WP/21/110

IMF Working Paper

The Macroeconomic Impacts of Digitalization in Sub-Saharan Africa: Evidence from Submarine Cables

By Félix F. Simione and Yiruo Li

INTERNATIONAL MONETARY FUND

 WP/21/110

IMF Working Paper

African Department

The Macroeconomic Impacts of Digitalization in Sub-Saharan Africa: Evidence from Submarine Cables

Prepared by Félix F. Simione and Yiruo Li[1]

Authorized for distribution by Karen Ongley

April 2021

Abstract

In recent decades, Sub-Saharan Africa (SSA) has experienced rapid advances in digitalization, with unprecedented growth in internet penetration. This paper investigates the impact of internet penetration in SSA on real GDP growth, productivity, and sector output and employment. The analysis treats the arrival of submarine cables to the region as a natural experiment to estimate the impacts of internet penetration. The findings suggest that higher internet penetration increases real per capita GDP growth and productivity, and impacts the composition of both output and employment, with an increasing share from the services sector.

JEL Classification Numbers: E23, E24, J62, O11, O33, O47, O55.
Keywords: Sub-Saharan Africa, economic growth, productivity, digital divide, technology adoption, labor mobility.
Author's Email Address: FSimione@imf.org; yli29@gwu.edu

[1] Authors would like to thank the IMF team of the Spring 2020 Regional Economic Outlook for Sub-Saharan Africa (REO SSA) project, the Inclusive Growth Network seminar participants, and the Research Advisory Group in the IMF African Department for their helpful comments. This Working Paper delves more deeply into part of the empirical analysis underpinning the published REO SSA 2020 chapter on digitalization in SSA. Yiruo Li conducted this work while still working in the African Department of the IMF.

Contents

I. INTRODUCTION

Digitalization has become an important theme in policy debate circles, and more so amid the COVID-19 pandemic. One of the several dimensions of digitalization is the use of the internet as a tool to collect, store, process, and exchange information. Higher internet penetration can be viewed as a positive shock to the productivity of both labor and capital. As such, in modern economies, the internet shapes production and service delivery. The internet is also critical for business continuity. As witnessed during the COVID-19 pandemic, in countries with reliable and widespread internet access, the ability to telework and conduct remote transactions allowed authorities to enforce lockdowns more easily in response to the outbreak.

Despite its myriads benefits, internet access remains a hurdle in many developing countries, particularly in SSA where access rates in 2017 ranged from as low as 1.3 percent of the population to 62 percent (compared to an average of 68 percent in advanced economies). Fortunately, the arrival of submarine cables is changing the landscape of internet penetration in the region. These cables, most of which landed in coastal SSA countries in 2009-2012, dramatically reduced the access cost and increased the download speed, ultimately fostering an escalation in internet penetration.

But has the increase in internet penetration *de facto* boosted economic growth in SSA? Past empirical studies, most of them reviewed in Minges (2015), attempted to address this question by estimating cross-country regressions to uncover the relationship between broadband internet penetration and economic growth. While some studies use different approaches to address the endogeneity in the relationship between digitalization and economic growth (e.g., Czernich et al. 2011; Katz and Callorda 2018), others do not (e.g., Qiang et al. 2009; Zaballos and López-Rivas 2012). Addressing this empirical challenge often requires the identification of an exogenous variation in internet penetration.

This paper argues that the arrival of submarine cables in SSA can be explored to construct instrumental variables for internet penetration. Being connected to such cables provides an exogeneous source of variation in internet penetration due to its differentiated impact in coastal (treatment group) and landlocked (control group) countries[2], as well as due to a weak relationship between the data traffic capacity of the cables and the level of income of the countries they connect. The paper then uses this exogeneous variation to estimate the impact of internet penetration on economic growth, productivity, and sector value added and employment shares.

The findings show a large and significant effect of internet penetration on real per capita GDP growth and productivity at the aggregate and sector levels. From a sectoral perspective, it finds that higher internet penetration is positively associated with the share of services in the economy and negatively associated with the share of industry. The effect is insignificant for the share of agriculture. In terms of employment, internet penetration does not have a significant impact on industry employment (as percent of total employment), but it does have a positive association with services employment, while being negatively associated with agriculture employment. The paper further finds that the positive association with services employment is higher for females than for males. These findings support the renewed policy

[2] Section III will elaborate on the features of the control and treatment groups.

attention to digitalization reforms in SSA, including to expand access to high-quality and reliable internet connectivity as a tool to boost growth, raise productivity, and foster diversification of production and employment. The latter is particularly relevant in SSA where economies are highly concentrated in the production of primary commodities subject to cyclical volatility.

This paper relates to the studies by Hjort and Poulsen (2019) and Cariolle (2018). These studies explore the arrival of submarine cables as a natural experiment impacting internet outcomes in SSA at the firm and country levels, respectively. Compared to Hjort and Poulsen (2019), this paper explores a modified submarine cables experiment[3] covering a larger sample of countries and more years. Whereas Hjort and Poulsen assess the impacts of being connected to fast internet on employment and productivity at the firm level, this paper focuses on a country-level sample and, given this feature, it estimates the macro-level impacts of internet penetration on economic growth, productivity and sectoral output and employment shares. Also, whereas Cariolle (2018) examines the impact of submarine cables on internet penetration, this paper goes an extra step. It uses the submarine cables-based estimates of internet penetration under a modified experiment to assess the impact of internet penetration on the macroeconomic variables mentioned earlier. To our knowledge, this paper is the first to explore the submarine cables experiment to assess country-level macroeconomic impacts of internet penetration in SSA.

While the paper's empirical strategy is chosen to attempt to address endogeneity, the results are sensitive to the specification of the instrumental variable and the underlying samples. Also, the natural experiment that the paper explores is not randomized in a strict sense as it is based on observational data. This means that claiming causality is still subject to fully controlling for relevant cross-country differences, which is challenging in empirical applications with observational data, particularly in SSA.

The remainder of the paper is as follows. Section II reviews the relevant literature. Section III provides a brief review of the arrival of submarine cables in SSA as a background for the empirical strategy used in the paper. Section IV describes the empirical strategy and the data. Section V and VI report the results and discuss their robustness, respectively. Section VII concludes and discusses policy implications.

[3] The term "experiment" here does not mean a randomized controlled trial or experiment. What we term "experiment" will be further explained throughout the paper, and particularly in section IV.

II. LITERATURE REVIEW

Digitalization refers to the widespread use of digital technologies, such as the internet and mobile phones, to collect, store, analyze, and exchange information (World Bank 2016; Brookings 2017). Evidence has shown that digitalization has profoundly influenced economic growth, productivity, labor market outcomes, and the sectoral distribution of output. Within the growth accounting framework, digitalization can affect economic growth through improved productivity by complementing and nurturing human labor and capital (Barro 1991; Oulton 2010; Akerman et al. 2015; Tisdell 2017). While the growth accounting framework is useful in empirical work, it relies on assumptions that do not always hold especially in developing countries (Stryszowski 2012). Empirical studies have therefore relied more on econometric methods through which a set of extensive country-specific factors correlated to both growth and digitalization can be controlled for more explicitly.

In most empirical applications, broadband internet penetration and related variables have been used as a proxy for digitalization. Country-level empirical studies have widely corroborated the positive effect of broadband internet penetration on economic growth. Using a sample of 120 countries and linear regressions, Qiang et al. (2009) find that a 10 percent increase in fixed broadband penetration is associated with a 1.2 percentage point increase in the average real per capita GDP growth during 1980-2006 for developed countries. This effect increases to 1.4 percentage points for developing countries. Czernich et al. (2011) focus on OECD countries and find that a 10-percentage point increase in broadband penetration increases per capita GDP growth by 0.9-1.5 percentage points. Zaballos and López-Rivas (2012) apply a non-linear model for 26 Latin American and Caribbean countries. Their results link a 10 percent increase in broadband penetration to a 3.19 percent increase in per capita GDP.

The fact that studies have yielded estimates of growth impacts at various magnitudes may imply a non-linear relationship between digital penetration and economic returns. Katz and Callorda (2018) note that digitalization returns on growth follow an inverted U-shape curve, suggesting the presence of a saturation point. They argue that mobile broadband technology has reached high saturation in most countries and these countries have entered the phase of diminishing returns. Fixed broadband, in contrast, is still in a phase of increasing returns. The literature also points out that, to reap digital dividends, a country needs to reach a sufficiently high level of digital penetration, before which the impact on growth can be relatively small (Roller and Weavers 2001; Koutroumpis 2009).

Digitalization has also been linked to structural transformation in the economy and uneven gender effects. As discussed in Mathess and Kunkel (2020), technology-driven changes affect relative prices and productivity across sectors which, in turn, triggers cross-sector shifts of labor and output. However, as the study notes, cross-sector structural change can also result from factors unrelated to digitalization such as changes in aggregate income, changes in input-output or sectoral linkages, international trade, supply for skilled workers, and transaction costs which partly relate to institutional soundness.

Houngbonon and Liang (2018) find that, following the wide adoption of broadband internet in France, the overall unemployment rate remained unchanged, while the services and manufacturing sectors became the net job creator and loser, respectively, implying a structural

change in employment. With regards to gender effects, Klonner and Nolen (2008) analyze the positive employment impact of mobile networks in South Africa and reveal that women mainly drive the gains in employment. Sovbetov (2018), based on a study in Turkey, finds that increases in e-commerce transactions are associated with significant female employment growth.

Another string of studies on digitalization focuses on productivity. At the country level, a recent assessment is compiled in Dieppe (2020) building on an updated global productivity database. The findings suggest that technology-driven improvements have economically meaningful and statistically significant effects on labor productivity growth. At the individual and firm levels, digitalization has been associated with uneven productivity gains favoring innovation (Franklin et al. 2009), highly productive firms (Paunov and Rollo 2016), and skilled workers in executing non-routine tasks (Akerman et al. 2015)

Digitalization also influences labor market outcomes through complementing or substituting human labor. The overall employment impact of digital technologies, therefore, can be decomposed into simultaneous job creation and destruction. Regarding job destruction, Akerman et al. (2015) find that, with the roll-out of broadband, unskilled workers were displaced and saw a reduction in wages. Dieppe (2020) argues that in the short run, the job substitution effect dominates for both advanced economies and emerging markets and developing economies. At the same time, job creation impacts are reported in both developed countries (Lehr et al. 2006; Crandall et al. 2007) and the developing world (Klonner and Nolen 2008; Hjort and Poulsen 2019). The job creation and destruction effects attributable to digitalization may ultimately drive uneven employment shifts across sectors. Many studies in advanced economies find a positive impact on employment in the services sector where tasks are less easily automated (Acemoglu 1999; Autor and Dorn 2009; Goos et al. 2014).

Overall, previous studies have broadly confirmed the positive impacts of digitalization on growth and productivity. The surveyed literature suggests that the magnitude of the impacts differs with respect to the level of digitalization, sector, skill level and gender.

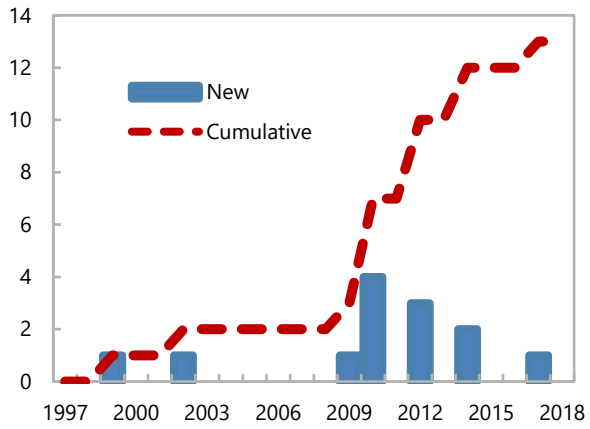
III. THE ARRIVAL OF SUBMARINE CABLES IN THE COAST OF AFRICA

In mid-January 2020, Bloomberg reported: "African Internet Slows After Undersea Cables Break".[4] The alluded cables were WACS and SAT3/WASC, two cable networks in the Atlantic Ocean connecting SSA countries to Europe. Their temporary failure slowed internet speed in several SSA countries. The event was a reminder of the critical role submarine cables play in internet connectivity. Submarine cables (SCs) are a network of fiber cables under the sea connecting different countries and continents. They move internet traffic with greater efficiency (faster speed, lower cost) compared to satellite networks that prevailed in the past. As shown in Figures 1 and 2, most SCs in SSA became operational during 2009-12, although

[4] For details, see Bloomberg's link (January 17, 2020): https://www.bloomberg.com/news/articles/2020-01-17/african-internet-slowdown-continues-after-undersea-cables-break.

some countries had been connected to SCs earlier (e.g. South Africa, Nigeria, Benin, and Cameroon).

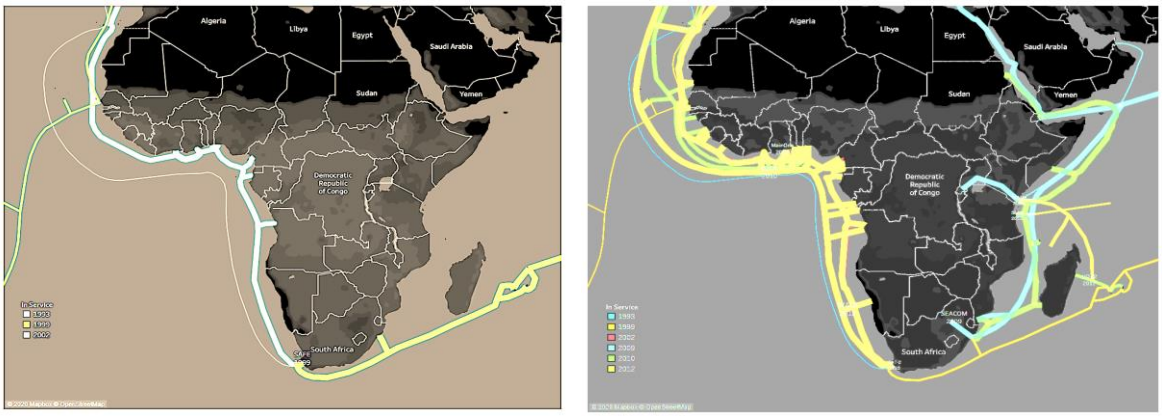
Figure 1. SSA: Number of Submarine Cables Connecting Two Countries or More



Sources: cablemap.info and authors calculations.

Figure 2. SSA: The Arrival of Submarine Cables

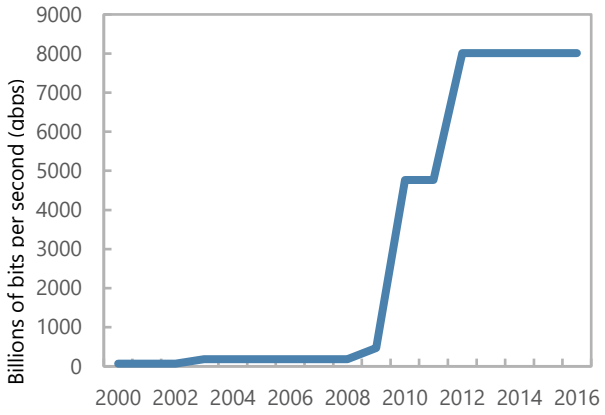
Before 2009 **2009 Onwards**



Sources: cablemap.info and authors' mapping based on *Mapbox* and *OpenStreetMap*.

While being connected to a SC offers efficiency advantages relative to satellite networks, the capacity of the cable also matters. Cables with higher data capacity move larger volumes of internet content per unit of time (typically measured in terms of bits per second). With technological advances in the global information and technology industry, the capacity of SCs has evolved over time (Figure 3). Cables that arrived in SSA after 2009 have a dramatically higher capacity. The average cable capacity increased from 165 billion bits per second in 2002-2008 to above 6000 billion bits per second in 2009-2018.

Figure 3. SSA: Submarine Cable Capacity in Costal Countries



Sources: cablemap.info and author's calculations.

SSA witnessed important gains in internet services following the SCs revolution during 2009-12. Prior to the arrival of SCs, access to the internet in SSA was on average 10 times more expensive than in other regions. This changed considerably with the arrival of SCs. For instance, the arrival of EASSy (East African Cable System)—a cables network currently linking eastern African countries to Europe and Asia—helped expand internet access for 20 coastal and landlocked countries and lower broadband costs by as much as 90 percent (International Finance Corporation 2019; World Bank 2018). The average internet speed also increased substantially (Figures 4 and 5).

Figure 4. SSA: Monthly Subscription Fee for Fixed-Broadband Internet (US$), Average

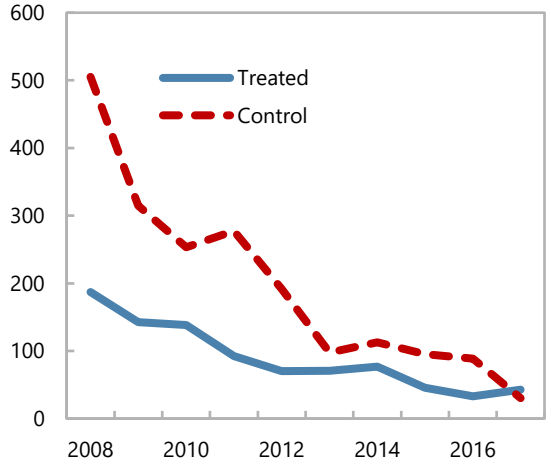
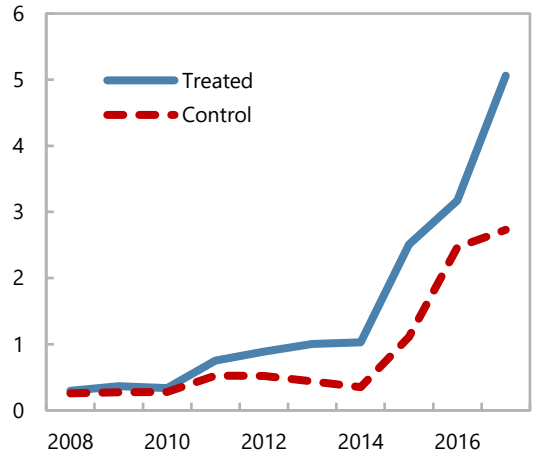


Figure 5. SSA: Fixed (Wired)-Broadband Speed (Mbit/s)



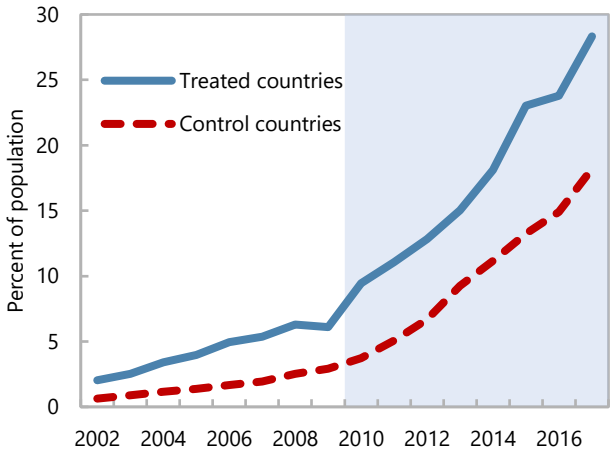
Sources: ITU and authors' calculations. "Treated" comprises countries that became directly connected to a SC starting in 2009. "Control" comprises those that were not directly connected. Section IV will further elaborate on this.

An aspect of SCs that is particularly relevant to this paper is that they are usually deployed regionally. Due to very high fixed costs involved in the underlying investments and operations, SCs often need to take advantage of economies of scale by connecting as many adjacent countries as possible (World Bank 2018; Cariolle 2018). It is therefore rare for a cable to skip a country along its route. For example, the cable ACE (Africa Coast to Europe submarine cable) has connected countries that are geographically continuous in the western coast of SSA. Regardless of differences in income level, each of those countries benefit from the same level of cable capacity. Another relevant fact for this paper is that most SC projects have sourced funding internationally. For instance, both EASSy and ACE received large funding from the World Bank.

A third relevant aspect for this paper is that SSA coastal countries, given their privileged location, connected directly to SCs and earlier than landlocked countries. Landlocked countries connected later and indirectly through the telecommunications network of coastal countries. A key reason for the delayed connection in landlocked countries is the limited network spillovers from neighboring coastal countries. Each SSA country is typically covered by a single backbone network, so the transmission of internet content between two countries' networks is subject to existing collaboration agreements, in the absence of which significant transit fees apply (Hjort and Poulsen 2019). Despite differences in the timing of connection, by the end of 2018 (i) almost all SSA countries were connected to at least one SC either directly (coastal countries) or indirectly through coastal countries (landlocked countries); (ii) most coastal countries were connected to multiple cables and (iii) there were 14 active cables and many more expected to come onstream.[5]

Given the above, the upward trend over time of internet penetration in SSA (measured as the percent of the population using the internet) is not surprising. On average, the internet penetration rate increased from 8.6 percent in 2002 to about 25 percent in 2017, with most of this increase occurring after 2009 when several SCs arrived (Figure 6). Also, not surprisingly, countries directly connected to a SC (mostly coastal) experienced a faster increase in internet penetration after 2009 than (landlocked) countries indirectly connected.

Figure 6. SSA: Percent of the Population Using the Internet



Sources: ITU and authors' calculations.

[5] For details, see cablemap.info: https://cablemap.info/_default.aspx. We only use cables that were active in SSA at the time of research.

The positive impact of SCs arrival on the internet penetration and speed in SSA has also been confirmed empirically in Cariolle (2018) and Hjort and Poulsen (2019). As this paper will further argue, the increase in internet penetration caused by the arrival of SCs after 2009, and the increase in the capacity of SCs over time, can be viewed as a quasi-experiment that is independent to macroeconomic conditions. In the next section, the paper exploits this quasi-experiment by laying out an empirical strategy to estimate the effect of internet penetration on a number of macroeconomic variables.

IV. EMPIRICAL STRATEGY AND DATA

The paper adopts as a proxy for digitalization the level of internet penetration defined as the percent of the population with access to the internet in a country. Internet penetration captures one of the several dimensions of digitalization, and the related time series is the most consistently available for SSA countries.[6] The paper estimates the effect of internet penetration on real per capita GDP growth, productivity, sector value added, and employment shares. To address the endogeneity between internet penetration and the macroeconomic variables of interest, the paper explores a natural experiment - the arrival of SCs in SSA - which affected the pace of internet penetration across countries. The paper defines as "treated" those countries that became directly connected to a SC starting in 2009. The "non-treated" (control) countries are by exclusion those that were indirectly connected to a SC. Treated countries are typically coastal countries, while control countries are landlocked.[7] The paper constructs two instrumental variables that capture (i) the unprecedented arrival of SCs starting in 2009 and (ii) the data capacity of the SCs measured in bits per second, a metric of internet data volume per unit of time[8].

Identification strategy

The following identification strategies underlie the instrumental variables. First, absent the unprecedented arrival of SCs in coastal SSA countries during 2009-12, the difference in internet penetration rates between coastal and landlocked countries would have remained broadly unchanged.[9] Second, the timing of the arrival of SCs and their capacity depend little

[6] As noted in the Literature Review section, several studies use internet penetration or related variables as a proxy for digitalization. Recent examples include Ouedraogo and Sy (2020).

[7] The exceptions are Rwanda and Uganda, two landlocked countries that are directly connected to a SC. Rwanda is connected to the same SC that Kenya (a higher income country) is connected to. Talks between the two countries started in 2009 over a partnership to connect Rwanda through Kenya, and the connection project was completed in 2010. Uganda became connected in late 2009 with the SC SEACOM in the coast of Tanzania.

[8] Hereafter the terms "data capacity", "cable capacity" and "SCs capacity" will be used interchangeably.

[9] In this identification strategy, which makes use of the simple difference-in-differences framework as in Duflo (2001), the level of internet penetration does not need to be, and in fact it is not, strictly the same between the two groups of countries in the starting year of the pre-treatment period. Factors other than submarine cables also

on macroeconomic conditions in countries connected by the cables. By corollary, the arrival of the cables and their capacity impact macroeconomic conditions only through their effect on internet penetration.

A few confounding variables could challenge the identification strategies. Coastal countries (directly connected to SCs) are more open to trade than landlocked countries, and trade openness is arguably correlated with some macroeconomic variables such as national income. Similarly, the unprecedented arrival of SCs starting in 2009 coincides broadly with the 2008-09 global financial crisis and the 2014-15 commodity price shock, both affecting macroeconomic conditions differently across countries. The regression analysis controls for those and other potential confounders.

Model specification

The following two-stage least square (2SLS) equations are estimated:

$$Internet_{i,t} = \alpha_0 + \alpha_1 Treatment_i + \alpha_2 Post2009_t + \alpha_3 (Treatment * Post2009)_{i,t} + \alpha_4 X_{i,t} + \varepsilon_{i,t} \quad (1)$$

$$Internet_{i,t} = \gamma_0 + \gamma_1 Cable_{i,t} + \emptyset_t + \vartheta_i + \gamma_2 X_{i,t} + \mu_{i,t} \quad (1')$$

$$Macro_{i,t} = \rho_0 + \rho_1 \widehat{Internet}_{i,t} + \rho_2 X_{i,t} + \omega_{i,t} \quad (2)$$

Equation (1) is the first-stage equation of the 2SLS, and is given by a fuzzy difference-in-differences (DiD) specification (as in Duflo 2001) where $Internet_{i,t}$ is internet penetration in country i at time t, $Treatment_i$ is a dummy variable that takes the value 1 if country i is directly connected to a SC and zero otherwise. $Post2009_t$ is a dummy variable that takes the value 1 if the time period is 2009 onwards and zero otherwise, $(Treatment * Post2009)_{i,t}$ is the interaction term between the two variables (i.e., being directly connected to a SC in 2009 onwards), $X_{i,t}$ is a vector of control variables, including lagged per capita income, public investment, trade openness, terms of trade, population density, institutional quality, human capital, credit to private sector, and fiscal position.[10]

All equations are estimated with robust standard errors (clustered at the country level). $\varepsilon_{i,t}$ is the error term, and α_0, α_1, α_2, α_3 and α_4 are parameters. Equation (1) by design removes any country- and time-fixed effects as it features two groups of countries that are observed in two

affect the difference between the two groups. Provided that such factors are controlled for, which is done in the regression framework discussed below, the difference-in-differences analysis can be employed.

[10] The list of control variables changes slightly depending on whether the outcome variable is growth and productivity, or structural change as captured by sector shares in value added and employment. In the latter case, some controls (fiscal balance, private sector credit and population density) are not included as the reviewed structural change literature (e.g. Matthess and Kunkel 2020) does not identify them as relevant for structural change.

time periods under a DiD specification. The parameter α_3 is of particular interest as it gives the average DiD effect on internet penetration of being directly connected to a SC after 2009[11]. As long as no omitted factor changed differently across treatment and control groups after 2009, the OLS estimation of Equation (1) is unbiased, so the fitted values of the dependent variable can be used as an exogenous explanatory variable in the second stage Equation (2).

Equation (1') explores an alternative instrumental variable (IV) for internet penetration where $Cable_{i,t}$ is the data capacity of SCs in country i at time t. This specification controls for time- and country-fixed effects ($\emptyset_t$ and ϑ_i) and other variables ($X_{i,t}$). Equation (2) is the second-stage equation. $Macro_{i,t}$ captures selected macroeconomic indicators for country i at time t (real GDP per capita growth, total factor and labor productivity, and sector shares in value added and employment as a measure to capture structural change effects). $\widehat{Internet}_{i,t}$ is the fitted, not the actual, internet penetration rate from the first stage equations (1) and (1'). $X_{i,t}$ is a vector of the same control variables used in the relevant first-stage equation.

While the two IVs underlying the first-stage equations relate to SCs, they are different as they capture different dimensions of internet connectivity. Whether one captures merely being directly connected to a submarine cable after 2009 (i.e., the extensive margin of technology adoption), the other captures the intensity of such connectivity as measured by the capacity of SCs (the intensive margin). In this sense, equation (1) has the advantage of more explicitly exploring the exogenous SCs arrival and adoption process, whereas equation (1') has the advantage of capturing the intensity of adoption within the group of adopters (mostly coastal countries as noted earlier). So, the two IVs entail different samples. Given all these differences, they are not substitutes and the analysis places the same importance on both.[12]

Data

Annex Table A1 provides details on the data used in the empirical analysis. The data on internet penetration (percent of the population with access to the internet) is obtained from the International Telecommunication Union. SCs data is obtained from *cablemap.info*[13], an online platform mapping the world's submarine telecommunications cable networks. The data is cross-checked in the official websites of the SCs and, using the mapping tools *Mapbox* and *OpenStreetMap,* SCs maps for SSA are generated for the periods pre- and post-2009. The macroeconomic variables of interest are obtained from the International Monetary Fund-IMF

[11] From now onwards, the expression "after 2009" includes the year 2009.

[12] The extensive and intensive margin of technology adoption is a widely studied phenomenon. For example, Comin and Mestieri (2010), studying the impact of technology adoption on economic growth, find that the variability across countries in the intensive margin is higher than in the extensive margin. The cross-country variation in intensive margin of adoption accounts for around 40 percent of the variation of per capita income.

[13] Accessible at https://cablemap.info/_default.aspx

(real per capita GDP, PPP 2011 international dollar), United Nations (sector value added, constant 2010 prices), World Bank and International Labor Organization (sector employment shares and productivity), and Penn World Tables 9.1 (total employment measured as number of persons engaged).

Annex Table A2 presents the list of countries and years covered in the sample. The largest sample used in the baseline regression analysis covers 33 SSA countries over the period 2005-2017. The baseline specification, among other control variables, includes the World Bank's Country Policy and Institutional Assessment (CPIA) rating for which consistent data is available from the year 2005.[14] In the robustness analysis, the paper runs alternative specifications with alternative institutional variables covering more years and countries. Unlike other candidates, the CPIA is based on actual policies rather than on perceptions, and the underlying series is desirably less correlated with the proposed IVs explored in the paper. Two different subsamples are also considered depending on which instrumental variable is used. As a few countries lack internet penetration data for some years, the resulting dataset is an unbalanced panel.

Annex Table A3 presents summary statistics for the main variables of interest. Overall, the data contains significant variation across countries and over time. We focus particularly on years 2008 (the most recent pre-treatment year) and 2017 (the most recent post-treatment year for which data is largely available). The average internet penetration rate for the full sample increases from 5.0 percent in 2008 to 24.9 percent in 2017. Penetration rates in 2017 range from 1.3 percent to 62 percent, with a higher average for treatment countries (28.3 percent) than control countries (18.2 percent). Interestingly, between 2008 and 2017 treatment countries experienced a faster increase in penetration (22 percentage points) than control countries (15.2 percentage points).

Other key variables in the analysis also exhibit significant variation. The average year-on-year per capita GDP growth in 2000-2017 is 1.9 percent in the full sample, with slightly higher growth for treatment countries (2 percent) than control countries (1.7 percent). In 2017, compared to 2008, the average per capita GDP growth declined in treatment countries, while remaining broadly unchanged in control countries. There are also slight differences in the composition of total output and employment in 2017. Treatment countries have a higher service share of total value added (25.8 percent) and employment (41.5 percent), compared to control countries (22.4 percent and 29.2 percent, respectively). These and other differences between treatment and control countries may reflect factors beyond internet penetration (e.g., income-driven structural transformation), reinforcing the importance of identifying causality as we attempt with the empirical strategy described earlier.

[14] The CPIA database covers all SSA countries except South Africa, Botswana, Namibia, Mauritius, Equatorial Guinea, Eswatini, Gabon and Seychelles. The countries in the sample altogether account for slightly over 75 percent of SSA's 2017 nominal GDP in PPP terms.

V. RESULTS

A. Instrument Relevance and Exogeneity

Instrument relevance (first-stage regressions)

The paper assesses first the extent to which the arrival of SCs affected internet penetration. Figure 6 above showed that internet penetration increased substantially after 2009. Annex Table A4 reports the results of a simple DiD model, which shows that internet penetration rose by 4.4 percentage points more in treatment group countries than in control group countries after 2009. In Annex Table A5 (column 1) further DiD analysis is performed in a regression framework that allows to control for potential confounding variables. The DiD estimate is lower (3.2) but remains sizeable and statistically significant. An important highlight is that the dummy for being directly connected to SCs (variable *Treatment*) is in itself insignificant, turning significant only when interacted with *Post2009,* which gives the DiD effect relevant for the underlying identification strategy discussed earlier. The finding of a positive DiD effect is similar to the evidence in Cariolle (2018), although this paper explores a slightly modified SCs experiment.

A key reason why the control group (landlocked) countries experienced a slower increase in internet penetration was discussed earlier and relates to the limited telecommunications network spillovers from neighboring coastal countries and existing collaboration agreements. The estimated coefficients are statistically significant and carry the expected sign for four control variables (real per capita GDP, years of schooling, institutional quality and credit to the private sector). For the remaining controls, the coefficients are not significant. In column 2 of Annex Table A5, we estimate the extent to which internet penetration responds to the installed capacity of SCs within the sample of coastal countries. The estimated coefficient (0.0008) is positive and highly statistically significant. The coefficients are also significant on real per capita GDP and trade openness (albeit with unexpected sign), while for the remaining control variables are insignificant. Data measurement issues, as well as different samples under the two IVs, could explain the unexpected signs and insignificance of some control variables.

The above findings suggest that the timing of arrival of SCs and the data capacity of SCs available in each country are relevant instrumental variables for internet penetration, as also corroborated by the high F statistics in the underlying regressions. Both variables pass the instrument relevance test whether scrutinized through the standard F statistic or the Montiel-Pflueger (robust) F statistic considered more reliable in testing weak instrumental variables.

Instrument exogeneity

The analysis also presents evidence that both the arrival of SCs and the data capacity of the cables are exogenous to macroeconomic variables such as GDP. A chart of average income by cable (Figure 7) and a scatterplot of income and cable capacity (Figure 8) do not show a

systematic relationship. The absence of a relationship is not surprising given that, as mentioned earlier: (i) SCs are deployed regionally to take advantage of economies of scale, so it is rare that a cable skips a country along its route; (ii) SCs are usually internationally funded instead of domestically funded; and (iii) the most deterministic factor of cable capacity is the technology available at the time.

Figure 7. SSA: Income Distribution by Submarine Cable

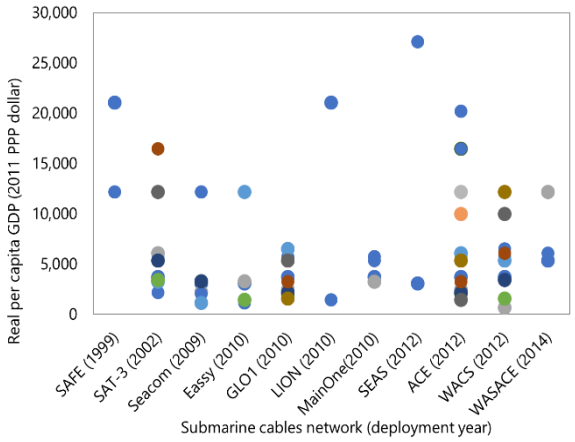
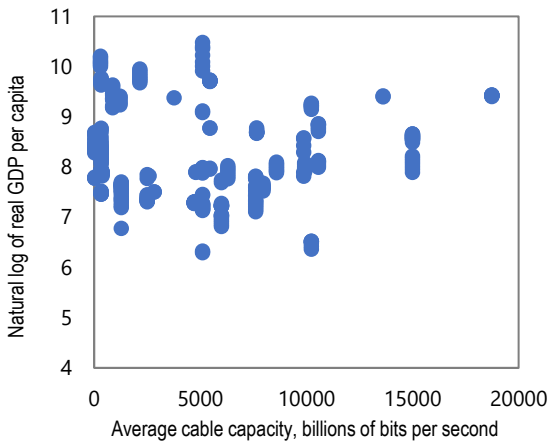


Figure 8. SSA: Correlation between Income and Cable Capacity



Sources: World Bank, Greg's Cable Map, and authors' calculations.
Note: Each dot represents one country in one year. Each color denotes a different country in LHS chart.

B. Macroeconomic Impact

Growth effects

The findings on the impact of internet penetration on real per capita GDP growth are mixed. Using as instrumental variable the arrival of SCs after 2009, there is not a statistically significant impact. However, there is a significant and large impact when employing as instrumental variable the capacity of the cables. Each percentage point increase in internet penetration leads to a 0.37 percentage point increase in real per capita GDP growth, all else equal.[15] Table 1 reports the detailed results. The control variables that are statistically significant broadly carry the expected sign. In particular, the negative coefficient on GDP suggests that lower income countries grow faster than higher income ones. Also, higher trade openness and fiscal balance are associated with higher growth. The estimated large impact of internet penetration on growth potentially reflects the fact that this paper's sample consists fully of SSA countries whereas other studies consider mostly advanced economies. In SSA

[15] Several studies estimate the relationship between growth and broadband internet penetration. The estimated coefficients range between 0.05 and 0.2, with very few exceptions of negative coefficients. However, most of the studies rely on a broader country sample (mostly non-SSA), and do not address the endogeneity between income and internet penetration. The ITU (2012) and World Bank (2016) provide an extensive review of the empirical work.

internet penetration has arguably not reached the saturation point—as noted earlier and as discussed in Katz and Callorda (2018)—and hence yields higher growth returns than in advanced economies.

Table 1. Internet Penetration and GDP Growth

Dependent variable:	Real per capita GDP growth	
	(1: IV1)	(2: IV2)
Internet penetration (% of population)	0.0530	0.3749**
	(0.0628)	(0.1661)
Lagged real per capita GDP (log)	-0.8094	-23.6526***
	(0.6866)	(5.3957)
Investment (% of GDP)	0.0655***	0.0484
	(0.0215)	(0.0748)
Mean years of schooling	0.0331	-2.2996
	(0.1594)	(2.8593)
Trade openness (% of GDP)	0.0215*	0.1212***
	(0.0112)	(0.0298)
Institutional quality (CPIA rating, higher=better)	2.1367***	1.8312
	(0.6999)	(1.2742)
Terms of trade (change)	0.0207*	0.0185
	(0.0117)	(0.0115)
Population density (people per square km)	0.0020	-0.0058
	(0.0024)	(0.0393)
Credit to private sector (% of GDP)	-0.0716**	-0.1920
	(0.0301)	(0.1569)
Fiscal balance (% of GDP)	0.0967***	0.0823*
	(0.0327)	(0.0443)
Sample	2005-17	2005-17
Year/Country fixed effects	No [1/]	Yes
Observations	395	265
R-squared	0.1119	0.3174
No. of countries	32	21

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2). Difference-in-differences specification in the first stage of IV1.

1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Output structure effects

The findings are also mixed with regards to the impact on the sectoral composition of output. The estimates are significant when employing as instrumental variable the arrival of SCs after 2009, but not the cable capacity. As reported in Table 2, higher internet penetration significantly increases the share of services in total value added (column 1), while reducing the share of industry (column 3). The impact is insignificant for agriculture (column 5). The control variables that are statistically significant mostly carry the expected sign. For example, higher income is associated with higher industry share and lower agriculture share (columns 3 and 5). This is consistent with the empirical findings reviewed in Matthess and Kunkel (2020) pointing that the income elasticity in agriculture is lower than in manufacturing, leading to a declining share of the former as income grows. Higher schooling achievement is associated with lower agriculture share (column 5), consistent with the view that structural change tends to move output away from low-skill labor sectors. Also, higher trade openness has a positive and negative effect on industry and agriculture value added, respectively (columns 3 and 5).[16]

The factors behind the positive and negative impact of internet penetration on the share of services and industry, respectively, are an empirical matter that goes beyond this paper. A potential explanation relates to the relative prices and productivity channel of structural transformation.[17] In addition to promoting new type of services, internet penetration, as it will be shown in further findings later, increases the productivity (potentially reducing relative prices and increasing the demand) of some existing services, and therefore increasing the services share of output. All else equal, this results in the declining share of other sectors.[18]

A more detailed look at the subsectors within services and industry is presented in Table 3.[19] The estimates suggest that the positive impact found on services (reported in Table 2) is driven by the subsectors of "transportation, storage and communication". This is not surprising given that most internet-related businesses are classified in the national accounting framework within communication-related subsectors, but there is a potential that transportation services could also benefit from higher internet penetration. Within industry, no significant impact for the

[16] As discussed in Matthess and Kunkel (2020), the current empirical literature is not straightforward on the effect of international trade on structural change, which in the past was believed to be driven mostly by the sector where countries held comparative advantage. With the emergence of new organizational form of international trade (eg., Global Value Chain and related power structure within), identifying the relationship has become more an empirical than theoretical matter.

[17] The channels of structural transformation were discussed in the literature review section.

[18] It is worth noting that the findings suggest that the "share" of industry value added (i.e., its relative importance on total value added), not its level, decline.

[19] On Table 3 the analysis is limited by available data with regards to how far the sectors of "services" and "industry" can be disaggregated. The data does not allow the disaggregation of "trade, restaurants and hotels" and "transportation, communication and storage".

three subsectors ("mining and utilities", "manufacturing", and "construction") are detected, although the signs of the estimated coefficients are negative and consistent with the coefficient for industry reported in Table 2. That is, while the share of industry significantly declines with internet penetration (Table 2), the declines in the three industry subsectors are too small to be statistically significant individually (Table 3).

Table 2. Internet Penetration and Sector Value Added Shares

Dependent variables:	Services and trade		Industry		Agriculture	
	(1: IV1)	(2: IV2)	(3: IV1)	(4: IV2)	(5: IV1)	(6: IV2)
Individuals using the Internet (% of population)	0.1634*	0.0162	-0.3233**	-0.0796	0.0365	0.0626
	(0.0875)	(0.1763)	(0.1359)	(0.2226)	(0.1176)	(0.1205)
Lagged real per capita GDP (log)	1.4160	1.7902	7.7093*	11.0699**	-6.7858***	-8.5469**
	(1.6148)	(3.5898)	(4.1317)	(4.5717)	(2.3792)	(3.8618)
Investment (% of GDP)	-0.0284	0.0680***	0.0186	-0.1275***	0.0915	0.0473
	(0.0997)	(0.0199)	(0.1145)	(0.0487)	(0.1084)	(0.0333)
Mean years of schooling	0.1823	-0.2663	0.8988	2.3146	-2.0119**	1.0066
	(0.6427)	(1.5653)	(1.1629)	(2.1766)	(0.9646)	(1.1473)
Trade openness (% of GDP)	-0.0055	-0.0325	0.1343***	0.0453	-0.1613***	-0.0284
	(0.0296)	(0.0219)	(0.0515)	(0.0426)	(0.0339)	(0.0253)
Institutional quality (CPIA rating, higher=better)	0.9403	-0.5990	-4.5714	0.2175	-3.0569	-2.0673**
	(1.8348)	(0.9941)	(2.7859)	(1.4045)	(2.1557)	(0.9546)
Terms of trade (change)	-0.0077	-0.0056	0.0199	0.0149	-0.0140	-0.0172***
	(0.0151)	(0.0069)	(0.0202)	(0.0109)	(0.0157)	(0.0061)
Sample	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17
Year/Country fixed effects	No[1]	Yes	No[1]	Yes	No[1]	Yes
Observations	406	268	406	268	406	268
R-squared	0.1345	0.9495	0.4454	0.9406	0.5755	0.9736
No. of countries	33	21	33	21	33	21

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2). Difference-in-differences specification in the first stage of IV1. Dependent variables are in share of total value added.

1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Table 3. Internet Penetration and Sector Value Added Shares Cont.

Dependent variables:	Services		Industry		
	Trade, restaurants, hotels	Transport, storage & communication	Mining and utilities	Manufacturing	Construction
	(1)	(2)	(3)	(4)	(5)
Individuals using the Internet (% of population)	0.0498	0.1136**	-0.2099	-0.0597	-0.0537
	(0.0809)	(0.0490)	(0.1644)	(0.0654)	(0.0567)
Lagged real per capita GDP (log)	2.5791*	-1.1631	6.7153*	-1.1245	2.1185***
	(1.4016)	(0.9148)	(3.8731)	(1.4851)	(0.7471)
Investment (% of GDP)	-0.0468	0.0184	0.0421	-0.0913*	0.0677
	(0.1051)	(0.0297)	(0.1157)	(0.0551)	(0.0421)
Mean years of schooling	-0.5526	0.7349***	0.1397	0.5701	0.1891
	(0.6544)	(0.2551)	(0.9784)	(0.5893)	(0.2751)
Trade openness (% of GDP)	-0.0252	0.0197	0.1456***	-0.0051	-0.0062
	(0.0273)	(0.0180)	(0.0549)	(0.0222)	(0.0144)
Institutional quality (CPIA rating, higher=better)	-0.0241	0.9644	-6.5299**	0.3344	1.6241**
	(1.4915)	(0.9639)	(2.6736)	(1.4162)	(0.8131)
Terms of trade (change)	-0.0049	-0.0027	0.0005	0.0056	0.0137**
	(0.0129)	(0.0075)	(0.0174)	(0.0135)	(0.0068)
Sample	2005-17	2005-17	2005-17	2005-17	2005-17
Year/Country fixed effects	No[1]	No[1]	No[1]	No[1]	No[1]
Observations	406	406	406	406	406
R-squared	0.0865	0.3327	0.4763	0.0626	0.3443
No. of countries	33	33	33	33	33

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1). Dependent variables are in share of total value added.

1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Employment structure effects

As reported in Table 4, there is no significant impact—positive or negative—on the level of total employment (columns 1 and 2).[20] However, there are notable sectoral dynamics. Internet penetration does not have a significant impact on industry employment as a share of total employment (columns 3 and 4), but it does have a significantly positive impact on services employment, and a negative impact on agriculture employment. The estimates suggest that each percentage point increase in internet penetration leads to an increase of 0.3-0.4 percentage points in the share of services employment (columns 7 and 8) and a reduction of 0.4 percentage

[20] This finding contrasts with Dieppe (2020) who reports a small negative impact on employment from technology-driven improvements for emerging markets and developing country economies. Different sample coverage, estimation methodology and measure of technology shock could explain the difference.

points in the share of agriculture employment (column 5).[21] This suggests that the decline in the agriculture share is broadly offset by the increase in the services share. Furthermore, within services, the impact is higher for females than for males (columns 9-12). Specifically, increased internet penetration is associated with a larger increase in the share of women working in the services sector—the shift to more employment in services is two and half times larger for women than men.

While the findings of employment share shift towards services is consistent with the view that the "servicification" of manufacturing, including from higher digitalization, is pushing employment towards the services sector (Matthess and Kunkel 2020), fully understanding the drivers would benefit from further research. An important question is whether the shift towards services is driven by displaced agriculture workers, or result from industry workers moving into services and, simultaneously, agriculture workers moving into industry.

Table 4. Internet Penetration and Sector Employment Shares

Dependent variables:	Total employment (log)		Percent of total employment									
			Industry		Agriculture		Services		Services, Male		Services, Female	
	(1: IV1)	(2: IV2)	(3: IV1)	(4: IV2)	(5: IV1)	(6: IV2)	(7: IV1)	(8: IV2)	(9: IV1)	(10: IV2)	(11: IV1)	(12: IV2)
Individuals using the Internet (% of population)	0.0150	0.0003	0.0629	-0.0305	-0.4336**	-0.4124	0.3377**	0.3627**	0.1968*	0.2431	0.4944***	0.4962***
	(0.0124)	(0.0036)	(0.0821)	(0.0768)	(0.2034)	(0.4012)	(0.1321)	(0.1556)	(0.1106)	(0.1579)	(0.1752)	(0.1632)
Lagged real per capita GDP (log)	-0.3503	0.1080	3.2641**	3.6296*	-12.9278***	2.8195	13.7796***	-2.1886	11.2349***	-2.7987	16.8909***	-1.6997
	(0.4557)	(0.0795)	(1.5109)	(2.1618)	(3.3230)	(11.0100)	(2.6638)	(4.6890)	(2.2891)	(4.4350)	(3.7181)	(5.1854)
Investment (% of GDP)	0.0257	-0.0005	-0.0319	0.0012	0.5338***	0.0334	-0.2781**	-0.0365	-0.2336**	-0.0469	-0.3456**	-0.0250
	(0.0164)	(0.0007)	(0.0706)	(0.0191)	(0.1804)	(0.0653)	(0.1184)	(0.0441)	(0.1149)	(0.0402)	(0.1433)	(0.0565)
Mean years of schooling	0.2031*	0.0051	-0.3556	-0.2901	-0.0116	3.2641	-0.8498	-3.8713**	-0.0610	-3.0809**	-1.8883	-4.8883*
	(0.1135)	(0.0284)	(0.7198)	(0.5447)	(1.5112)	(2.8878)	(1.0119)	(1.8852)	(0.9113)	(1.5527)	(1.2781)	(2.5561)
Trade openness (% of GDP)	-0.0163	-0.0009*	0.0420	0.0088	-0.2818***	-0.0730	0.0515	0.0562**	0.0278	0.0450*	0.0839	0.0628**
	(0.0099)	(0.0005)	(0.0312)	(0.0113)	(0.0874)	(0.0498)	(0.0580)	(0.0248)	(0.0616)	(0.0234)	(0.0658)	(0.0289)
Institutional quality (CPIA rating, higher=better)	-0.1706	-0.0443***	0.9652	-1.4338**	-8.1720*	1.9829	2.1187	-1.4920	0.3203	-1.1813	4.8166	-1.6220
	(0.6315)	(0.0168)	(2.1154)	(0.5890)	(4.5872)	(2.1663)	(3.3216)	(1.3153)	(2.7491)	(1.2457)	(4.3954)	(1.4242)
Terms of trade (change)	0.0050**	0.0000	-0.0151	0.0017	0.0693**	0.0013	-0.0289	-0.0051	-0.0329**	-0.0069	-0.0257	-0.0026
	(0.0025)	(0.0001)	(0.0120)	(0.0031)	(0.0305)	(0.0118)	(0.0182)	(0.0071)	(0.0160)	(0.0064)	(0.0270)	(0.0081)
Sample	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17
Year/Country fixed effects	No[1]	Yes	No[1]	Yes	No[1]	Yes	No[1]	Yes	No[1]	Yes	No[1]	Yes
Observations	401	268	406	268	406	268	406	268	406	268	406	268
R-squared	0.1295	0.9997	0.2269	0.9774	0.5875	0.9723	0.5840	0.9761	0.5192	0.9670	0.5858	0.9815
No. of countries	32	21	33	21	33	21	33	21	33	21	33	21

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2).

1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

[21] This is consistent with the finding that employment losses are more likely in sectors where tasks are easily automated. Several studies find evidence of increased employment in the services sector where automation of existing jobs is less likely compared to other sectors (Acemoglu 1999; Autor and Dorn 2009; Goos et al. 2014).

Productivity effects

The paper also assesses whether the positive effect of internet penetration on growth (reported in Table 1) can be explained by higher productivity. It looks at both aggregate total factor productivity (TFP), and aggregate and sector-level labor productivity measured as output per employment.[22] The findings show that, on aggregate, higher internet penetration has a positive and significant impact on TFP. As reported in Table 5, each percentage point increase in internet penetration leads to around 1.5 percent increase in TFP (column 1). With regards to labor productivity, the impact is insignificant at the aggregate level (column 2). However, when disaggregating the analysis into different subsectors, there is a significant and positive impact on utilities, trade and transportation.[23] For every percentage point increase in internet penetration, labor productivity increases by 6.9 percent, 1.4 percent and 4.9 percent, respectively, in the sectors of utilities, trade and transportation (columns 6, 8 and 9). Surprisingly, the coefficient on agriculture productivity is negative and statistically significant (column 3). This is counterintuitive considering the hope that digital adoption will contribute to raise farm productivity in SSA. Measurement errors from higher informality in agriculture may bias the estimated coefficient.[24] We find no significant impacts for the remaining sectors. Table 5 only reports the results using the cable capacity IV. The results are not significant when using the submarine cable arrival IV or specifying the dependent variables in terms of growth rates.

[22] Due to data availability for labor productivity, sector-level regressions include fewer (11) countries.

[23] The results are insignificant when a broader sectoral classification (comprising only "agriculture", "industry" and "services") is used (results not reported here). The reported significance at a more disaggregated level on Table 5 suggests that productivity in different subsectors within "industry" or "services" can be impacted differently by digitalization even if the more aggregate sector-level impact is insignificant.

[24] Measurement error in the dependent variable should not bias the estimate under the assumption that the error is random with respect to any independent variables. This assumption is unlikely to hold in the case of the regression in column (2) where the dependent variable entails a measurement of agriculture output and employment in SSA, which is highly informal as shown in the ILO (2018) data. Measurement errors from high informality are likely negatively correlated with other explanatory variables capturing income, institutional and educational development. This has the potential to cause a downward bias in the estimated coefficient.

Table 5. Internet Penetration and Productivity

Dependent variables:	Total factor productivity (log)	Labor productivity (log)	Labor productivity by sector (log)								
			Agriculture	Mining	Manufacturing	Utility	Construction	Trade	Transportation	Finance	Other services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Individuals using the Internet (% of population)	0.0147***	-0.0030	-0.0212***	0.0206	0.0302	0.0687**	-0.0057	0.0143**	0.0493**	-0.0322	-0.0077
	(0.0035)	(0.0054)	(0.0058)	(0.0368)	(0.0206)	(0.0285)	(0.0165)	(0.0069)	(0.0223)	(0.0267)	(0.0136)
Lagged real per capita GDP (log)	0.1260	1.0201***	0.8381**	2.8752***	1.3360*	-1.5320	0.2280	1.1676***	0.1912	-0.6084**	0.8542**
	(0.2087)	(0.2950)	(0.4065)	(0.9389)	(0.7000)	(0.9879)	(0.5263)	(0.3730)	(0.7421)	(0.2602)	(0.4171)
Investment (% of GDP)	-0.0079***	-0.0034*	-0.0045*	-0.0130**	-0.0024	-0.0036	-0.0024	-0.0008	0.0037	-0.0062	-0.0018
	(0.0022)	(0.0019)	(0.0026)	(0.0054)	(0.0063)	(0.0051)	(0.0031)	(0.0021)	(0.0027)	(0.0058)	(0.0034)
Mean years of schooling	-0.0059	-0.1458**	-0.2079	-0.8701***	-0.0807	-0.8415***	0.1350	-0.0052	-0.0623	0.3876	-0.0023
	(0.0854)	(0.0689)	(0.1361)	(0.3121)	(0.1759)	(0.2263)	(0.2094)	(0.1254)	(0.2598)	(0.2553)	(0.1598)
Trade openness (% of GDP)	-0.0001	0.0006	-0.0008	0.0241***	-0.0154***	0.0067*	-0.0030*	-0.0008	-0.0049	-0.0021	-0.0006
	(0.0011)	(0.0007)	(0.0011)	(0.0042)	(0.0021)	(0.0034)	(0.0016)	(0.0019)	(0.0032)	(0.0022)	(0.0031)
Institutional quality (CPIA rating, higher=better)	-0.0047	0.0193	0.0900	0.2406	-0.4007*	-0.0636	-0.0708	-0.0761	-0.1199	0.2162***	-0.1611
	(0.0635)	(0.0779)	(0.0981)	(0.1642)	(0.2144)	(0.1593)	(0.0774)	(0.0961)	(0.1970)	(0.0719)	(0.1144)
Terms of trade (change)	0.0002	0.0004	-0.0003	-0.0008	-0.0015	-0.0015	-0.0001	0.0011	0.0021	0.0012	-0.0013
	(0.0003)	(0.0006)	(0.0006)	(0.0021)	(0.0015)	(0.0018)	(0.0005)	(0.0008)	(0.0017)	(0.0009)	(0.0009)
Population density (people per square km)	0.0021*	-0.0000	-0.0007	-0.0148***	0.0121***	0.0034	-0.0087***	-0.0047**	0.0017	0.0001	0.0007
	(0.0012)	(0.0011)	(0.0019)	(0.0052)	(0.0035)	(0.0042)	(0.0025)	(0.0019)	(0.0020)	(0.0034)	(0.0019)
Credit to private sector (% of GDP)	0.0075	0.0000	0.0009	0.0034	0.0116	0.0235**	0.0137**	0.0136***	0.0331**	0.0116**	-0.0060
	(0.0052)	(0.0043)	(0.0064)	(0.0181)	(0.0166)	(0.0111)	(0.0067)	(0.0052)	(0.0150)	(0.0058)	(0.0109)
Fiscal balance (% of GDP)	-0.0029**	0.0037*	0.0062**	0.0222***	-0.0010	0.0029	-0.0086***	0.0022	-0.0081**	-0.0016	-0.0026
	(0.0013)	(0.0021)	(0.0028)	(0.0072)	(0.0060)	(0.0067)	(0.0033)	(0.0035)	(0.0035)	(0.0032)	(0.0028)
Sample	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17	2005-17
Year/Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	151	138	138	138	138	138	138	138	138	138	138
R-squared	0.4799	0.9831	0.9674	0.9536	0.9130	0.9558	0.9818	0.9597	0.7611	0.9599	0.8329
No. of countries	12	11	11	11	11	11	11	11	11	11	11

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. IV is defined as total installed capacity of submarine cables (IV2). The results (not reported here) are not significant when using the submarine cable arrival (IV1). The sample in column (1) is not the same as in other columns. For the sector regressions, the number of countries drops to 11 due to lack of data for many SSA countries in the World Bank's (Dieppe 2020) productivity database.

Overall, the findings reported above suggest that internet penetration does significantly affect real per capita GDP growth, value added and employment structure, and productivity. The results are sensitive to the IV used for internet penetration. As discussed earlier, while the two IVs relate to SCs, they are different as they measure different margins of SCs adoption (extensive vs. intensive) and entail different samples. One potential indirect implication from the findings is that the intensive margin seems to matter more to growth and productivity, while the extensive margin seems to matter more to output and employment composition effects.[25] However, the different composition of the samples under the two IVs challenges this interpretation. Further in-depth research would be warranted.

VI. Robustness

This section presents further checks that the identification strategy and estimation methodology used in the paper address concerns over endogeneity and related issues. The analysis shows that the key assumption in the identification strategy based on the arrival of SCs—the parallel trend assumption—holds. It also shows that the estimated coefficients remain significant and broadly unchanged after controlling directly or indirectly for key confounding shocks such as the 2014-15 terms of trade shock and the 2008-09 global financial crisis. The results are

[25] See an earlier footnote in Section IV on the extensive and intensive margin on technology adoption.

broadly unchanged when controlling for access to electricity—an important correlate of both internet access and GDP growth—and when dropping from the sample the high-digitalized-and-high-income countries. Finally, the estimates are also robust to restricting the sample endpoint to earlier years. However, some of the regressions are sensitive to using alternative control variables for institutions. These robustness checks are discussed in detail below.

Parallel trend assumption

If the pre- and post-2009 and treatment-control identification strategy is reasonable, there should be no significant DiD coefficient within the pre-treatment (2000-2008) subperiods. This is inspected in Annex Table A6. The DiD estimate is insignificant in the pre-treatment subperiod where we set the year 2005 (roughly halfway in the pre-treatment subperiod) as time cutoff. This backs the parallel trend assumption: had the large-scale arrival of SCs not occurred after 2009, the difference in internet penetration rates between treatment and control group countries would have remained broadly unchanged.

The 2014-15 terms of trade shock and the 2008-09 global financial crisis

The treatment period (post-2009) coincides with two external shocks that might have affected growth differently across treated and control countries: (1) the 2014-15 terms of trade shock and (2) the 2008-09 financial crisis. The paper addresses the potential bias from (1) by including terms of trade as a control variable in the regressions. To address (2), alternative regressions are run excluding countries heavily affected by the financial crisis—South Africa, Nigeria, Mauritius, Angola and Kenya.[26] Annex Table A7 shows the results. The estimated coefficient on internet penetration remains sizeable and significant for the growth regression.[27] It becomes insignificant for the productivity regression; however, this may reflect the much lower sample size compared to the already small initial size.

Access to electricity

Digitalization is intrinsically related to electricity given that digital equipment typically requires electrical power to function. A possible concern is that countries with higher SCs capacity and greater availability of other digital equipment may be those with higher access to electricity. To the extent that electricity is a direct input to production (and hence growth), omitting it from the regression could bias the results. The paper argued earlier that the arrival of the SCs and their capacity are exogenous to countries' income status which would in principle imply exogeneity to electricity access as well. To further support this, Annex Table

[26] South Africa and Mauritius (where growth declined substantially amid the global financial crisis) are from the start not included in the sample used in this study due to lack of CPIA data. For this particular robustness check exercise, we further exclude three countries where growth fell markedly amid the financial crisis- Nigeria, Angola and Kenya.

[27] Results for the remaining value added and employment regressions (not shown in the paper) also do not change significantly.

A8 shows the results from growth and productivity regressions where access to electricity is added as a control variable. The estimated coefficient of internet penetration remains broadly unchanged and statistically significant for both the growth and productivity regressions.

Income-driven digitalization

A legitimate concern relates to reverse causality. Fast-growing countries and those with higher income may have adopted digital policies earlier than other countries, which could upward bias the estimated impact of internet penetration on growth. It was argued earlier that the IVs are exogenous to growth and income. Also, the cable arrival IV regression is set in a DiD scheme which removes fixed effects such as some countries growing constantly faster than others over time. In addition, the analysis directly controls for determinants of growth as done in standard growth regressions. As a further evidence against reverse causality concerns, Annex Table A9 runs additional regressions excluding from the sample the "champions" of internet penetration, i.e., those countries where the very high penetration rates (above 40 percent in 2017, compared to the sample average of 25 percent) may partly reflect their higher income status: Botswana, Cabo Verde, Côte D'Ivoire, Gabon, Mauritius, Namibia, Nigeria, Senegal, Seychelles and South Africa.[28] The estimated coefficient of internet penetration on growth and productivity remains statistically significant and sizeable.

Sensitivity to sample end-year (treatment window)

Most treatment countries in our sample became connected between 2009 and 2012, and the dramatic increase in total cable capacity also occurred during the same period. However, the identification strategy takes as the treatment period the whole post-2008 period. This assumes that part of the increase in internet penetration during 2013-17 reflects the arrival of SCs and the increase in their capacity in earlier years. This is a reasonable assumption given the commonly observed lagged effects in adjusting to technology shocks. However, the results hold also when relaxing this assumption. Annex Table A10 reports the estimated coefficients from the growth regression using the cable capacity IV. The coefficients remain statistically significant irrespective of the sample end-year, the only exception being the year 2013. Also, as expected, the size of the coefficients decline as the sample extends towards more recent years during which the number of new cables and their capacity remained broadly unchanged.[29]

Sample selection

As noted earlier, the control variable for institutional quality restricts the sample to CPIA countries. While CPIA countries account for sightly over 75 percent of SSA GDP, the omission

[28] These are all "Middle Income Countries" according to the IMF classification. By excluding them from the sample, we focus on a subsample of countries with smaller variation in income.

[29] These findings are similar for the productivity regression, except that the coefficient is also significant when 2013 is the sample end-year.

of some countries like South Africa (19 percent of SSA GDP) could cause sample selection bias. To determine whether this could be an issue, the paper considers another measure of institutions from the World Bank's World Governance Indicators—the index of regulatory quality (IRQ)—which have greater country and year coverage.[30] Under such specifications, the impact of internet penetration on real per capita GDP growth and total factor productivity becomes statistically insignificant, albeit maintaining the expected sign and with still significant IV in the first stage regression (not reported). However, the coefficients are significant for labor productivity at the aggregate level, and for selected sectors as before, with the difference that manufacturing now gains significance and transportation loses it (Annex Table A11). The significance remains broadly unchanged for the regressions on sector value added (Annex Table A12-13) and employment shares (Annex Table A14). The higher correlation between the IRQ and the IVs (as compared to the CPIA), which leads to less significant IVs, could explain the loss of significance in some second-stage regressions. Therefore, an important caveat regarding the external validity of the baseline findings is that they may not necessarily generalize to non-CPIA countries.

VII. CONCLUSIONS AND POLICY IMPLICATIONS

This paper assessed the impact of internet penetration on macroeconomic performance in SSA. It finds a significant, large, and robust impact. Each percentage point increase in internet penetration leads to a 0.37 percentage point increase in real per capita GDP growth, all else equal. This large estimate potentially reflects the fact that the sample consists entirely of SSA countries where returns to digital technologies are likely higher, whereas studies finding lower estimates mostly cover non-SSA and, in many instances, more advanced economies. As the recently arrived submarine cables in SSA use new technology with higher data capacity, quality effects may also be driving the larger estimate.

The paper also finds evidence that higher internet penetration increases labor productivity, particularly in the sectors of utilities, trade, and transportation. It further finds that higher internet penetration significantly boosts the share of services (as percent of total value added), with a positive impact on the subsectors of "transportation, storage and communication". Yet, higher internet penetration reduces the share of industry, while there is no significant impact on the share of agriculture in output.

The paper finds that internet penetration does not have a significant impact on the total level of employment. However, it finds a significant impact on services employment as a share of total employment. The services share increases by 0.3-0.4 percentage points for each percentage point increase in internet penetration. Furthermore, the impact on services employment is higher for female than for male—the shift to services is two and half times

[30] Alternative indicators within the World Governance Indicators were considered and the results are not significantly different than when using the IRQ.

larger for women than men. As for agriculture and industry employment, we find, respectively, a negative and insignificant impact.

The findings on employment imply that higher internet penetration affects the composition, but not the level, of total employment. More specifically, net gains in services sector jobs broadly offset agricultural job losses, while the impact on industry jobs is neutral. An important question is whether the new jobs created in services are filled by displaced agriculture workers or result from industry workers moving into services and agriculture workers moving into industry. Another question is whether the increase in the share of services employment also reflects higher wage returns in services driven by more digital-ready workers. Moreover, within the female sample, it would be useful to distinguish the increase in the share of services employment due to (i) new women joining the labor market from (ii) women moving from other sectors into services due to higher flexibility in service jobs allowing them to better reconcile household work. Tackling these questions, which merit a separate follow-up research, can provide additional insights in designing policies on how best to invest in people's skills to yield the benefits of digitalization.

The results are sensitive to the instrumental variable used for internet penetration. For example, the estimated effects on GDP growth and productivity only hold when employing as instrumental variable the submarine cable capacity. While the two instrumental variables relate to submarine cables, they are different as they measure different margins of submarine cables adoption (extensive vs. intensive) and entail different samples. One potential indirect implication could be that the intensive margin, i.e., the intensity of connectivity to submarine cables, matters more to growth and productivity than merely being connected to the cables (the extensive margin). However, the different composition of the samples under the two instrumental variables challenges this interpretation. Further in-depth research would be warranted.

The findings overall support the renewed attention to digitalization reforms in SSA, including to expand access to high-quality and reliable internet connectivity, as a tool to boost growth, raise labor productivity and foster diversification of the production structure. The latter is particularly relevant in SSA where economies are highly concentrated in the production of primary commodities subject to significant cyclical volatility. From a labor market perspective, the results suggest that digitalization reforms could lead to compositional shifts in employment towards services for both men and women. For SSA women particularly, higher internet connectivity could pave the way for a more flexible work arrangement and more online business opportunities, which may explain why it is mostly women in the study that drive the employment shift towards services. However, digitalization does not happen by itself. It requires complementary investments in foundational infrastructure, such as electricity, and critical investments in digital literacy skills to ensure workers can take advantage of the new digital opportunities.

References

Acemoglu, D. (1999). Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence. American Economic Review 89 (5): 1259-78.

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

Autor, D. H. (2001). Wiring the labor market. The Journal of Economic Perspectives, 15(1), 25-40.

Autor, D. H., and Dorn, D. (2009). Inequality and Specialization: The Growth of Low-skill Service Jobs in the United States. NBER Working Paper Series, 15150.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. The quarterly journal of economics, 106(2), 407-443.

Brookings Institution 2017. "Digitalization and the American Workforce." Brookings, Washington, DC.

Cariolle, J. (2018), Telecommunication Submarine-Cable Deployment and the Digital Divide in Sub Saharan Africa. Ferdi Working Paper P241, May 2019.

Comin, D. and Mestieri, M. (2010), An Intensive Exploration of Technology Diffusion. NBER Working Paper 16379.

Crandall, R., W. Lehr and R. Litan (2007), The Effects of Broadband Deployment on Output and Employment: A Cross-sectional Analysis of U.S. Data", Issues in Economic Policy, The Brookings Institution, No 6.

Czernich, N., Falck, O., Kretschmer, T., and Woessmann, L. (2011). Broadband infrastructure and economic growth. The Economic Journal, 121(552), 505-532. Cite 801

Dieppe, A. (2020). Global Productivity: Trends, Drivers and Policies. Advance Edition. Washington, DC: World Bank. License: Creative Commons Attribution CC BY 3.0 IGO.

Duflo, E. (2001), Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. American Economic Review. September.

Franklin, M., P. Stam, and T. Clayton (2009), ICT Impact Assessment by Linking Data, Economic and Labour Market Review, Vol 3 No.10, 18-27.

Goos, M., A. Manning, and A. Salomons. (2014). Explaining Job Polarization: Routine-Biased
Technological Change and Offshoring. American Economic Review 104 (8): 2509–26.

Hjort J. and Poulsen J. (2019), The Arrival of Fast Internet and Employment in Africa. American Economic Review, 109 (3): 1032-79.

Houngbonon, Georges V.; Liang, Julienne (2018) : The Impact of Broadband Internet on Employment in France, 29th European Regional Conference of the International Telecommunications Society (ITS): "Towards a digital future: Turning technology into markets?", Trento, Italy, 1st - 4th August 2018, International Telecommunications Society (ITS), Trento

Hu, Y., and Yao, J. (2019). Illuminating economic growth. International Monetary Fund.

International Finance Corporation (2019, August). Bringing Africa Up to High Speed. Retrieved from: https://www.ifc.org/wps/wcm/connect/news_ext_content/ifc_external_corporate_site/news+and+events/news/cm-stories/cm-connecting-africa#page0.

International Labour Organization (2018). Women and men in the informal economy: a statistical picture (third edition)/International Labour Office – Geneva: International Labor Organization, 2018

Katz, R. (2012). Impact of broadband on the economy. ITU Broadband Series.

Katz, R., and Callorda, F. (2018). The economic contribution of broadband, digitization and ICT regulation. International Telecommunication Union, published in Switzerland, Geneva.

Klonner, S. and Nolen, P. (2008). Does ICT Benefit the Poor? Evidence from South Africa. Working paper.

Koutroumpis, P. (2009). The economic impact of broadband on growth: A simultaneous approach. Telecommunications policy, 33(9), 471-485.

Lehr, W., C. A. Osorio, S. E. Gillett, and M. A. Sirbu (2006), Measuring Broadband's Economic Impact, Final report prepared for the U.S. Department of Commerce, Economic Development Administration.

Mark Muro, Sifan Liu, Jacob Whiton, and Siddharth Kulkarni. (2017). Digitalization and the American Workforce. Metropolitan Policy Program at Brookings.

Mathess, M. and Kunkel, S. (2020). Structural change and digitalization in developing countries: conceptually linking the two transformations. Technology in Society 63 (2020) 101428

Minges, Michael. 2015. Exploring the Relationship between Broadband and Economic Growth. WDR 2016 Background Paper; World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/23638 License: CC BY 3.0 IGO.

Ouedraogo, R. and Sy. Amadou (2020). Can Digitalization Help Deter Corruption in Africa? IMF Working Paper No. 20/68

Oulton, N., (2010), Long Term Implications of the ICT Revolution: Applying the Lessons of Growth Theory and Growth Accounting, CEP Discussion Paper No 1027.

Paunov, C., and Rollo, V. (2016). Has the internet fostered inclusive innovation in the developing world? World Development, 78, 587-609.

Qiang, C. Z. W., Rossotto, C. M., and Kimura, K. (2009). Economic Impacts of Broadband. Information and Communications for Development 2009: Extending reach and increasing impact, 3, 35-50.

Roller, L. H., and Waverman, L. (2001). Telecommunications infrastructure and economic development: A simultaneous approach. American economic review, 91(4), 909-923.

Sovbetov, Y. (2018). Impact of Digital Economy on Female Employment: Evidence from Turkey. International Economic Journal, 32(2), 256-270.

Stryszowski, P. (2012). The impact of internet in OECD countries. OECD Publishing

Tisdell, C. (2017). Information technology's impacts on productivity and welfare: a review. International journal of social economics.

World Bank. (2016). World Development Report 2016: Digital Dividends. Washington, DC.

World Bank. (2018). Innovative Business Models for Expanding Fiber-Optic Networks and Closing the Access Gaps.

Zaballos, A. G., and López-Rivas, R. (2012). Socioeconomic Impact of Broadband in Latin American and Caribbean Countries. Inter-American Development Bank, 220.

Annex

Table A1: Data and Sources

Variable	Sources
Internet use (percent of population using the internet)	International Telecommunication Union
Per capita GDP, constant prices, PPP 2011 international dollar	International Monetary Fund
Terms of trade (percent change)	International Monetary Fund
Fiscal Balance (percent of GDP)	International Monetary Fund
Share of agriculture in total value added (percent)	United Nations, authors' calculations
Share of services in total value added[1] (percent)	United Nations, authors' calculations
Share of industry in total value added[2] (percent)	United Nations, authors' calculations
Average years of schooling (number of years)	United Nations
Share of industry in total employment (percent)	International Labour Organization
Share of services in total employment (percent)	International Labour Organization
Share of services in total male employment (percent)	International Labour Organization
Share of services in total female employment (percent)	International Labour Organization
Total employment (number of persons engaged)	Penn World Tables 9.1
Total investment (percent of GDP)	International Monetary Fund
Trade openness (percent of GDP)	World Bank (WDI)
Population density (people per square Km of land area)	World Bank (WDI)
CPIA quality of public administration rating	World Bank (CPIA)
Regulatory quality indicator	World Bank (WGI)
Labor productivity (GDP per employment, PPP-adjusted)	World Bank: Dieppe (2020)
Total factor productivity (constant 2011 prices)	Penn World Tables 9.1
Credit to private sector (percent of GDP)	World Bank (WDI)
Submarine cables	Greg's Cable Map

Note: The vintage of most databases is as of October 2019.

1/ Services include transports, storage, communications, wholesale, retail trade, restaurants and hotels.

2/ Industry includes mining, manufacturing, utilities and construction.

Table A2: Country Coverage, 2005-17

Country	Included in the sample	Country	Included in the sample
Angola	Yes*	Kenya	Yes*
Benin	Yes*	Lesotho	Yes
Botswana	No	Liberia	No
Burkina Faso	Yes	Madagascar	Yes*
Burundi	Yes	Malawi	Yes
Cabo Verde	Yes*	Mali	Yes
Cameroon	Yes*	Mauritius	No
Central African Republic	Yes	Mozambique	Yes*
Chad	Yes	Namibia	No
Comoros	Yes*	Niger	Yes
Congo, Dem. Rep.	Yes*	Nigeria	Yes*
Congo, Rep.	No	Rwanda	Yes*
Cote d'Ivoire	Yes*	Sao Tome and Principe	No
Equatorial Guinea	No	Senegal	Yes*
Eritrea	Yes	Seychelles	No
Eswatini	No	Sierra Leone	Yes*
Ethiopia	No	South Africa	No
Gabon	No	Tanzania	Yes*
Gambia, The	Yes*	Togo	Yes*
Ghana	Yes*	Uganda	Yes*
Guinea	Yes*	Zambia	Yes
Guinea-Bissau	Yes*	Zimbabwe	No

Note: This study originated under the IMF's Spring 2020 "Regional Economic Outlook (REO) for Sub-Saharan Africa". The analysis covers SSA countries in line with the REO country groupings. Some countries are not included in the sample due to lack of data. They include non- CPIA countries and others for which internet penetration, or a specific control variable data are not available. All "Yes" countries make up the sample under the "being directly connected to a submarine cable" IV1. For the "cable capacity" IV2, the sample covers only coastal countries in this table that have a direct connection to a submarine cable ("Yes*" countries), including Rwanda and Uganda which have a direct connection despite being landlocked. In the sectoral productivity regressions (Table 5 in the main text), the number of countries drops to 11 (Angola, Cameroon, Ghana, Kenya, Mozambique, Nigeria, Rwanda, Senegal, Sierra Leone, Tanzania, Uganda) due to lack of data for many SSA countries in the World Bank's (Dieppe 2020) productivity database. Thus, some of the regression results are not directly comparable.

Table A3: Descriptive Statistics

	Mean	Mean 2008	Mean 2017	Std. Dev.	N
Full Sample					
Internet use (percent of population)	8.6	5.0	24.9	12.2	789
Real GDP (Log)	7.9	7.9	8.0	1.0	836
Growth of real GDP per capita (%)	1.9	2.1	1.2	4.9	792
Share of agriculture of total value added (%)	23.8	24.0	21.4	15.5	792
Share of service in total value added (percent)	23.3	23.5	24.6	8.0	792
Share of manufacturing in total value added (%)	26.9	27.1	26.3	14.5	792
Share of manufacturing employment in total (%)	12.2	11.9	12.7	6.9	817
Share of service employment in total (%)	34.5	33.7	37.2	15.9	817
Share of service in total male employment (%)	32.6	32.0	34.4	13.5	817
Share of service in total female employment (%)	37.2	36.3	40.9	20.6	817
Total investment (percent of GDP)	22.8	23.5	22.3	10.2	808
Trade openness (percent of GDP)	75.5	80.5	72.8	38.2	770
population density (people per square km)	97.8	94.5	116.8	122.7	829
CPIA quality of public administration rating	2.9	2.9	2.9	0.5	495
Average years of schooling	4.6	4.6	5.3	2.1	775
Terms of trade (percent change)	2.7	1.4	3.6	15.0	774
Credit to private sector (percent of GDP)	26.0	19.4	72.7	114.8	809
Treatment					
Internet use (percent of population)	10.2	6.3	28.3	13.5	521
Real GDP (Log)	8.0	8.0	8.2	1.0	551
Growth of real GDP per capita (percent)	2.0	2.7	1.4	4.7	522
Share of agriculture of total value added (percent)	23.2	23.4	20.9	15.6	522
Share of service in total value added (percent)	24.2	24.5	25.8	8.8	522
Share of manufacturing in total value added (%)	26.4	26.3	25.7	15.3	522
Share of manufacturing employment in total (%)	12.7	12.5	13.0	6.6	532
Share of service employment in total (%)	38.6	37.9	41.5	14.9	532
Share of service in total male employment (%)	36.3	35.9	38.2	11.9	532
Share of service in total female employment (%)	41.9	41.1	45.9	20.6	532
Total investment (percent of GDP)	22.9	24.3	22.6	10.1	532
Trade openness (percent of GDP)	76.5	82.1	74.0	39.3	518
population density (people per square Km)	116.8	113.7	137.0	134.0	551
CPIA quality of public administration rating	3.0	2.9	3.0	0.5	313
Average years of schooling	4.9	4.8	5.7	1.9	510
Terms of trade (percent change)	2.8	-0.5	6.3	15.2	522
Credit to private sector (percent of GDP)	31.2	22.4	94.9	139.3	546
Control					
Internet use (percent of population)	5.4	2.5	18.2	8.3	268
Real GDP (Log)	7.6	7.6	7.7	0.9	285
Growth of real GDP per capita (percent)	1.7	0.7	0.6	5.4	270
Share of agriculture of total value added (percent)	25.2	25.2	22.3	15.3	270
Share of service in total value added (percent)	21.5	21.5	22.4	5.9	270
Share of manufacturing in total value added (%)	27.8	28.6	27.5	12.9	270
Share of manufacturing employment in total (%)	11.2	10.9	12.1	7.4	285
Share of service employment in total (percent)	26.8	25.7	29.2	14.9	285
Share of service in total male employment (%)	25.6	24.6	27.4	13.6	285
Share of service in total female employment (%)	28.5	27.3	31.5	17.6	285
Total investment (percent of GDP)	22.5	22.1	21.9	10.4	276
Trade openness (percent of GDP)	73.4	77.2	70.1	35.9	252
population density (people per square km)	60.1	57.3	75.0	85.2	278
CPIA quality of public administration rating	2.8	2.8	2.8	0.5	182
Average years of schooling	4.1	4.1	4.7	2.3	265
Terms of trade (percent change)	2.4	5.1	-2.1	14.7	252
Credit to private sector (percent of GDP)	15.3	13.1	19.0	9.6	263

Sources: ITU, IMF, World Bank, ILO and authors' calculations

Note: "Treatment" countries are defined as those directly connected to a submarine cable (mostly coastal, with the exception of Rwanda and Uganda). "Control" countries are those not directly connected to a submarine cable (landlocked).

Table A4: Difference-in-Differences in Internet Penetration (percent)

	(1) Treatment	(2) Control	(1) - (2) Difference
(a) 2009-17	16.4	9.5	7.0
(b) 2002-2008	4.1	1.5	2.6
(a) - (b) Difference	12.4	8.0	4.4

Table A5: First Stage of the Two Instruments

Dependent variable: Internet penetration (%)	(1: IV1)	(2:IV2)
Treatment*Post2009	3.1707*	
	(1.7412)	
Treatment	-1.5677	
	(1.3320)	
Post2009	3.6847***	
	(1.1392)	
Submarine cable capacity (gbps per second)		0.0008***
		(0.0002)
Lagged real per capita GDP (log)	4.7816***	15.3078**
	(1.1640)	(7.4566)
Investment (% of GDP)	-0.0770	-0.0509
	(0.0684)	(0.1286)
Mean years of schooling	0.9560**	4.2246
	(0.4377)	(6.5026)
Trade openness (% of GDP)	-0.0319	-0.0917*
	(0.0217)	(0.0506)
Institutional quality (CPIA rating, higher=better)	2.8751**	1.0618
	(1.1846)	(2.8658)
Terms of trade (change)	-0.0205	-0.0119
	(0.0157)	(0.0177)
Population density (people per square km)	0.0046	-0.0028
	(0.0042)	(0.0794)
Credit to private sector (% of GDP)	0.2773***	0.3369
	(0.0777)	(0.3204)
Fiscal balance (% of GDP)	-0.0066	-0.0168
	(0.0678)	(0.0579)
Sample	2005-17	2005-17
Year/Country fixed effects	No [1/]	Yes
Observations	395	265
R-squared	0.5920	0.8520
F-statistic	12.79	102.95
Montiel-Pflueger effective F-statistic	14.548	14.675
No. of countries	32	21

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2).
1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Table A6: Placebo Test in 2005

Dependent variable: Internet penetration (%)	(1)	(2)	(3)	(4)
Treatment*Post2009	4.7937*		5.1257*	
	(2.7515)		(2.7397)	
Treatment	-1.4109		-1.6345	
	(1.5433)		(1.6898)	
Post2009	6.4293***		6.4703***	
	(2.0726)		(2.0391)	
Treatment*Post2005		0.8906		0.7301
		(0.8380)		(0.7926)
Treatment		0.2518		0.3721
		(0.8758)		(0.9628)
Post 2005		1.1700**		1.3001**
		(0.5711)		(0.5930)
Sample	2001-17	2001-08	2002-17	2002-08
Year/Country fixed effects	No [1/]	No [1/]	No [1/]	No [1/]
Observations	640	293	591	253
R-squared	0.6144	0.5781	0.6204	0.5913
Montiel-Pflueger effective F-statistic	20.408	4.001	21.81	3.985
No. of countries	40	40	39	39

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The sample in column (1) covers the period 2000-2017, column (2) covers the pre-treatment period (2000-08). Compared to earlier regressions, these omit institutional quality (CPIA) as a control variable in order to include the years 2000-2004 for which CPIA data is not available. Columns (3) and (4) are the same as (1) and (2), respectively, except that (3) and (4) include an additional control for institutions using the WGI indicator (regulatory quality index) which has greater country and year coverage than the CPIA. The results are consistent.
1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Table A7: Internet Use and the Arrival of Submarine Cables (removing countries heavily affected by the 2008 global financial crisis)

Dependent variables:	Real per capita GDP growth		Total factor productivity (log)
	(1: IV1)	(2: IV2)	(3: IV2)
Internet use (% pop.)	0.1033	0.5324***	0.0028
	(0.0707)	(0.1863)	(0.0029)
Sample	2005-17	2005-17	2005-18
Year/Country fixed effects	No[1]	Yes	Yes
Observations	360	230	116
R-squared	0.1306	0.1790	0.7058

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2). The difference-in-differences specification is in the first stage of IV1. Beyond South Africa and Mauritius which are not in the baseline sample, other countries that were heavily impacted by the 2008 GFC shocks are Nigeria, Angola and Kenya.
1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Table A8: Internet Penetration and Growth (controlling for access to electricity)

Dependent variables:	Real per capita GDP growth		Total factor productivity (log)
	(1: IV1)	(2: IV2)	(3: IV2)
Internet use (% pop.)	0.0396	0.3685**	0.0174***
	(0.0651)	(0.1655)	(0.0038)
Sample	2005-17	2005-17	2005-18
Year/Country fixed effects	No[1]	Yes	Yes
Observations	394	264	151
R-squared	0.1183	0.3226	0.4157

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2). Regressions include access to electricity as additional control. The difference-in-differences specification is in the first stage of IV1.
1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Table A9: Internet Penetration and Growth (excl. high internet penetration countries)

Dependent variable:	Real per capita GDP growth		Total factor productivity (log)
	(1: IV1)	(2: IV2)	(3: IV2)
Internet use (% pop.)	0.0794	0.6957**	0.0246*
	(0.0968)	(0.3246)	(0.0141)
Sample	2005-17	2005-17	2005-17
Year/Country fixed effects	No[1]	Yes	Yes
Observations	213	213	112
R-squared	0.1152	0.3317	0.4665

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as the interaction between being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2). Beyond those not in the baseline sample (Botswana, Mauritius, Namibia, South Africa, Gabon and Seychelles), other high internet penetration countries include Cabo Verde, Cote d'Ivoire, Nigeria and Senegal.
1/ The difference-in-differences specification removes fixed effects.

Table A10: Internet Penetration and Growth (sensitivity test on sample end-year)

Dependent variable:	Real per capita GDP growth					
Sample end-year:	2012	2013	2014	2015	2016	2017
Internet use (% pop.)	0.9608*	0.9709	0.8355*	0.6782**	0.5018**	0.3749**
	(0.5701)	(0.6560)	(0.4324)	(0.2766)	(0.2185)	(0.1661)
Sample	2005-12	2005-13	2005-14	2005-15	2005-16	2005-17
Year/Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	166	187	207	226	245	265
R-squared	0.3908	0.3095	0.2532	0.2468	0.2925	0.3174

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as total installed capacity of submarine cables.

Table A11: Internet Penetration and Productivity (replacing CPIA)

Dependent variables:	Labor	Labor productivity by sector (log)								
	productivity (log)	Agriculture	Mining	Manufacturing	Utility	Construction	Trade	Transportation	Finance	Other services
	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Individuals using the Internet (% of population)	0.0049*	0.0004	0.0405	0.0314***	0.0410***	0.0084	0.0133*	0.0235	-0.0167	-0.0026
	(0.0029)	(0.0050)	(0.0264)	(0.0121)	(0.0146)	(0.0076)	(0.0077)	(0.0207)	(0.0148)	(0.0059)
Lagged real per capita GDP (log)	1.3610***	1.1183***	3.0008**	1.8947**	-0.9639	0.7912	1.7109***	1.1396	-0.7362**	1.0436**
	(0.3002)	(0.4228)	(1.2622)	(0.8756)	(1.2492)	(0.5894)	(0.4441)	(0.9392)	(0.3299)	(0.4733)
Investment (% of GDP)	-0.0017	-0.0006	-0.0012	0.0070	0.0009	0.0057*	0.0019	0.0069	-0.0058	-0.0048**
	(0.0018)	(0.0028)	(0.0104)	(0.0072)	(0.0053)	(0.0033)	(0.0031)	(0.0044)	(0.0062)	(0.0022)
Mean years of schooling	-0.1841***	-0.1737**	-0.6084*	-0.1881	-0.9171***	-0.1598	-0.2536*	-0.5279**	0.1915	-0.1138
	(0.0614)	(0.0866)	(0.3210)	(0.1824)	(0.3191)	(0.1433)	(0.1377)	(0.2604)	(0.2048)	(0.0906)
Trade openness (% of GDP)	0.0005	-0.0007	0.0203***	-0.0185***	0.0018	-0.0048***	-0.0032	-0.0081**	-0.0016	-0.0010
	(0.0008)	(0.0016)	(0.0047)	(0.0054)	(0.0030)	(0.0016)	(0.0026)	(0.0040)	(0.0026)	(0.0017)
Regulatory Quality (estimate, higher = better)	0.0926	-0.0608	0.4978	-0.0854	0.0615	0.1690	-0.1489	0.0564	-0.4132	0.4998***
	(0.0877)	(0.1478)	(0.4376)	(0.2398)	(0.3028)	(0.1073)	(0.1550)	(0.1799)	(0.3563)	(0.1374)
Terms of trade (change)	0.0008	0.0003	-0.0008	0.0012	-0.0016	-0.0008	-0.0003	0.0010	-0.0001	-0.0013**
	(0.0009)	(0.0008)	(0.0026)	(0.0021)	(0.0025)	(0.0007)	(0.0011)	(0.0018)	(0.0005)	(0.0006)
Population density (people per square km)	-0.0002	0.0002	-0.0119**	0.0145***	0.0025	-0.0085***	-0.0041*	0.0010	0.0032	-0.0042***
	(0.0008)	(0.0015)	(0.0053)	(0.0056)	(0.0040)	(0.0024)	(0.0024)	(0.0031)	(0.0045)	(0.0015)
Credit to private sector (% of GDP)	-0.0016	0.0010	-0.0085	-0.0020	0.0031	-0.0045	0.0017	0.0098	0.0076**	-0.0010
	(0.0017)	(0.0034)	(0.0123)	(0.0083)	(0.0095)	(0.0045)	(0.0039)	(0.0080)	(0.0033)	(0.0030)
Fiscal balance (% of GDP)	0.0016	-0.0001	0.0153*	0.0008	0.0036	-0.0037	0.0000	-0.0071*	-0.0042	-0.0022
	(0.0028)	(0.0027)	(0.0078)	(0.0063)	(0.0085)	(0.0049)	(0.0037)	(0.0043)	(0.0031)	(0.0032)
Sample	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17
Year/Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	222	222	222	222	222	222	222	222	222	222
R-squared	0.9892	0.9834	0.9217	0.8958	0.9266	0.9570	0.9682	0.8361	0.9279	0.9369
No. of countries	14	14	14	14	14	14	14	14	14	14

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as total installed capacity of submarine cables (IV2).

1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Table A12: Internet Penetration and Sector Value Added (replacing CPIA)

Dependent variables:	Services and trade		Industry		Agriculture	
	(1: IV1)	(2: IV2)	(3: IV1)	(4: IV2)	(5: IV1)	(6: IV2)
Individuals using the Internet (% of population)	0.1912***	-0.0186	-0.3471***	0.1263	0.0292	0.0319
	(0.0627)	(0.1234)	(0.0940)	(0.2291)	(0.0699)	(0.1378)
Lagged real per capita GDP (log)	-2.7514	4.7946**	14.6462***	12.7758**	-8.6703***	-7.0418**
	(1.6856)	(2.2678)	(1.9666)	(5.5590)	(1.3623)	(3.1670)
Investment (% of GDP)	-0.0033	0.0483**	0.2182**	-0.0565	-0.0626	-0.0016
	(0.0898)	(0.0245)	(0.0958)	(0.0611)	(0.0939)	(0.0382)
Mean years of schooling	0.1480	0.2541	-0.7667	-4.2389	-1.1283	0.3268
	(0.8220)	(1.9808)	(0.7868)	(3.9231)	(0.7860)	(1.9590)
Trade openness (% of GDP)	0.0575	-0.0369**	-0.0325	0.0664*	-0.0701*	-0.0285
	(0.0364)	(0.0169)	(0.0697)	(0.0400)	(0.0362)	(0.0260)
Regulatory Quality (estimate, higher = better)	1.4704	0.1242	-8.7018***	2.2932	-0.0586	1.2841
	(1.9969)	(1.6396)	(2.4750)	(4.1653)	(1.7092)	(1.5839)
Terms of trade (change)	-0.0191	-0.0006	0.0566***	0.0503*	-0.0125	-0.0203**
	(0.0133)	(0.0078)	(0.0219)	(0.0264)	(0.0126)	(0.0092)
Sample	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17
Year/Country fixed effects	No 1/	Yes	No 1/	Yes	No 1/	Yes
Observations	602	405	602	405	602	405
R-squared	0.1786	0.9636	0.6358	0.9544	0.7475	0.9814
No. of countries	40	26	40	26	40	26

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2). Difference-in-differences specification in the first stage of IV1. Dependent variables are in share of total value added.

1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Table A13: Internet Penetration and Sector Value Added, Cont. (replacing CPIA)

Dependent variables:	Services		Industry		
	Trade, restaurants, hotels	Transport, storage & communication	Mining and utilities	Manufacturing	Construction
	(1)	(2)	(3)	(4)	(5)
Individuals using the Internet (% of population)	0.0726	0.1185***	-0.3482***	-0.0163	0.0173
	(0.0561)	(0.0306)	(0.1343)	(0.0543)	(0.0340)
natural log of rgdp = L,	-0.9309	-1.8204***	12.3612***	1.3617	0.9234*
	(1.4822)	(0.6844)	(2.6360)	(1.4099)	(0.5579)
Investment (% of GDP)	-0.0483	0.0450	0.2729**	-0.1937**	0.1390***
	(0.0861)	(0.0363)	(0.1190)	(0.0922)	(0.0402)
Mean years of schooling	-0.2665	0.4145	-0.1802	-0.4612	-0.1254
	(0.7840)	(0.2737)	(0.8186)	(0.6342)	(0.2902)
Trade openness (% of GDP)	0.0349	0.0226*	-0.0441	0.0278	-0.0163
	(0.0319)	(0.0125)	(0.0714)	(0.0282)	(0.0120)
Regulatory Quality, Estimate	0.9906	0.4798	-9.9634***	2.0967	-0.8350
	(1.6755)	(0.7623)	(2.7793)	(1.5424)	(0.8283)
Terms of trade (change)	-0.0180*	-0.0011	0.0515**	-0.0050	0.0101
	(0.0108)	(0.0064)	(0.0260)	(0.0125)	(0.0070)
Sample	2002-17	2002-17	2002-17	2002-17	2002-17
Year FE	No 1/	No 1/	No 1/	No 1/	No 1/
Observations	602	602	602	602	602
R-squared	0.0697	0.3007	0.5273	0.1315	0.2204
No. of countries	40	40	40	40	40

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1). Dependent variables are in share of total value added.

1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.

Table A14: Internet Penetration and Employment Shares (replacing CPIA)

Dependent variables:	Total employment (log)		Percent of total employment									
			Industry		Agriculture		Services		Services, Male		Services, Female	
	(1: IV1)	(2: IV2)	(3: IV1)	(4: IV2)	(5: IV1)	(6: IV2)	(7: IV1)	(8: IV2)	(9: IV1)	(10: IV2)	(11: IV1)	(12: IV2)
Individuals using the Internet (% of population)	0.0157	-0.0021	0.0011	-0.1067	-0.2835**	-0.1763	0.1695*	0.2257*	0.0671	0.1833	0.2833**	0.2299
	(0.0095)	(0.0034)	(0.0593)	(0.0880)	(0.1332)	(0.2569)	(0.0883)	(0.1330)	(0.0773)	(0.1349)	(0.1187)	(0.1566)
natural log of rgdp = L,	-0.6224**	0.0860	3.2645**	4.4111***	-9.8688***	-0.2155	11.3999***	0.7508	9.8675***	-1.2145	13.2744***	2.7448
	(0.2511)	(0.0796)	(1.3062)	(1.7082)	(3.6762)	(6.3574)	(3.2018)	(4.8629)	(3.1600)	(4.8976)	(3.8456)	(5.0934)
Investment (% of GDP)	0.0351**	0.0001	-0.1317**	0.0195	0.5002***	0.0263	-0.3016***	-0.0402	-0.2825***	-0.0370	-0.3479**	-0.0457
	(0.0176)	(0.0008)	(0.0611)	(0.0177)	(0.1402)	(0.0630)	(0.1141)	(0.0453)	(0.0981)	(0.0442)	(0.1575)	(0.0500)
Mean years of schooling	0.1361	-0.0734	-0.1690	-0.8646	-0.4105	2.3007	-0.4820	-1.8523	0.2195	-1.5649	-1.4627	-1.6709
	(0.1238)	(0.0535)	(0.6841)	(1.3962)	(1.5215)	(3.6248)	(1.0732)	(2.6664)	(0.9935)	(2.4692)	(1.3106)	(3.2338)
Trade openness (% of GDP)	-0.0243***	-0.0017***	0.0593*	-0.0047	-0.2464***	-0.0871*	0.0577	0.0606*	0.0452	0.0538*	0.0789	0.0564
	(0.0071)	(0.0005)	(0.0330)	(0.0153)	(0.0742)	(0.0481)	(0.0529)	(0.0339)	(0.0548)	(0.0306)	(0.0676)	(0.0414)
Regulatory Quality, Estimate	0.0040	-0.0289	3.8455**	-1.9680*	-10.5546***	-3.0042	6.8634**	2.1003	3.7481	1.9052	11.5933***	2.7854
	(0.4639)	(0.0349)	(1.6328)	(1.0521)	(3.2970)	(3.0108)	(2.7310)	(2.1741)	(2.4852)	(2.1601)	(3.4924)	(2.2851)
Terms of trade (change)	0.0062**	-0.0000	-0.0204**	-0.0012	0.0776***	-0.0063	-0.0431***	0.0057	-0.0448***	0.0050	-0.0438**	0.0038
	(0.0025)	(0.0003)	(0.0098)	(0.0053)	(0.0229)	(0.0165)	(0.0155)	(0.0094)	(0.0154)	(0.0088)	(0.0211)	(0.0107)
Sample	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17	2002-17
Year/Country fixed effects	No[1]	Yes	No[1]	Yes	No[1]	Yes	No[1]	Yes	No[1]	Yes	No[1]	Yes
Observations	597	405	587	390	587	390	587	390	587	390	587	390
R-squared	0.3865	0.9994	0.4569	0.9744	0.7034	0.9795	0.6992	0.9746	0.6726	0.9637	0.6572	0.9822
No. of countries	39	26	39	25	39	25	39	25	39	25	39	25

Note: ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively. Standard errors are in parenthesis and are clustered at the country level. The IV is defined as being directly connected to a submarine cable in post-2009 period (IV1) or total installed capacity of submarine cables (IV2).

1/ The two-period two-group difference-in-differences specification in the first stage removes fixed effects.