatial methods are described. Results are presented in Section 7. Section 8 discusses our results in comparison with related research, while Section 9 suggests policy implications. Section 10 concludes.

2 Literature review

We contribute to two main strands of the literature. First, we add to the broad literature assessing the impact of infrastructure on economic outcomes. For developing countries, Hjort and Tian (2021) give an overview of the effects of internet connectivity, dividing this literature into supply-side and demand-side mechanisms and overall impact of connectivity. Our study is, to the best of our knowledge, the only one investigating the overall impact of internet availability on local economic growth. Most closely related to our work is Hjort and Poulsen (2019), who study the employment effects of broadband Internet on an individual level. 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. We further show that internet availability even at basic speeds and in a setting with low adoption rates benefits regional economic development.

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. He applies the slope of the terrain as an instrument for broadband expansion and detects growth in population, employment, the average wage, and the employment rate. Focusing on labor market effects, Atasoy (2013) discovers that gaining access to broadband services in a county is associated with approximately a 1.8 percentage points increase in the employment rate. This contrasts with Czernich (2014), who finds no effect on the unemployment rate for Germany. On the firm level, Akerman et al. (2015) identify different effects of broadband Internet depending on the skill level of workers on labor market outcomes and productivity for Norway. For Germany, Bertschek et al. (2013) detect broadband Internet effects on the firms' innovation activity, but not on their labor productivity. Colombo et al. (2013) find that the productivity performance of small and medium enterprises in Italy is not influenced by basic broadband applications. However, depending on the sector, advanced broadband applications do influence productivity. On broadband adoption, Grimes et al. (2012) discover an increase in firm productivity.

Related to Internet are mobile phones which are, especially in SSA, the most important and widely used digital device. Jensen (2007) shows that the adoption of mobile phones by fishermen and wholesalers in Kerala, India, led to a reduction in price dispersion. He also finds that the use of mobile phones led to complete elimination of waste and near adherence to the Law of One Price, which increased both 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).

There is also a large body of related literature on the effect of non-digital infrastructure on economic outcomes in developing countries. This literature is more established for developing countries than the literature on digital infrastructure. Assessed infrastructure includes transportation infrastructure (roads, highways, and railways), electrification, and water supply. Most important in our context is Storeygard (2016) as he estimates the effect of transportation costs in SSA cities and towns. He finds that reducing transportation costs increases growth of local economic activity. However, the economic literature has found mixed results. Ghani et al. (2016) find similar results for India, while Banerjee et al. (2020) find growth unchanged but a positive effect on levels, and Faber (2014) even finds a decrease in both in China. Aggarwal (2018) studies the development of paved roads in rural India and finds that paved roads lead to lower prices, higher market integration, and higher use of agricultural technologies. Finally, Donaldson (2018) investigates the effect of railroads in colonial India. He finds that the railroads decreased trade costs and hence increased interregional and international trade, as well as increased real income levels. Electrification is another important issue in developing countries. Its provision demands also high infrastructure expenses. Electrification in rural areas was also analyzed by Dinkelman (2011) for South Africa. She shows that electrification increases female employment. Similar effects are found by Grogan and Sadanand (2013) for Nicaragua. For India, Rud (2012) looked at increased manufacturing output by electricity through the channel of electric pump sets. Lipscomb et al. (2013) find large positive effects of electrification on development through broad-based improvement in labor productivity in Brazil. Finally, Duflo and Pande (2007) show the positive effect of irrigation dams on agricultural production and how these can reduce rural poverty in India.

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 had a positive impact on today's economy and Henderson et al. (2012) who indicated that the hinterland is growing faster than coastal areas and that primate cities are not growing faster than their hinterland, a strand of literature now focuses on the catch-up from secondary to primate cities, with no conclusive results. While Christiaensen and Todo (2014); Christiaensen and Kanbur (2017); Fetzer et al. (2016) show that secondary cities are meaningful to reduced poverty, Bluhm and Krause (2018) show with an adjusted nighttime light data approach that primate cities remain the economic centers.

Economic productivity is typically higher in agglomerations for several reasons, for example, thick local labor markets, knowledge spillovers, low transportation costs, but also due to local provision of amenities (see e.g., Albouy, 2016; Clark et al., 2002; Deller et al., 2008). Thinking of digital infrastructure as a local amenity, our findings indicate that the benefits of internet availability are spatially highly concentrated, accruing to connected locations (towns) only, with important implications for regional inequality both across cities and between cities and rural areas.

3 Background

There are three major components of Internet infrastructure determining availability and bandwidth of Internet in a given location. First, international sub-marine fiber-optic cables connect SSA countries to the global Internet backbone.[2] Second, within-country inter-regional fiber cables form the national backbone. Precondition for internet availability in a location is an access point (AP) to the national backbone. Finally, individual users in a location are reached via the 'last mile' infrastructure.

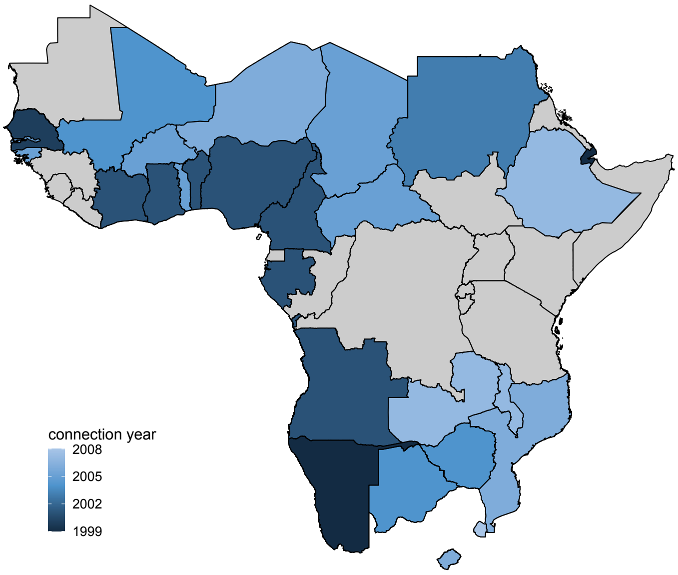
3.1 International backbone: sub-marine cables

Since the vast majority of web pages and applications is hosted on servers located in North America or Europe, almost all African Internet traffic is routed inter-continentally (Kende and Rose, 2015; Chavula et al., 2015). Before the first fiber-optic sub-marine cables (SMCs) landed on African shores, the only way to

[2] 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.

connect to the Internet on the continent was via satellite.[3] While being largely unconstrained by geography and local infrastructure, satellite connection is costly and allows only for very narrow bandwidths. With SMCs—often a joint effort of governments, private investors, and/or multinational organizations—Internet connection was first brought to SSA at a noticeable scale.

Figure 1: Sub-marine cable 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.

As shown in Figure 1, the first wave of internet-enabled SMCs arrived in SSA countries only in the early 2000s. These 'first-generation' cables had the capacity to provide Internet at basic speeds. The biggest of them was SAT-3 and started operating in 2001. It featured landing points on the shores of eight SSA countries on the western coast of Africa.[4] These landing points—typically one per country—constitute the starting point for the respective national backbones (cf. Section 3.2). Until the late 2000s, most SSA countries were connected to the Internet via these 'first-generation' SMCs.[5]

[3]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.

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

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

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

Inter-regional cables are almost always built along pre-existing infrastructure (roads, but also railroads, the electric grid, and pipelines) to minimize construction costs. Even though the goal was to connect nodal cities, in many cases, towns on the route of inter-regional cables got Internet access as well due to their fortunate location between to nodal cities. Our empirical strategy (cf. Section 5) focuses on these on-route towns which get Internet connection because of their location next to an inter-regional cable but are not nodal cities themselves.

3.3 Local transmission: 'last mile' infrastructure

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

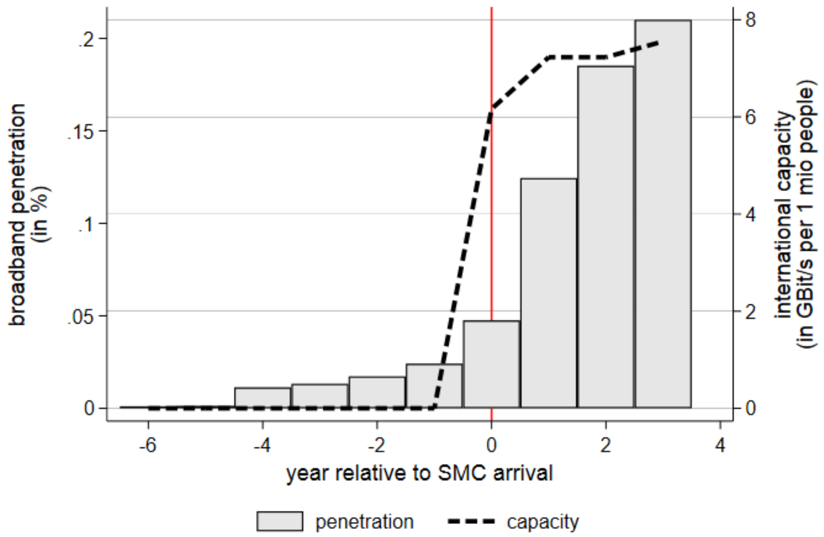
Figure 2 shows how the bandwidth and usage increase in countries that were served by a first generation

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

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

SMC. Though, figures being relatively low in absolute values, the relative change is remarkable. Furthermore, one can notice that both start increasing when the first SMC was constructed (1999 in Senegal) and that both lines increase rather jointly. Although broadband usage is low among the population, broadband penetration in firms might be completely different. Unfortunately, we do not have data on the adoption of broadband in SSA firms. However, 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.

Figure 2: International capacity and broadband adoption



Notes: International capacity is calculated from SMC capacities and assigned to each country by using population-weighted shares. Measures are calculated relative to the arrival of the first SMC in each country and then aggregated via taking the weighted mean. Weights are population size in 2010.

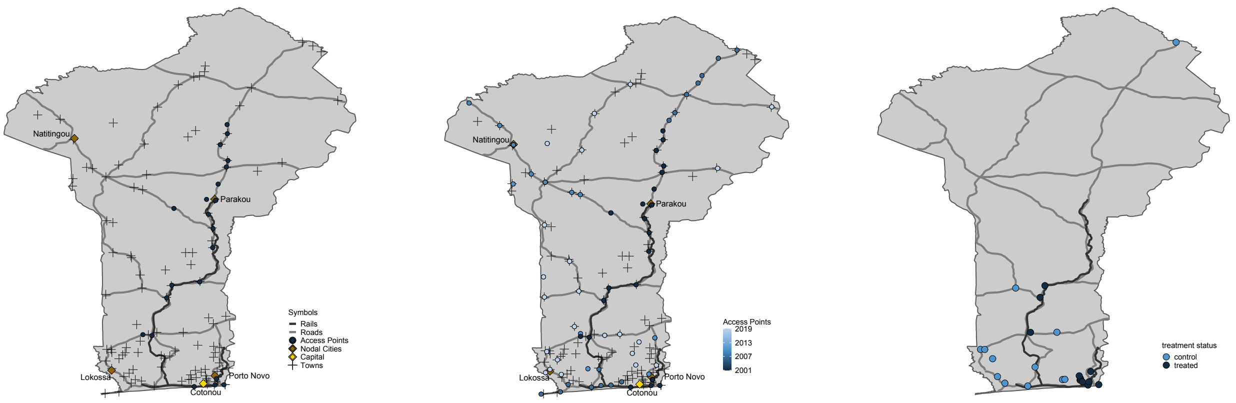
4 The case of Benin

This section gives an overview of one example, how the backbone network was rolled out and how it influenced internet usage. This example is Benin, one of the countries that was connected by the SAT-3 SMC. This cable brought an international connection of 45 mbps (Chabossou, 2007).

The rollout of the national backbone network 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).

Following Chabossou (2007), the SAT-3 SMC landed in Cotonou, Benin's biggest city, the location of the seat of government, and 40 km 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 km optical fibre cable was built 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 3) and connected further regional towns on its way. On the way further smaller towns, for instance Savalou with 30,000 inhabitants were connected. Next, from Parakou connections to the borders of 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 network and access points were still constructed in 2001. The connection towards Burkina Faso and Togo was constructed through Natitingou, the capital of 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.[8] Consequently, Benin Telecoms SA investment in the telecommunications sector peaked in 2001 with more than $80 billion.

Figure 3: Rollout in Benin



(a) APs in 2001 **(b)** All APs **(c)** Treatment / control towns

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.

All transmission happens via Benin Telecoms SA. They offer data transmission networks to mostly commer-

[8] Towns like Nikki, Ségbana, or Banikoara, were connected to the national backbone only in 2019, but had internet access at very low speeds via VSAR satellites from 2001 on.

cial clients (banks, hotels, ministries etc.) in packets.[9] Having grown exponentially, thousands of cybercafes offer wireless 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 cybercafes (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.

5 Empirical strategy

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

To address these endogeneity concerns, we leverage 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 nation-wide, to investigate the effect of internet availability. Following Hjort and Poulsen (2019), we argue that the exact timing of SMC arrival is essentially random.[10] First, for each individual SMC, it is exogenous to national planners when exactly it is put into operation and starts providing (international) internet connection. This is because each SMC typically connects many countries and the connection years are highly uncertain due to unforeseen delays in construction and coordination difficulties among consortium members.[11][12] Second, within SMCs, the connection years are mainly determined by the country's geographical location within SSA because SMCs come from Europe, either through the west passing Spain or Portugal or through the east passing Egypt or the Arabic peninsula, and connect SSA countries according to their location at Africa's coastline. Moreover, landlocked countries get their connection through the backbone network of their neighboring countries and rely therefore on the construction speed there. This construction speed again is exogenous for the respective landlocked country.

[9] 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.

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

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

[12] For example, the cable EASSy was delayed by five years due to coordination difficulties among consortium members (Poppe, 2009).

This enables us to use a difference-in-differences (DiD) design. We compare towns that already have access to the national backbone when the SMC arrives to a control group of similar towns getting an access point (AP) in later years (first difference) before and after the arrival of an SMC (second difference). The definition of towns in the treatment and control group is depicted in Figure A.2. Importantly, all towns in our analysis get connected eventually, mainly because of their favourable location between nodal cities. Hence, we also exclude towns that are still waiting for an internet connection today. Moreover, towns being connected in the three years under observation after the arrival of the SMC are excluded from the control group as well as they would contaminate the control group but do not get the full treatment and would thereby confound our analysis. We focus on incidentally connected towns, i.e., towns close to an AP that are not (endogenously connected) nodal cities (cf. Section 3.2) . We focus our analysis on these towns because they are very similar to each other in key characteristics such as size and infrastructure. Therefore, we exclude nodal cities: the landing point, the capital, regional capitals, and economic centers (cities with a population of more than 100,000 inhabitants). Since national backbones are self-contained and (coastal) countries have own landing points each country has a specific treatment date. Therefore, depending on the nation-wide connection date, towns in different countries are treated at different points in time. For robustness, we vary the threshold of 100,000 inhabitants as the definition of an economic center, we include towns to the treatment group by the year they get an AP, and vary the last possible year of connection for the control group, such that treated towns are not compared with very late connected towns.

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

$$y_{ict} = \beta_0 + \beta_1 \left(smc_{ct} \times access_{ic} \right) + \boldsymbol{X}_{ict} \beta + \alpha_{ic} + \delta_{ct} + \varepsilon_{ict} \tag{1}$$

where y_{ict} is economic growth of town i in country c in calendar year t as proxied by nighttime light luminosity measured by satellites (cf. Section 6). The dummy variable smc_{ct} indicates if country c has access to an SMC in calendar year t. The variable $access_{ic}$ is one if town i in country c is located within 10 km distance to an AP that was established in the year of SMC arrival or before. Contrary, the indicator is zero if town i in country c is located close to an AP that was established in the years after the SMC arrival. Thus, the interaction term $smc_{ct} \times access_{ic}$ indicates internet availability in town i in country c in calendar year t. The coefficient of interest is β_1. It captures the effect of internet availability on local economic activity. $\boldsymbol{X}_{ict}$ contains time-varying control variables, such as mobile-phone coverage. We include two types of fixed effects into the model. Time-constant differences across towns are captured by town fixed effects α_{ic}. Differences across calendar years common to all towns within a country are absorbed by country-year fixed effects δ_{ct}. Note that this allows for country-specific time trends and variations in satellite sensor quality over years. ε_{ict} is an error term. Like in many other DiD applications, our panel data are serially correlated in the time dimension. Hence, we use cluster-robust standard errors whereby we cluster at the town level.

The key identifying assumption for this DiD model is that treatment and control towns would have evolved similarly in the absence of treatment (parallel trends assumption). This assumption cannot be tested. Its plausibility can, however, be examined by testing for pre-treatment differences in time trends between the treatment and control group. Therefore, we look at the dynamic impact of internet availability on local economic activity using an event study design:

$$y_{ict} = \beta_0 + \sum_{j=\underline{T}}^{\overline{T}} \beta_{1j}(t_j \times access_{ic}) + \boldsymbol{X}_{ict}\beta + \alpha_{ic} + \delta_{ct} + \varepsilon_{ict} \quad (2)$$

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

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, 2019; 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 trend assumption in the presence of treatment effect dynamics. This cannot happen when applying more rigorous fixed effects. If country-year fixed effects are applied, there is only one treatment (one arrival date of the SMC). However, we can relax the fixed effects. To account for the following threat to identification, we also perform alternative approaches proposed by Callaway and Sant'Anna (2020), Roth and Sant'Anna (2021), and Sun and Abraham (2020). For instance, Callaway and Sant'Anna (2020) 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.

6 Data and spatial methods

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

(APs) to the national fiber-cable backbone network. Moreover, we use data on towns' built-up area, merged with characteristics, such as administrative status and population, and on 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). This section describes the data sets we use, the processing steps we apply, and shows descriptive statistics relevant for our analysis.

6.1 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.[13] 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.

Since geographically and chronologically granular data on economic activity in SSA is lacking—especially for the period we investigate—we deploy nighttime light (NTL) satellite data. This data measures human-caused NTL emissions in a geographically high resolution and on a yearly basis. There have been two major programs that collected NTL data. First, NTL data was collected in the *Defense Meteorological Satellite Program* (DMSP) *Operational Linescan System* (OLS) between 1992 and 2013. Second, the follow-up program *Visible Infrared Imaging Radiometer Suite* (VIIRS) on the *Suomi National Polar-orbiting Partnership* satellite started in April 2012. The instruments of DMPS-OLS satellites measure light intensity on an integer scale from 0 to 63 with pixels covering 30 arc-second grid cells (an area of .86 km^2 at the equator). The data is then combined to yearly composite images. With VIIRS, both the spatial (pixels are smaller) and radiometric resolutions (both dark and bright spots are recorded better) as well as the temporal resolution have been improved.

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 DHS cluster locations as well as for grid cells of roughly 50×50 km.

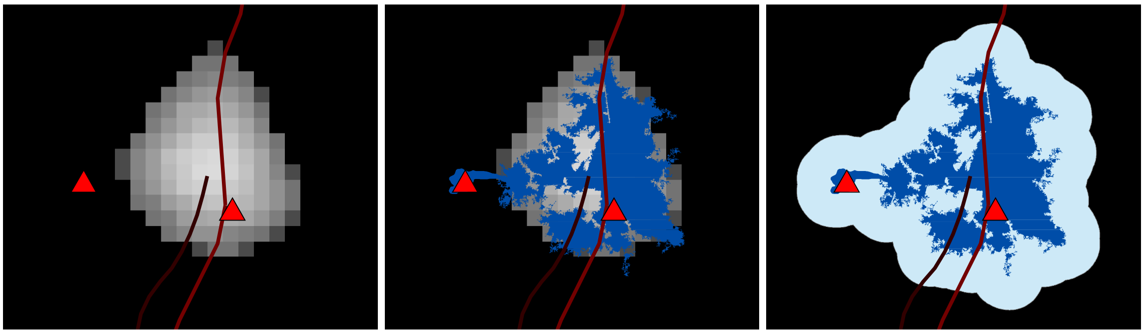
We use the harmonization of the two sources by Li et al. (2020) to get consistent yearly NTL data from 1992 to 2018. Though, we do not need that long time frame for our main results, it allows us to investigate medium-term effects. Moreover, we exploit the advantage of this procedure that noise from aurora, fires, boats, and other temporal lights are excluded. Another advantage of this data is that the DMPS-OLS data has been inter-calibrated globally from 1992 to 2013 as well, making it temporally consistent. To harmonize the two satellite programs, Li et al. (2020) recalculated the newer and better VIIRS data to the (spatial and radiometric) resolution of the original DMPS-OLS data. Therefore, the lower DMPS-OLS resolutions apply.

[13] https://africapolis.org

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 2km to the built-up area, such that the growth of light emissions in the extensive margin is properly captured.[14] 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 4 shows for Parakou, Benin, its NTL emission, built-up area, and infrastructure. A road connecting Parakou with its neighbouring cities (blue line) and the access points (red triangles) constructed in 2001 are shown in all panels. Panel (a) shows moreover the nighttime lights for the year 2001, where a brighter gray reflects higher light intensity. Panel (b) adds Parakou's built-up area from *Africapolis*. It shows that the blurring of the nighttime light data exceeds the built-up boundaries. Therefore, we draw a buffer of 2 km around the area (shown in Panel(c)). This allows us to take all light emissions into account.

Figure 4: Data example: Parakou, Benin



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

Notes: Panels (a)-(c) show our data for Parakou, Benin, in 2001. Panel (a) shows the constructed access points (red triangles) and nighttime lights for the year 2001 (connection year of Benin). The dark red line represents a major road connecting Parakou with its neighbouring cities and the darker red line the railway connection to the south. The black-to-white scale indicates light intensity, with brighter colors reflecting higher light intensities. Panel (b) adds its built-up area from *Africapolis* (shown in green). Finally, Panel (c) shows in blue a 2 km buffer around the built-up area.

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

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

economic activity (intensive margin).[15]

6.2 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. The data comes primarily from *Submarine Cable Map*.[16] Table B.1 shows the country-specific connection dates of all SSA countries that were connected before 2009. If the country was connected via an SMC, the SMC landing point is stated as well. We geo-coded this point to merge it to the respective built-up area. The connection year of through neighboring countries was also mainly taken from *Submarine Cable Map*. In the last column, the upgrade year of the next SMC is shown. These SMCs had a lot higher capacities and landed in SSA between 2009 and 2012. The geo-location of the APs for the national fiber-cable backbone comes from *Africa Bandwidth Maps* and are mapped in Figure A.1.[17] The database contains the most comprehensive set of APs for Africa, covers the period starting from 2009, and is updated on a yearly basis. The data is directly sourced from the network operators. As APs existing in 2009 were largely established earlier, we conducted an extensive review of network deployment projects for each country. Thereby, we determined the construction years of the APs 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 APs were built in the year the SMC landed, which is still sufficient for our approach. This makes it possible to identify which towns already were connected to the national fiber-cable backbone when the first wave of SMCs arrived. We match APs to towns via their geo-location: First, we calculate the distance between the towns' border and the closest AP. Then, we assign a national fiber-cable backbone connection to towns within a distance of less than 10 km.

To date, there are 2.708 APs in SSA countries. About half of them were constructed since 2013. Especially in bigger cities, more than one AP is usually built to account for the limited capacity of each AP. In 2019, for example, although 189 new APs were constructed, only 27 new cities and towns were connected. In total, around 900 cities and towns have an AP close by.

6.3 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 SSA countries, which are listed in Table B.1. Among the first countries are Djibouti, where an SMC landed in 1999, Namibia, which was connected by a trans-national fiber cable from South Africa in 1999, and Senegal, which got a connection to the Internet in 2000 with an individual-country SMC. In 2001, seven more countries were connected by a

[15] As specified in Section 5, we apply the logarithm of each outcome measure.
[16] https://www.submarinecablemap.com
[17] http://www.africabandwidthmaps.com

single SMC, the SAT-3 cable. In the following years until 2008, 17 more countries got an SMC connection or were connected through their neighboring countries.

However, not all countries that were connected until 2008 had built 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 connection right after the connection. This reduces the number of countries in our analysis to 23.[18] Moreover, 11 countries built at least one AP before SMC arrival, but only in nodal cities.[19] Finally, we cannot consider Namibia in our analysis because it did not construct APs after getting Internet access. 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. For example, the first connected country in our sample is Senegal, which was connected in 2000. This leaves us with 7 pre-treatment years. Hence, on the left with a truncation at 7 years before the treatment year, we have no data loss. This limits our pre-treatment comparison to a period of 7 years. Malawi and Mozambique only have 2 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 3 post-treatment years leaves us with a sample of 10 countries.[20] A longer post-treatment period would shrink the sample further. Therefore, we estimate the main specification with 3 post-treatment years. For robustness, we will relax these restrictions.

6.4 Descriptive statistics

For the estimation sample in the balanced panel, Figure A.3 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 10 countries, five are coastal and five are landlocked.[21]

We focus our analysis on mid-sized towns. From 510 agglomerations, for which NTL data is detected in each year, in the ten countries in the estimation sample, 143 were connected to the Internet via an AP before the country was connected via SMC or a neighboring country. Therefore, they are part of our treatment group. 70 of these agglomerations are nodal cities. Another 147 towns got an AP in the subsequent years

[18] Central African Republic has not yet built a national backbone infrastructure. In Lesotho, the APs were built in 2009 three years after being connected through South Africa. In Djibouti, the first APs were built in 2007, which is 8 years after the first SMC connection. Nigeria built its first APs in 2003, which is 2 years after the arrival of the first SMC.

[19] Guinea-Bissau, Lesotho, and Swaziland built all APs until today only in nodal cities.

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

[21] Sudan is a special case as 10 cities are in the control group but only one city is in the treatment group. We account for that by...grouping countries? robustness leave-out?)

and are therefore in the control group. The remaining 220 agglomerations are still not connected. Further 102 towns were connected in the three years after SMC arrival and are therefore not considered as they would confound our control group.

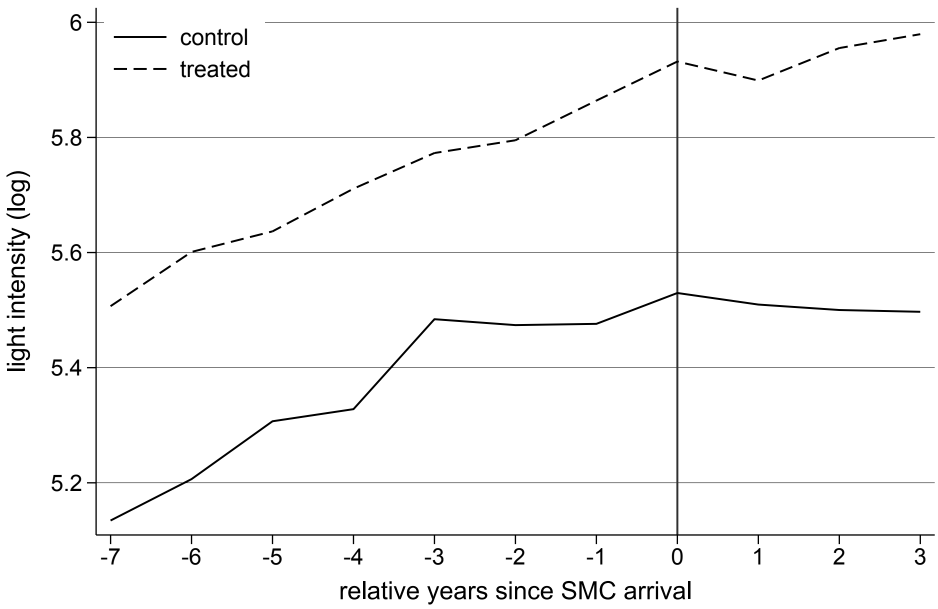
Figure A.4 compares the average size of cities and towns by the year they get an AP, relative to the SMC connection year. In the early years, until the SMC arrives, many nodal cities are connected besides the towns in the treatment group. While connected nodal cities are bigger on average in the early years and decline in their size in subsequent years, towns in the treatment group only have a population of around 20,000 inhabitants on average. Control towns, which are connected in the subsequent years after the observation period of 3 years, vary for all subsequent years around a population of 20,000 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 as there clearly can be identified for nodal cities. This suggests that treated towns are not selected into treatment because of their population size. Importantly, nodal cities are still connected in further years after the arrival of the first SMC. Showing that the rollout continues to other parts of the country as well. Their size decreases after the first two years as capital cities are usually connected first and a lot bigger than other nodal cities. Nonetheless, their size is still greater 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). 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 30,00 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 distance to the closest AP is with 9.43 km smaller than 10 km. On average, this distance is a lot smaller with 1.26 km. More than half of the towns have an AP even within the built-up area and most cities have it within 2 km (1.21 km 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 km (3.8 km 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. The variables come either from *Natural Earth (NE)*, *OpenStreetMap (OSM)*, and *Africa Infrastructure Country Diagnostic (AICD)*. Therefore, the data is time-invariant and not dated before the treatment but was collected in 2007 (AICD), 2010 (NE), or more recently (OSM).

Before presenting the estimation results, we show the development of cities and town over time. We use the main outcome measure, log light intensity, averaged over the different groups but do not include any fixed effects or controls. We focus on the treatment and control group only (Figure 5). Figure 5 shows that in the early pre-treatment years both groups grew with similar rates. In more recent years before the treatment, this growth stagnated. This stagnation holds on for the control group, while the treated cities grew again in the second and third year after the treatment. One reason for the stagnation might be a general decrease in economic growth induced by the increasing oil price from the mid-2000s to 2008, which was followed by the international financial crisis. Another reason could be measurement error in the date of the AP construction year. This measurement error would be biased towards earlier years and might explain the increase in t-3 and t-2. Finally, at the end of the observation period, the before rather small gap between these towns grew by about .1 from about .4 to about .5 on a logarithmic scale.

Figure 5: 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 11 years (7 before and 3 after the treatment year). The measurement is the logarithm of light intensity.

7 Results

7.1 Main effects

We estimate the effect of internet availability on local economic growth. We are particularly interested in the effect of early internet availability brought by the ‘first-generation’ SMCs. We estimate a linear model on a balanced panel by difference-in-differences. Besides town fixed effects, country-year fixed effects are used and standard errors are clustered at the town level. We measure economic activity by the logarithm of the sum of light intensities as the main outcome. We discuss other outcomes like mean light intensity (intensive margin) and sum of lit pixels (extensive margin) in Subsection 7.3.

Table 1 shows the main results. In Column (1), the most basic specification, nodal cities and the landing point are still included. We then stepwise eliminate nodal cities 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. In Column (3), we also remove regional capitals. In Column (4), we further remove cities of more than 100,000 inhabitants. Finally, we add as a control variable the share of the towns’ area with GSM mobile network coverage (Column (5)) to account for other (tele-)communication technologies.

In line with our expectations, we find an economically and statistically positive effect of Internet at basic speeds on local economic growth. From Columns (1) to (5), the nodal cities are dropped column by column. Therefore, the number of cities we estimate on, decreases from 290 to 220 towns. Hence, 70 nodal cities are dropped. As only nodal cities are removed, which mostly have an AP, the share of treated towns on the overall population decreases from 49.3 to 44.5 percent. We find a slightly increasing effect size, that is 5 percent for the whole sample and increases to 7 percent for our preferred specification without nodal cities. For bigger cities, the estimate is less precisely estimated as other factors might influence economic development. For these cities, and especially for capital cities, no counterfactual exists in the control group. Especially, nodal cities are bigger on average, both in comparison to other treatment and control group towns. Hence, they differ in levels and therefore possibly have lower growth rates.

In our preferred specification (Column (5)), towns which were connected to the Internet in the year of an SMC arrival become 7 percent brighter than towns without internet availability. This finding supports our initial claim that towns which get incidentally connected to the Internet growth faster in comparison to otherwise comparable towns. The mobile coverage control is statistically insignificant and smaller in size. It makes the estimation more precise as it controls for differences in another telecommunication technology. As it slightly increases the point estimates of the main effect, we are not worried that the main effect is in fact transported through mobile coverage. But we will discuss this issue later in Section D.

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

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

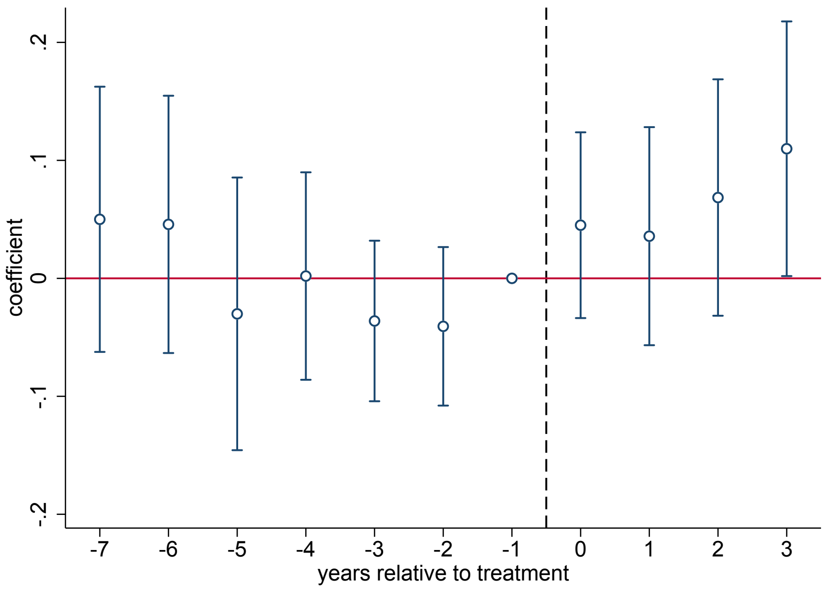
	(1)	(2)	(3)	(4)	(5)
post x treated	0.0462	0.0532*	0.0591*	0.0633*	0.0703**
	(0.0286)	(0.0302)	(0.0328)	(0.0344)	(0.0349)
GSM coverage					0.0486
					(0.0342)
Observations	3,190	3,069	2,563	2,420	2,420
R-squared	0.971	0.961	0.958	0.943	0.943
#countries	10	10	10	10	10
#cities	290	279	233	220	220
share treated	.493	.473	.468	.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 km buffer. GSM mobile-phone coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing-point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and all specifications include town and country-year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

7.2 Event study

For our preferred specification (Column (5)), Figure 6 presents event-study coefficients on the same periods we have a balanced panel. We cannot estimate this effect for a longer time period as second-generation SMCs arrive in the years after the initial SMC arrival to upgrade internet capacity and speeds. From this dynamic perspective, there is no evidence for a potential fading out of the effect. 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 SMC arrival, conditional on fixed effects. In the year of the SMC arrival, the point estimate turns positive but remains statistically insignificant until year three. In the years after SMC arrival, the point estimates are between .05 and .1 and have a slight tendency to increase to nearly .15 in year three after the connection, when the effect gets statistically significant at the 5% level. The increase of the effect over time is inline with the expectation that internet adoption takes time to increase after the connection and that the effect could be induced by spillovers and not only by the adopting firms directly.

Figure 6: Event-study coefficients



Notes: Coefficients for event-study specification of Column (5) from Table 1. Robust standard errors clustered by town for 95% confidence interval reported as bars.

7.3 Mechanism

Next, we analyse what drives the effects. Therefore, we first investigate the effects of the intensive and extensive margin of regional growth and the relation to population growth. Second, we show how internet availability affects the industry composition and that these changes might be a channel through which internet availability affects regional economic growth.

7.3.1 Population, Intensive, and Extensive Margin

Next, we investigate the intensive and extensive margin of regional development. Therefore, we take the mean luminosity (Table 2, Columns (3) and (4)) and the sum of lit pixels (Table 2, Columns (5) and (6)) as outcome measures. The main outcome, the sum of luminosity, is shown again (Table 2, Columns (1) and (2)). All outcomes are logged. In odd columns, we estimate the specifications as in Column (5) of Table 1. Hence, the sample is identical as well in all columns. In even columns, we add the population as a control. By doing so, we control for which part of the effect is driven by economic growth and which by growth in population.

The observed increase of towns having internet available in light intensity can thus be explained by both an increase in brightness (Columns (3) and (4)) and in size (Columns (5) and (6)). While towns' light intensity increases by 7 percent, their brightness increases by 5 percent and their size increases by 5 percent as well. 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 and shows slightly higher standard errors at a slightly lower point estimate. Adding population as a control, the main effect remains robust and statistically significant. Population has an elasticity of .4 on light intensity. The coefficient of the population control is positive and statistically significant at the 10% level (Column (2)). Therefore, the effect of internet availability is mainly inducing an economic development and not showing a population growth. While the main effect is similar for the intensive and the extensive margin, the effect of population is different. For the intensive margin, the effect of population is smaller and statistically insignificant (Column (4)). On the other hand, the population effect is positive and statistically significant for the extensive margin (Column (6)). Again, the main effect remains robust when adding population as a control. This suggests that towns increase in their size due to an increase in population. But that they also increase in their light intensity, which reflects an increase in productivity.

Table 2: Population, intensive, and extensive margin

	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
#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 in the combined models is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area and a 2 km 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-phone coverage is calculated as the percentage share of town area covered with signal. Population is measured as inhabitants per square kilometer within town area from *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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

7.3.2 Industry

Next, we can show how the regional industry composition changes when the Internet becomes available. We use survey data from IPMUS International, a collection of census microdata, to calculate the share of jobs in certain industries (agriculture, manufacturing, and services). As we rely on survey data, we focus our analysis on changes in the industry shares. In the Appendix, we will also discuss changes in levels, though they might depend more strongly on the survey selected individuals. The data contains surveys for 20 SSA countries from which 12 countries were connected by an SMC in 2008 or earlier. For most countries, there are two levels of subnational regions available to which we can merge the built-up areas. However, the survey does not cover all regions. For each country, only a very limited number of years is available as for most countries data is available with a frequency of 10 years. Therefore, we estimate a long difference with one survey year before the arrival of the SMCs and one afterwards. However, not all countries have survey data from a pre-treatment year. On the other hand, 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. The treatment is defined as above. Hence, we remove nodal cities and define a region as treated if a a town in that region has an AP close by by the year an SMC lands.

The results in Table 3 indicate that internet availability shifts jobs from agriculture to manufacturing and services. In Columns (1) to (3), we estimate on five countries and 91 towns. The share of jobs in agriculture declines by around 5 percentage points. In contrast, job shares in manufacturing and in services increase each by around 2 percentage points. In Columns (4) (6), we estimate on the countries from our main analysis. Here, the decrease in agriculture jobs is with around 6 percentage points even bigger. Consequently, the substitution to manufacturing and services is with 3 percentage points greater as well. This is inline with Hjort and Poulsen (2019), who finds an increase in employment in Ethiopian manufacturing firms and an increase in net firm entry in services.

Table 3: Employment shares by industry

	(1) agriculture	(2) manufacturing	(3) service	(4) agriculture	(5) manufacturing	(6) service
post x treated	-0.0466**	0.0234**	0.0231*	-0.0594***	0.0257***	0.0337**
	(0.0183)	(0.00898)	(0.0129)	(0.0211)	(0.00767)	(0.0162)
Observations	182	182	182	124	124	124
R-squared	0.972	0.910	0.976	0.973	0.948	0.974
#countries	5	5	5	3	3	3
#cities	91	91	91	62	62	62
share treated	.363	.363	.363	.419	.419	.419
City FE	✓	✓	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓	✓	✓
w/o capital+landing point	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓
w/o population >100k	✓	✓	✓	✓	✓	✓
late adopters	included	included	included	excluded	excluded	excluded

Notes: Regional industry composition comes from IPMUS International survey. 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

7.4 Robustness

7.4.1 Placebo test with nondigital infrastructure

In Table 4, we include additional controls. As was shown in Figure 3, APs usually are constructed along other infrastructure. In the case of Benin, this infrastructure is railroads and major roads. But also more generally, to save costs fibre cables are rolled out along existing infrastructure (cf. Section 3.2). We control for this other infrastructure to rule out that towns closer to this nondigital infrastructure grow faster when the SMC arrives, irrespective of whether they are in the treatment or the control group. Unlike the mobile network coverage (GSM), we do not have time-varying data on other infrastructures. We therefore intersect these controls with the post dummy for the time after the arrival of the SMC and construct placebo treatments. We follow two approaches to define placebo treatments. First, we control for the linear distance of the town to the next greater (paved) road (Column (2)), to the next railroad (Column (4)), and to the next electricity grid (Column (6)). Second, we repeat this exercise with a connection dummy for each infrastructure, indicating whether the distance is below 10 km (as we defined treated towns with APs) (Columns (3), (5), and (7)). Finally, we include both approaches jointly (Columns (8) and (9)). Column (1) shows our preferred specification without further infrastructure controls as baseline.

In each case, the estimate of the main effect remains close to .07 and statistically significant. In the case of roads, we loose slightly precision. Nevertheless, controlling for all infrastructures jointly, the estimate is again statistically significant at the 5% level and slightly higher than in the baseline. The estimates of the placebos are all statistically insignificant and very small in their economic significance. Moreover, the point estimate of mobile coverage remains robust when including any infrastructure placebo.

7.4.2 Further robustness

Different fixed effects As explained above, we apply country-year fixed effects to account for country-specific growth paths in their economies. For robustness, we re-estimate Equation (1) including the classical two-way fixed effects: towns and calendar years. This specification is less demanding in the set of fixed effects. The estimate presented in Column (2) of Table B.3 remains robust. However, the estimate of the control variable GSM coverage turns statistically highly significant. In Subsection 7.5, we estimate further models with classical two-way fixed effects which allow for a bigger sample containing more countries and estimate novel DiD / event-study designs (Roth and Sant'Anna, 2021; Callaway and Sant'Anna, 2020; 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 AP might generate further spillover effects in the town's surrounding area. In theory, it could also be the case that one AP serves more than one town, if both towns are located within 10 km to the AP. In our preferred specification, we cluster the standard errors at

Table 4: Robustness (competing infrastructure)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
post x treated	0.0703**	0.0631*	0.0647*	0.0766**	0.0942**	0.0753**	0.0733**	0.0729**	0.0903**
	(0.0349)	(0.0345)	(0.0349)	(0.0371)	(0.0376)	(0.0357)	(0.0356)	(0.0368)	(0.0375)
GSM coverage	0.0486	0.0449	0.0477	0.0482	0.0463	0.0481	0.0480	0.0440	0.0450
	(0.0342)	(0.0341)	(0.0341)	(0.0342)	(0.0340)	(0.0342)	(0.0343)	(0.0342)	(0.0340)
Road network (in 100km)		-0.197						-0.198	
		(0.215)						(0.211)	
Road network (dummy)			0.0928						0.104
			(0.0768)						(0.0757)
Railroad network (in 100km)				0.0112				0.0102	
				(0.0213)				(0.0222)	
Railroad network (dummy)					-0.0624				-0.0657
					(0.0385)				(0.0406)
Electricity grid (in 100km)						0.0415		0.0335	
						(0.0495)		(0.0526)	
Electricity grid (dummy)							-0.0227		-0.00718
							(0.0391)		(0.0405)
Observations	2,420	2,420	2,420	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	0.943	0.943	0.943
#countries	10	10	10	10	10	10	10	10	10
#cities	220	220	220	220	220	220	220	220	220
share treated	.445	.445	.445	.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 km buffer as in Table 1. GSM mobile-phone coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and all specifications 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the town level as the treatment, the access point construction, is occurring there. To take this into account, we apply a higher level of clustered standard errors as robustness. We re-estimate Equation (1) correcting standard errors for clusters at the level of states. Column (3) presents the estimates with a higher level of clustered standard errors. The standard error of our variable of interest increases only very slightly.

Linear time trends In Column (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 even increases slightly to .10 and is also statistically significant at the 5% level. Thus, we can show that after the SMC arrival, the connected towns grow in a nonlinear manner.

APs during the post-observation period Thus far, we have tested the robustness of our results to alternative assumptions about the variance-covariance matrix. In Column (5), we include towns that were connected only within three years after the SMC landed. In our main specification, these towns were excluded as they neither belong clearly to the treatment nor the control group and would thus confound our analysis. Including them, we count them as treated when they get an AP and therefore have internet available right away. Thus, for these towns, the treatment year is equal to the construction year of the AP and not of the SMC connection year. With this approach, we can add them without confounding our analysis. However, for these towns the assumption of the exogeneity of the treatment might not be valid anymore. Therefore, they are not included in the main specifications. However, for robustness, we investigate how including them affects our results. Moreover, within our 10 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.

No buffer In Column (6), we remove the 2 km 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 going back to the smaller area, we might lose some pixels at the town's 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 city growth does not predominantly happen at the city border. In fact, the estimate not only remains, but increases slightly to .08 and strongly statistically significant (at the 1% level).

7.4.3 Ethnic favoritism

A further concern could be that in the rollout specific ethnic groups were favored. Though, the exogenous shock comes from the arrival year of the sub-marine cables and the parallel trends in the event study do not underpin this concern. A threat could be that certain ethnic groups are also favored in any other dimension. Your strategy to overcome this threat is two-fold.

First, we can show descriptively that many countries construct APs for more than one ethnic group before the treatment period (Figure A.8). 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 APs. And Angola only built very few APs 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 AP, while towns with other ethnic groups remain on a worse growth path. The results are shown in Column (7) of Table B.3. The sample size does not shrink a lot, it shows that for most ethnic groups for which APs were constructed in the treatment period, APs 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.

7.5 External validity

7.6 Extending the countries

Applying the classical two-way fixed effect model of Column (2) in Table B.3, we can re-estimate Equation (1) on a broader sample. Comparing treatment and control group towns in the whole sample through the year fixed effect, we can allow also for countries containing either only control or treated towns. Table 5 repeats the step-wise selection of non-nodal cities from Table 1. With this approach, we estimate on 491 (including nodal cities) to 352 towns in 17 to 19 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 be seen easily that this sample contains more control towns as countries which started later with the rollout are added primarily. In fact, the number of treated towns increases only by three from 98 to 101 in Column (5).

Comparing the results with the ones from the stepwise procedure in Table 1, the estimate of the main effect again increases slightly from Column (1) to (4) and also when adding mobile coverage as control variable in Column (5). Moreover, the estimates are slightly bigger, reaching an effect size of nearly 10 percent. In contrast, the estimates are statistically significant throughout all columns at the 5% level and at the 1%

level in our preferred specification in Column (5). Here, the estimate of mobile coverage is statistically significant at the 10% level and positive with a medium sized point estimate. In comparison to Column (2) in Table B.3, where we applied the same fixed effects on the more restrictive sample, the estimate of the main effect increases by nearly .04 and is statistically significant at a higher level. The significance level of the mobile coverage control is lower. The estimated result is very comparable to the more restrictive sample. Nevertheless, it indicates that the effect of broadband internet is very comparable across SSA countries (irrespective of how they rollout the national backbone infrastructure).

Table 5: Robustness (external validity)

	(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
City 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 km buffer as in Table 1. GSM mobile-phone coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants and all specifications include town and year fixed effects. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The event study graph shows parallel-trends before the treatment (Figure A.5). Especially in the four years prior to the treatment, the point estimates are close to zero. Only the estimate in year -5 is statistically significant at the 5% level and is strongly negative. This is not the case in the main specification in Figure 6, where all pre-treatment coefficients are statistically insignificant. After the treatment, both event-study estimates are very comparable. In both cases, the point estimate increases especially two and even more three years after the treatment and becomes statistically significant at the 5% level only in the last observation period. As in Table 5, the point estimates slightly exceed the ones in the main specification in post-treatment period.

One concern is 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; Callaway and Sant'Anna, 2020; Sun and Abraham, 2020). Figure A.7 shows event-study estimates for the respective estimators. 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.5, only the point estimate of year 2 is statistically significant and the point estimate in year 3 also declines slightly in comparison to the year before. However, the point estimate in year 2 is bigger than in Figure A.5. Another differences to Figure A.5 is that in all three estimations only one pre-treatment estimate is statistically significant at the 5% level. Again, the point estimate in year -5 (in the approach by Sun and Abraham (2020)) is negative and statistically significant at the 5% level.

7.7 Extending the post-treatment period

Originally, we limit the sample when the second generation of SMCs arrive. So, we can estimate the effect of broadband internet at basic speeds. The event study estimates, Figure 6, show that the effect on local economic growth increased on a yearly basis. Therefore, we show for robustness how this effect evolves in two more subsequent years. One should note, however, that the effects in the last periods might be driven by fast internet induced by the new SMCs. On the other hand, not all countries already have fast internet available by the end of the new sample. Indeed, the average effect increases to .12. The event study estimates are shown in Appendix Figure A.6 and indicate the the growth rate increases further to more than 20 percent.

8 Discussion

Previous estimates of economic growth induced by broadband serve for comparison with our results. We find that cities with internet available growth 2 percentage points faster in their GDP than not connected cities. Though Czernich et al. (2011) also investigate GDP growth in OECD countries induced by broadband internet and timing and internet speed is very comparable, some factors make it difficult to compare the two results. Czernich et al. (2011) find that the broadband increased GDP per capita by 2.7—3.9 percent, implying a .9–1.5 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). 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). This is only very few years earlier than the landing of the SAT3-cable in 2001. Two major differences are that we cannot investigate broadband penetration and compare cities within countries and not broadband penetration 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. Though, in SSA broadband internet fails high penetration, it still counts as a general purpose technology as firms can access information such as global prices a lot easier with access to the Internet.

For SSA, Hjort and Poulsen (2019) estimates a 3.3 percent increase in economic activity of 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. Hence, though in both cases local economic activity is measured, the comparison is different. Finally, the selection of cities and towns is different 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 other issues. Finally, it cannot be rejected that the effects of the extensive margin—of pure connectivity—are still in play when the next generation of SMCs landed. Nevertheless, both studies show that SMCs that brought internet to SSA a different speeds had both a positive effect on the local economic growth there.

Finally, we want to compare our results to Storeygard (2016) who also estimates city growth. Though, not estimating the effects of a digital infrastructure, he is most closely related to our work with regard to the outcome measure. Therefore, we can state that a town experiencing an oil price shock of $70 and being 200 km away from the primate city can compensate this loss when having internet available in comparison to a not connected but otherwise identical city.

Both studies include only coastal countries.

9 Policy implication

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 of this infrastructure payoff the costs of the rollout. Our study comes in at this point. We showed that even smaller towns that were connected incidentally are growing faster than comparable towns without access points to make internet available. 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. The effects of internet are not bound to a high uptake. Hence, we recommend to rollout this infrastructure further even when only a low uptake is expected. We believe that even a small uptake by some firms generates 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.

10 Conclusion

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

We find that the connection of towns to the world wide web, on average, leads to an increase in light intensity of about 7 percent, relative to similar towns not (yet) connected. This translates into 2 percentage points higher growth in terms of GDP. Moreover, we differentiate growth in more pixels, where towns increase in their area (extensive margin), and in a higher average of the light intensity, which is associated with a higher density of economic activity (intensive margin). We find that towns with internet availability due to access to digital infrastructure typically grow on both margins, i.e. become brighter and geographically bigger. Further, our results suggest that this growth is partly driven by growing populations in connected towns.

References

Aggarwal, S. (2018, jul). Do rural roads create pathways out of poverty? Evidence from India. *Journal of Development Economics 133*, 375–395.

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

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

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

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

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

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

Bertschek, I., D. Cerquera, and G. J. Klein (2013). More bits - more bucks? Measuring the impact of broadband internet on firm performance. *Information Economics and Policy 25*(3), 190–203.

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

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

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

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

Callaway, B. and P. H. Sant'Anna (2020). Difference-in-Differences with multiple time periods. *Journal of Econometrics* (xxxx), 1–31.

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

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

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

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

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

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

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

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

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

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

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

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.

Duflo, E. and R. Pande (2007). Dams. *The Quarterly Journal of Economics 122*(No.2 May), 601–646.

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

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

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

Goodman-bacon, A. (2019). Andrew Goodman-Bacon July 2019. (July).

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

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

Hamilton Research (2020). Africa bandwidth maps.

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

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

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

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

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

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

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

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

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

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

Lipscomb, M., A. Mushfiq Mobarak, and T. Barham (2013). Development effects of electrification: Evidence from the topographic placement of hydropower plants in Brazil. *American Economic Journal: Applied Economics 5*(2), 200–231.

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

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

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

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

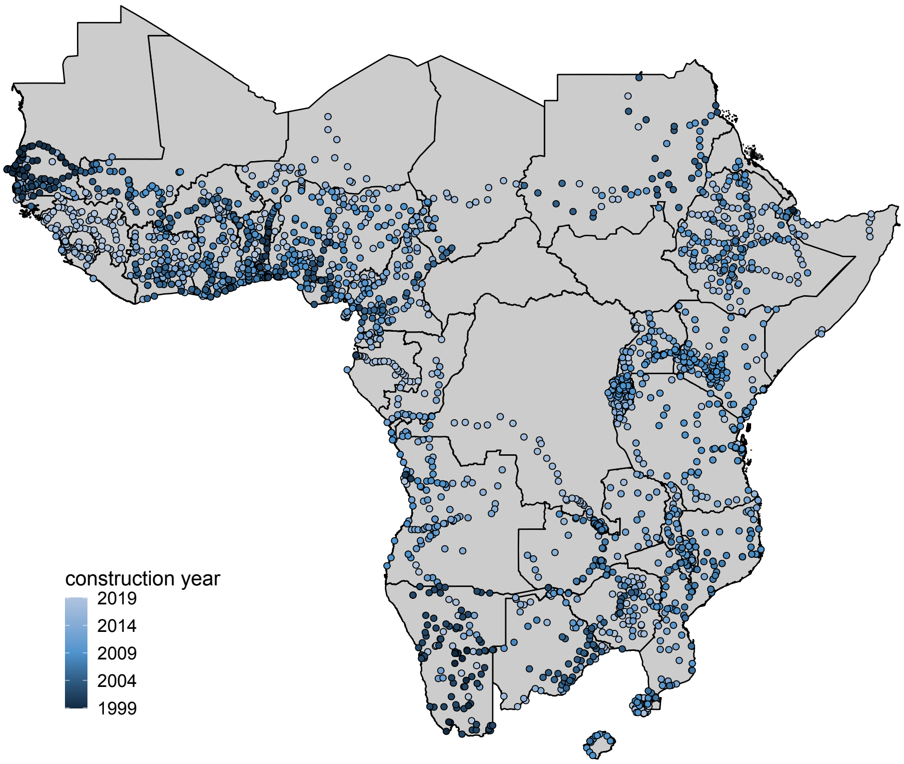
Storeygard, A. (2016). Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa. *Review of Economic Studies 83*(3), 1263–1295.

Sun, L. and S. Abraham (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.

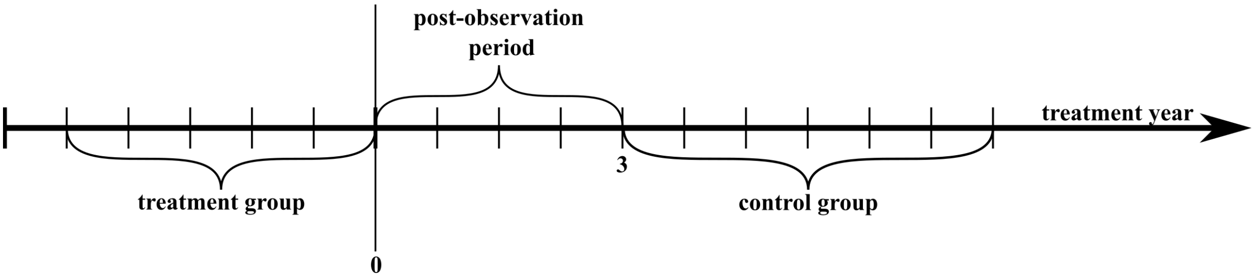
A Appendix: Figures

Figure A.1: Access points and their construction years



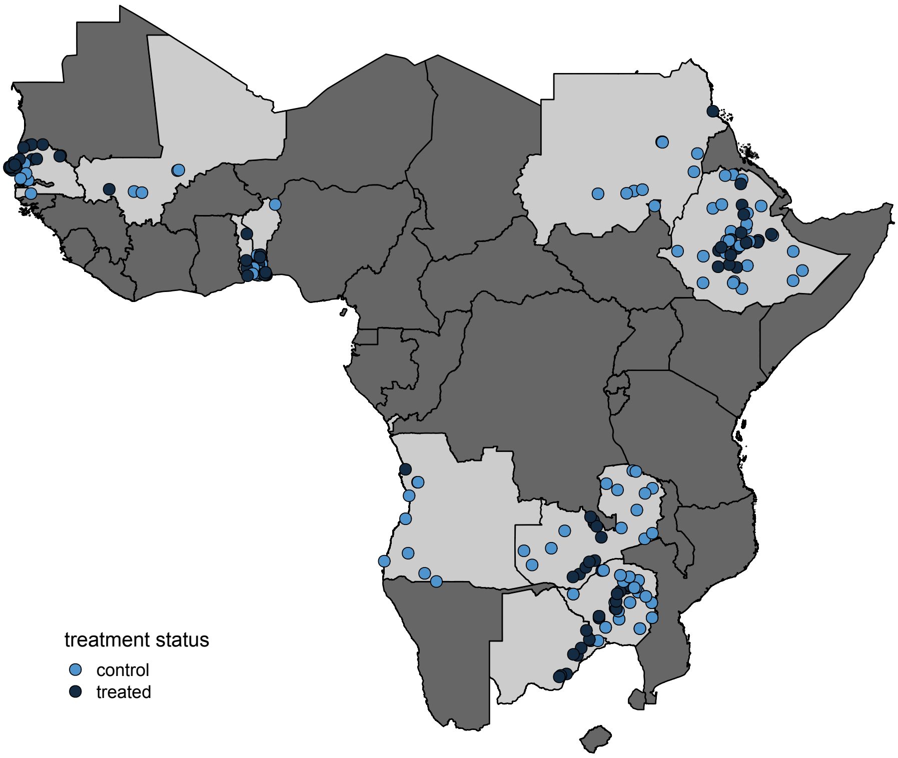
Notes: The figure depicts the location and construction date of all SSA access points. Brighter blue dots correspond to later constructed APs.

Figure A.2: Timeline identification



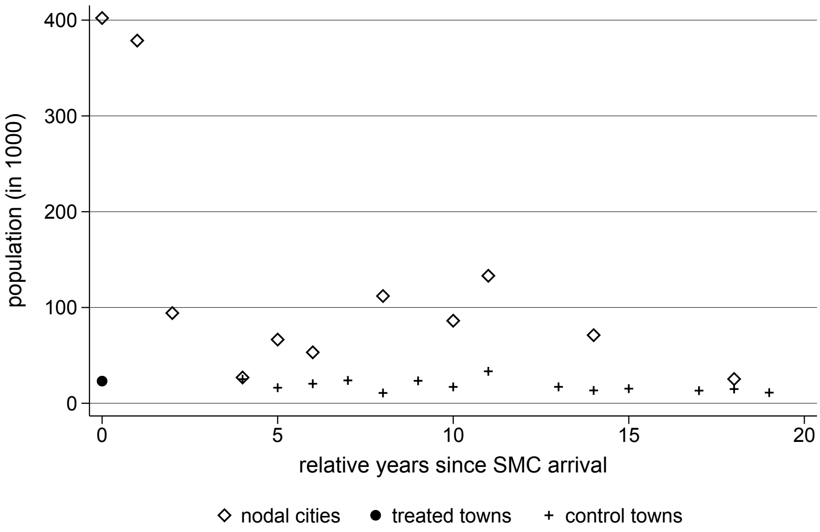
Notes: The figure gives an overview, of the timeline of our identification strategy.

Figure A.3: Countries and towns location in the estimation sample



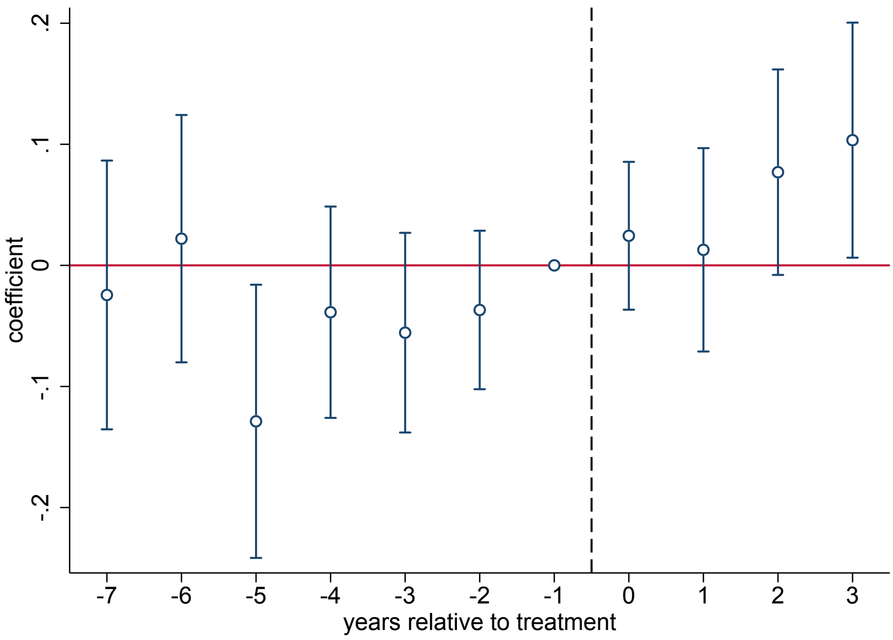
Notes: This figure depicts the countries in our analysis (brighter gray) and for each country the towns in the treatment and control group.

Figure A.4: Population size of connected cities and towns by year (relative to connection year)



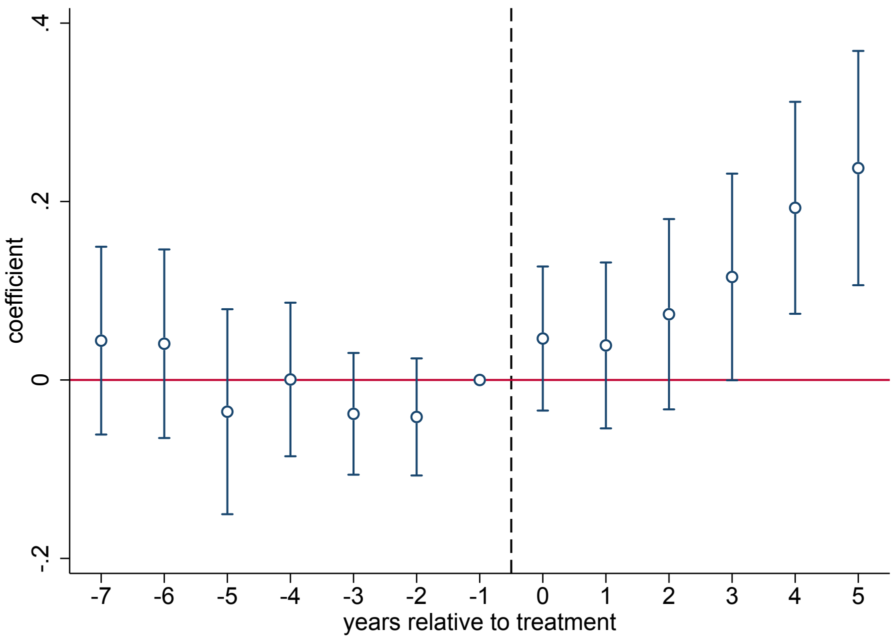
Notes: The figure depicts the average population size of connected cities and towns by year relative to the connection year. The black dot in the lower left corner represents the treated towns, while the control towns are represented by the plus symbol and the nodal cities by a diamond. For treated towns and nodal cities that were connected in earlier years than the arrival of an SMC are shown in year zero as well for clarity.

Figure A.5: Robustness event study (external validity)



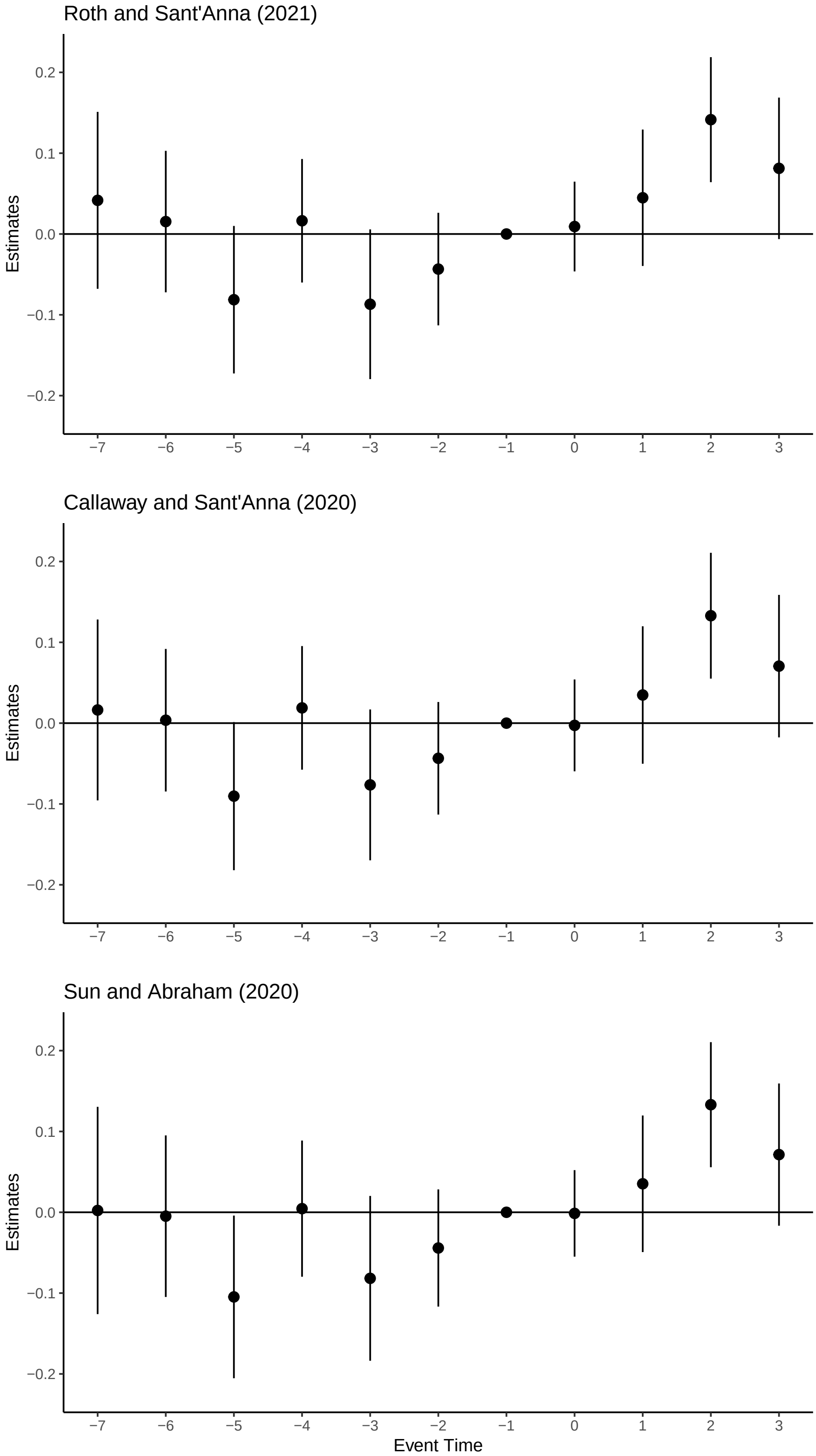
Notes: For external validity, more relaxed fixed effects are applied and therefore more countries are included.

Figure A.6: Event-study coefficients with longer post-treatment period



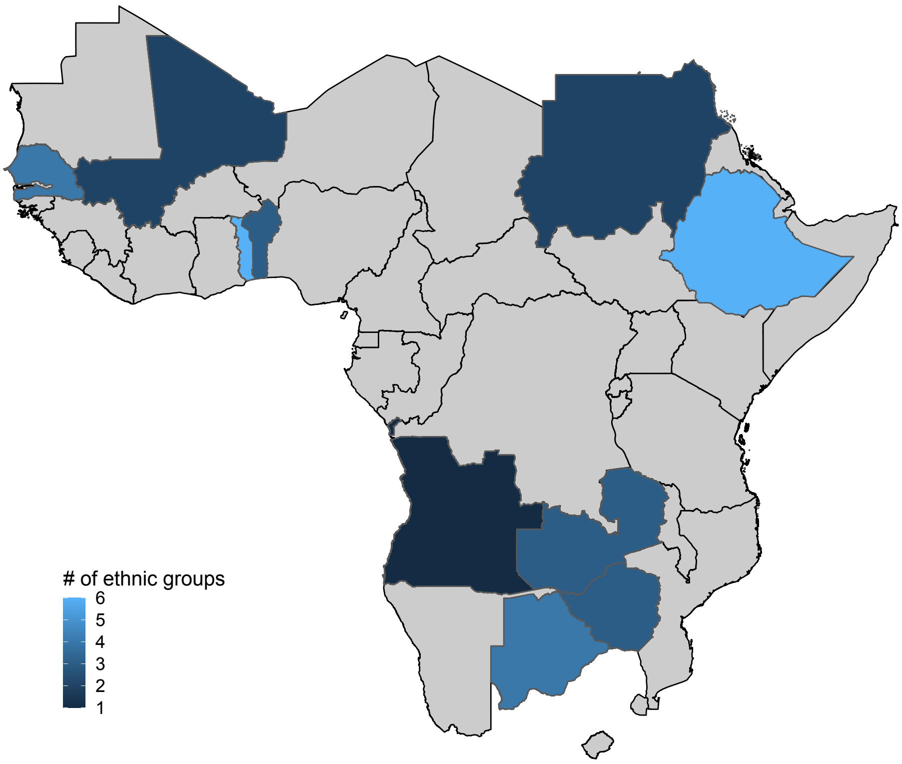
Notes: Coefficients for event study specification of Column (2) from Table B.3. Robust standard errors clustered by town for 95% confidence interval reported as bars.

Figure A.7: Robustness event study (external validity)



Notes: Staggered adoption, applying XX

Figure A.8: Ethnic groups



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

B Appendix: 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

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
post x treated	0.0703**	0.0580*	0.0703*	0.102**	0.0625*	0.0797***	0.0720*
	(0.0349)	(0.0338)	(0.0358)	(0.0449)	(0.0336)	(0.0265)	(0.0374)
GSM coverage	0.0486	0.0908***	0.0486	0.0193	0.0233	0.0193	0.0482
	(0.0342)	(0.0320)	(0.0330)	(0.0332)	(0.0310)	(0.0249)	(0.0408)
Observations	2,420	2,420	2,420	2,420	2,827	2,343	1,804
R-squared	0.943	0.927	0.943	0.962	0.951	0.978	0.939
#countries	10	10	10	10	10	10	9
#cities	220	220	220	220	257	213	164
share treated	.445	.445	.445	.445	.525	.502	.445
City FE	✓	✓	✓	✓	✓	✓	✓
Country x Year FE	✓		✓	✓	✓	✓	
w/o capital+landingpoint	✓	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓	✓
w/o population >100k	✓	✓	✓	✓	✓	✓	✓
Year FE		✓					
#ethnic group-countries							13
Ethnic Group-Country x Year FE							✓
no buffer						✓	
late APs					✓		
linear time trends				town-level			
Cluster			state-level				

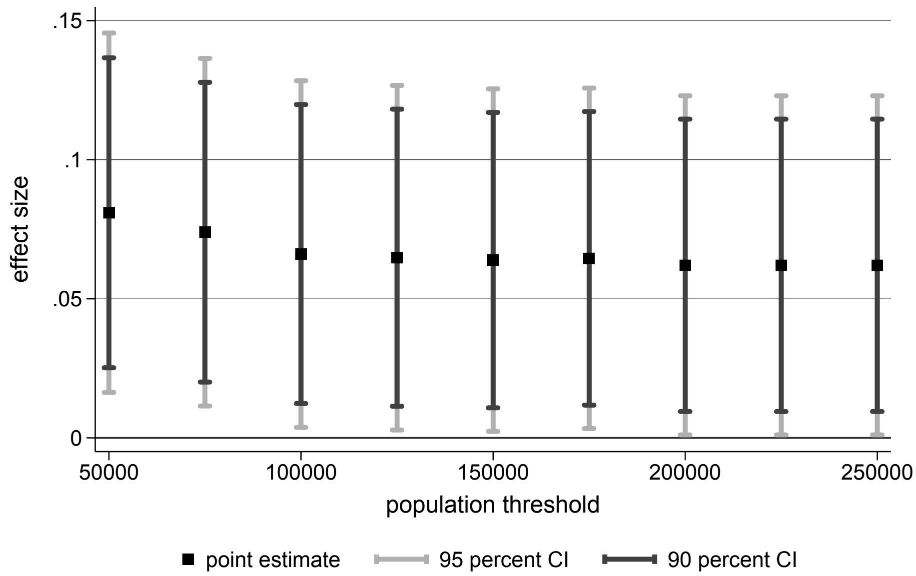
Notes: Light intensity is measured as the logarithmic sum of light intensities of DMPS-OLS pixels within town area. GSM mobile-phone coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing point cities and capitals, regional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

C Appendix: Further robustness checks

Spatial correlation Following Conley (1999) we calculate the standard error to account for spatial correlation. Results remain. Moreover, Moran's I indicates that the data is not spatially correlated.

Definition of nodal cities The main results show that the estimate does not change when restricted to towns with less than 100,000 inhabitants. However, this threshold is chosen arbitrarily. Therefore, we vary the population threshold as a further robustness check. Figure C.9 shows that the estimate remains independently of the chosen population threshold.

Figure C.9: Robustness nodal cities



Notes: Variation of population thresholds are shown. Coefficients for the specification of Column (5) from Table 1. Robust standard errors clustered by town.

Definition of control group One concern might be that very late connected towns might not be comparable to the treated towns. However, Table C.4 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 increases when only including towns that were connected shortly after the arrival of the SMC. This indicates that the control towns might be a heterogeneous group. Nonetheless, the restriction decreases the sample size. The last column repeats the main effect estimate.

Linear outcome measure Our results also hold when estimating the absolute light intensity instead of the logarithm (Table C.5). Especially, nodal cities, in the first columns of Table C.5, growth a lot stronger in

Table C.4: Robustness: Connected control towns

VARIABLES	(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
City FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
w/o capital+landingpoint	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
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-phone coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

absolute terms.

Table C.5: Robustness: Light intensity (absolute)

VARIABLES	(1)	(2)	(3)	(4)	(5)
post x treated	318.6***	79.87***	56.50***	42.99**	44.74***
	(111.4)	(23.00)	(19.72)	(17.26)	(17.19)
GSM coverage					12.25
					(14.63)
Observations	3,190	3,069	2,563	2,420	2,420
R-squared	0.987	0.979	0.986	0.963	0.963
#countries	10	10	10	10	10
#cities	290	279	233	220	220
share treated	.493	.473	.468	.445	.445
City FE	✓	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓	✓
w/o capital+landingpoint		✓	✓	✓	✓
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 km buffer. GSM mobile-phone coverage is calculated as the percentage share of town area covered with signal. Sample successively restricted by excluding landing point cities and capitals, regional capitals, and cities with more than 100,000 inhabitants. Robust standard errors clustered by town reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

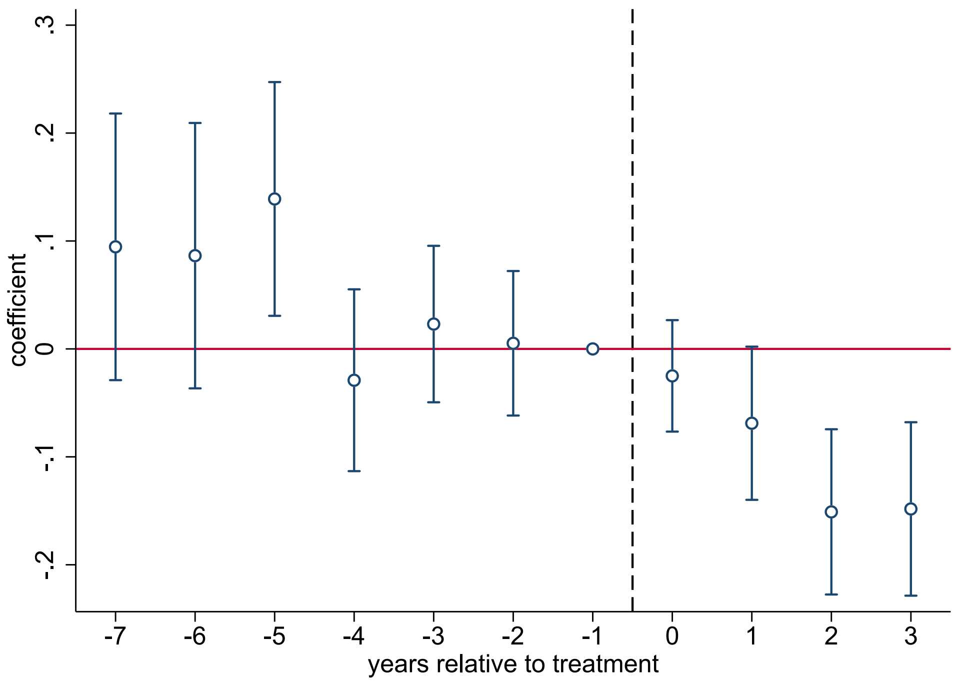
D Appendix: Mobile coverage

Table D.6: Mobile coverage

VARIABLES	(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
#cities	290	279	233	220	220	220	220
share treated	.493	.473	.468	.445	.445	.445	.445
City FE	✓	✓	✓	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓	✓	✓	✓
w/o capital+landingpoint		✓	✓	✓	✓	✓	✓
w/o regional capitals			✓	✓	✓	✓	✓
w/o population ¿100k				✓	✓	✓	✓

Notes: Mobile coverage as share of built-up area with the most basic technology (GSM/2G). 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 D.10: Event-study coefficients for mobile coverage



Notes: Coefficients for event study specification of Column (4) from Table D.6. Robust standard errors clustered by town for 95% confidence interval reported as bars.

E Appendix: Industry changes and firm creation

E.1 Industry changes

As discussed in the main section, regional industry shares change when Internet becomes available. Next, we discuss the level changes and their percent changes. Therefore, we add the total number of survey individuals per region in Columns (1) and (5). This shows whether in treated regions the number of surveyed individuals increases. Table E.7 shows that the total number of survey individuals does not increase in the whole sample nor in the reduced one. Though, the point estimate is positive in both cases, it is far from statistical significance. This strengthens the interpretation that indeed individuals being employed in agriculture, find a job in manufacturing and services. While in industry shares the substitution is equal between manufacturing and services, in absolute terms employment in services increases nearly three times more. Turning to the percentage changes, Table E.7 at least indicates an increase of 13 percent of survey individuals in the reduced sample. Here, the increase in manufacturing is predominant. It is highly statistically significant in both samples and the point estimate is bigger than the one for services. A decrease in agriculture is not statistically significant in the reduced sample. In conclusion, these figures could again be interpreted that the treated towns do not find a stronger population increase than the towns in the control group. However, their industry structure changes and might be the channel through which economic development happens. However, this analysis only shows the workers side and does not talk about firm creation.

Table E.7: Employment by industry (count)

VARIABLES	(1) total	(2) agriculture	(3) manufacturing	(4) service	(5) total	(6) agriculture	(7) manufacturing	(8) service
post x treated	150.6	-536.3**	159.4**	527.4**	287.8	-674.1**	226.9***	735.1***
	(345.0)	(259.8)	(66.52)	(208.4)	(449.2)	(329.3)	(83.25)	(268.9)
Observations	182	182	182	182	124	124	124	124
R-squared	0.965	0.971	0.844	0.883	0.957	0.968	0.853	0.856
#countries	5	5	5	5	3	3	3	3
#cities	91	91	91	91	62	62	62	62
share treated	.363	.363	.363	.363	.419	.419	.419	.419
City FE	✓	✓	✓	✓	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓	✓	✓	✓	✓
w/o capital+landingpoint	✓	✓	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓	✓	✓
w/o population ¿100k	✓	✓	✓	✓	✓	✓	✓	✓
late adopters	included	included	included	included	included	included	excluded	excluded

Notes: Regional industry composition comes from IPMUS International survey. 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 E.8: Employment by industry (count ln)

VARIABLES	(1) total	(2) agriculture	(3) manufacturing	(4) service	(5) total	(6) agriculture	(7) manufacturing	(8) service
post x treated	0.0861	-0.162*	0.276***	0.107	0.139*	-0.170	0.394***	0.199**
	(0.0625)	(0.0831)	(0.102)	(0.0799)	(0.0797)	(0.107)	(0.108)	(0.0933)
Observations	182	182	182	182	124	124	124	124
R-squared	0.968	0.970	0.946	0.956	0.958	0.962	0.960	0.958
#countries	5	5	5	5	3	3	3	3
#cities	91	91	91	91	62	62	62	62
share treated	.363	.363	.363	.363	.419	.419	.419	.419
City FE	✓	✓	✓	✓	✓	✓	✓	✓
Country x Year FE	✓	✓	✓	✓	✓	✓	✓	✓
w/o capital+landingpoint	✓	✓	✓	✓	✓	✓	✓	✓
w/o regional capitals	✓	✓	✓	✓	✓	✓	✓	✓
w/o population ¿100k	✓	✓	✓	✓	✓	✓	✓	✓
late adopters	included	included	included	included	included	included	excluded	excluded

Notes: Regional industry composition comes from IPMUS International survey. 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.