WORKING paper

BANQUE DE FRANCE
EUROSYSTÈME

Digital Vulnerability and Performance of Firms in Developing Countries

Joël Cariolle[1], Maëlan Le Goff[2], Olivier Santoni[3]

September 2018, WP #709

Abstract

This paper provides evidence of the positive effects of email use on firm performance. Using a pseudo-panel from repeated cross-sections of more than 30,000 firms, surveyed between 2006 and 2014 in some 40 developing and transition economies, we adopt an IV approach emphasizing firm's digital vulnerability to seismic shocks upon the telecommunications submarine cable network. Our results show that a 10% increase in email use incidence at the location level raises by 37-38% their total annual sales, by 22-23% sales per worker, by 12-14% the number of full-time workers. While these results are driven by the service sector, we find that a greater incidence of email use in locations increases the number of non-production and unskilled production workers in manufactures. Last, we find heterogenous but weaker evidence on the effect of email use on firms' direct exports share in total sales.

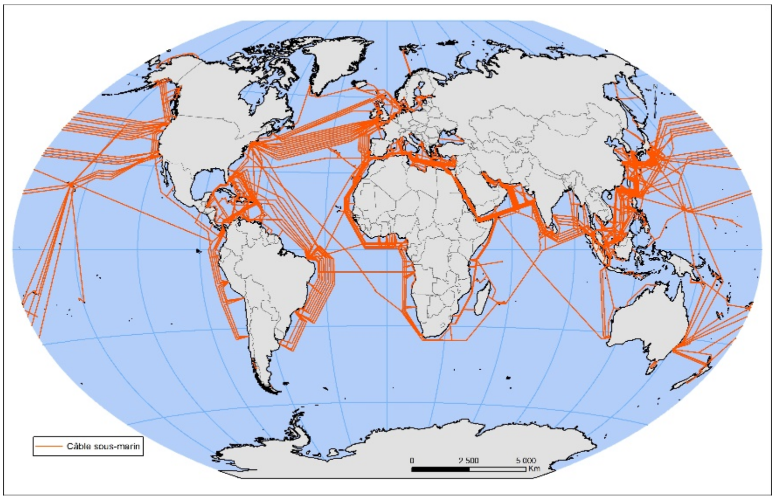
Keywords: Internet, Email, Firm performance, Submarine cables, Seaquakes

JEL classification: F02, O11, O33, O18, L25, L96

[1] Foundation for Studies and Research on International Development (FERDI), joel.cariolle@ferdi.fr;
[2] Banque de France, maelan.legoff@banque-france.fr;
[3] Foundation for Studies and Research on International Development (FERDI), olivier.santoni@ferdi.fr.
We would like to thank Raphaël Paris (Laboratoire Magma Volcan – UCA, CNRS) for the insightful discussions we had on seismic risk location. We also thank Simone Bertoli, Michaël Goujon, Mathilde Maurel, Jaime de Melo, and participants at the Ferdi-Banque de France research seminar, and at the DIAL 2017 and AFSE 2017 conferences, for their valuable comments. We are also grateful to Martine Bouchut (CERDI-UCA) for helping us in the programming of some variable calculation.

NON-TECHNICAL SUMMARY

Over the last three decades, international connectivity underwent a dramatic improvement promoted by the laying of around 400 fibre submarine cables (SMCs). Nowadays, more than 99% of the world's telecommunications – Internet content, phone and video calls, classified diplomatic messages – passes through SMCs. SMC international networks now irrigate a USD 20.4 trillion industry and connect 3 billion Internet users across the world. As a consequence, almost all coastal developing and transition countries are now plugged into the global Internet through SMCs. Fast-growing Asian and South American countries have been rapidly connected to Northern economies, and Africa's digital isolation from the rest of the world has rapidly fallen since 2005. This rapid expansion of the international broadband infrastructure network and the following boom in Internet services raises strong expectations for many low-income countries' economic catch-up, notably, through its potential for fostering innovation, productivity, trade, and job creation.



Worldwide SMC deployment in 2018.

However, countries' higher dependence on SMCs for international telecommunications has increased their vulnerability to SMC faults. SMC faults have local-level economic consequences, by disorganizing the economic activity, and depriving populations that are remote from core connectivity infrastructures of a fast, stable and cost-effective access to telecommunications. In this regard, seismic activity represents a major threat to the SMC network integrity, by shaking violently the underwater body, provoking turbidity currents, landslides, and seabed sand waves. Seismic risk around SMCs therefore represent an important concern for operators and governments, which in addition to economic costs, face high maintenance, repair, rerouting and insurance costs related to SMC failures.

In the 2000s, some papers, focused on aggregated data or on firm level data, find a positive effect of Internet access on international trade, productivity and growth. These positive impacts are mainly be explained by a reduction in transaction costs and improvements in knowledge spill-overs. This paper brings additional insights into this area of research by providing new evidence on the local impact of email use by firms on four outcomes – their revenue, their labour productivity, their exports and the size of their workforce – in developing countries. In addition, an innovation in this paper lies in its instrumental variable (IV) approach, emphasizing the firm's vulnerability to shocks upon the SMC network. Our instrument set indeed reflects local spatial inequalities in terms of Internet access, by combining data on the SMC network's exposure to seismic shocks with the firm's location distance to international connectivity infrastructures. To examine to what extent email use shapes firm outcomes, IV within fixed-effects estimations are therefore conducted using WBES data on firms' characteristics, outcomes and Internet usage.

Based on a large sample of firms located in more than 130 municipalities/provinces in some 40 developing and transition countries, IV estimates provide evidence of a large and positive impact of email use on firm performance at the location level. Indeed, we find that a 10% increase in the local incidence of email use among firms increases by 37-38% their revenues, by 22-23% average sales per worker, and by 12-14% the average number of full-time permanent workers. While these results are driven by the service sector, we find that a greater incidence of email use among firms at the location level increases the number of non-production and unskilled production workers in manufactures. Last, we provide heterogenous but weaker evidence on the effect of email use on firms' exports share in total sales, perhaps because of the heterogeneous effects of reduced communication costs on exporting firms.

Vulnérabilité Numérique et Performance des Firmes dans les Pays en Développement

Résumé

Cet article met en évidence des effets positifs importants de l'utilisation des emails sur la performance des entreprises au niveau local. En utilisant des données en coupes transversales répétées entre 2006 et 2014 portant sur plus de 30000 entreprises dans une quarantaines de pays en développement et en transition, nous avons mené des estimations en variables instrumentales reflétant la vulnérabilité numérique des entreprises aux chocs sismiques proches des câbles sous-marins télécommunication. Nos résultats révèlent qu'une hausse de 10% de l'utilisation des emails au niveau des localités améliore les ventes annuelles des entreprises de 37-38%, les ventes par travailleurs de 22-23% et de 12-14% le nombre d'emplois équivalent temps plein. Bien que ces résultats semblent tirés par le secteur des services, nous trouvons qu'une plus large utilisation d'Internet au niveau local augmente le nombre de travailleurs non affectés à la production et de travailleurs non qualifiés affectés à la production au sein du secteur manufacturier. Enfin, nous trouvons un effet hétérogène mais moins robuste de l'utilisation des emails sur la part des exportations directes dans les ventes des entreprises.

Mots-clés : NTIC, câbles sous-marin, performance des entreprises, pays en développement

1- Introduction

Over the last three decades, international connectivity underwent a dramatic improvement promoted by the laying of around 400 fibre submarine cables (SMCs). Nowadays, more than 99% of the world's telecommunications – Internet content, phone and video calls, classified diplomatic messages – passes through SMCs, irrigating a USD 20.4 trillion industry and connecting 3 billion Internet users across the world (Towela & Tesfaye, 2015).

Almost all coastal developing and transition countries have now access to the global Internet through SMCs. Fast-growing Asian and South American countries have been rapidly connected to Northern economies, while Africa's digital isolation from the rest of the world has rapidly fallen since 2009 (Cariolle, 2018). This densification of the SMC network (Graph 1) has stimulated digital ecosystems worldwide, and raised strong expectations for many low-income countries' economic catch-up, notably through fostered innovation, productivity, trade, and job creation.

This paper provides new evidence on the local impact of the use of emails by firms on four key outcomes: their total revenue, their labour productivity, their exports, and the size of their workforce. In fact, to our knowledge, very few studies have focused on this specific usage of Internet and its consequence for the private sector development in developing and transition countries (Paunov & Rollo, 2015, 2016). In addition, an innovation in this paper lies in its instrumental variable (IV) approach, emphasizing the firm's vulnerability to shocks upon the SMC network. Our instrument set indeed reflects local spatial heterogeneities in ICT access and usage, by combining data on the SMC network's exposure to seismic shocks with the firm's location distance to international connectivity infrastructures. Our results stress that this source of digital vulnerability has important consequences on firm outcomes, and therefore seem to be a core element of the local Internet-production nexus.

In fact, countries' higher dependence on SMCs for international telecommunications has increased their vulnerability to SMC faults. SMC faults have local-level economic consequences, by disorganizing the economic activity, and depriving populations that are remote from core connectivity infrastructures of a fast, stable and cost-effective access to telecommunications (Malecki, 2002; Grubesic & Murray, 2006; Grubesic et al., 2003; Gorman et al., 2004; Buys et al., 2009; Cariolle, 2018). In this regard, seismic activity represents a major threat to the SMC network integrity, by shaking violently the underwater body, provoking turbidity currents, landslides, and seabed sand waves (Soh et al, 2004; Carter et al, 2009; Clark, 2016; Pope et al, 2017; Aceto et al, 2018; Yincan et al, 2018). Seismic risk around SMCs therefore represent an important concern for operators and governments, which face high maintenance, repair, rerouting and insurance costs related to SMC failures (Carter et al, 2009; Widmer et al., 2010; Clark, 2016; Aceto et al., 2018).

The empirical analysis is based on data from the World Bank Enterprise Surveys on more than 30,000 surveyed firms, located in around 130 locations in some 40 developing and transition countries. We build a pseudo-panel dataset by aggregating repeated cross-sectional firm-level data at the location-level, and conduct IV within fixed-effects estimations and find large positive effects of email use incidence on local firm performance: a 10% increase in email use incidence within locations raises by 37-38% average firms' total sales, by 22-23% average sales per worker, by 12-14% the average number of full-time permanent workers. While these results are driven by the service sector, we find that a greater incidence of email use among firms at the location level increases the number of non-production and unskilled production workers in manufactures. This evidence stresses the positive local spillovers of Internet access, and nuances recent evidence on a skilled-biased digital revolution (Michaels et al, 2014; Akerman et al, 2015; Acemoglu and Restrepo, 2018; Hjort & Poulsen, 2019). Last, we provide mixed and somewhat weaker evidence on the effect of email use on firms' exports – highlighting a negative effect on service firms, but a positive effect on domestic SMEs and in intermediary-size cities – that could be explained by the heterogeneous effects of reduced communication costs on exporting firms (Melitz, 2003; Clarke & Wallsten, 2006; Fink et al, 2005, Bustos, 2011).

By controlling for location, year, country-year fixed effects, and by conducting a wide range of restrictions upon the sample composition and the instrument set calibration, we try to lower the concerns for omitted variable bias, reverse causality or measurement error. These tests are, among others, meant to reduce the risk that our results are not unduly biased by firms and infrastructure location choices, although we cannot entirely exclude the possibility of a (downward) bias in the estimated effects of email use on firm outcomes.

Graph 1. Worldwide SMC deployment, 1990-2005-2015.

1990 **2005** **2015**

Notes: Data on infrastructure deployment are drawn from Telegeography.

Our paper is organized as follows. In the next section, we review the literature on the macro and micro-level impacts of Internet and related telecommunication technologies on growth, trade and employment. In the third section, we present our data and some descriptive statistics. In the fourth section, we explain our identification strategy. The fifth section presents the main results and the last one concludes.

2- Literature review

ICTs are a general purpose technology expected to accelerate the development process by reducing transaction and information costs incurred by economic agents (Aker, 2017; Goldfarb & Tucker, 2019). We review below studies highlighting the prospects of a better access to ICTs in terms of growth, trade, and employment.

2.1. ICTs, productivity and growth

The literature pointing out the contribution of ICTs to growth and productivity in developed and developing economies is abundant. At the early stage of the Internet development, Litan and Rivlin (2001) were already projecting the positive impact of Internet access on cost-minimization in the US manufacturing and financial sectors, and on public service delivery. Their projections were confirmed for a larger sample of 21 OECD countries by Röller and Waverman (2001), who emphasize the positive contribution of the telecommunication infrastructures to their growth rates. Examining the impact of higher broadband penetration rates in 25 OECD countries, Czernich et al. (2011) get comparable and updated results. Choi and Yi (2009) provides additional evidence on the positive effect of Internet access on economic growth for a larger sample of 207 countries, while Niebel (2018) shows that this positive effect does not rely on the income group level.

At the micro-level, studies highlight the growth-potential of ICTs adoption by focusing on the effects of Internet access on firm performance, especially productivity and innovations.[1] Paunov and Rollo (2016, 2015) stress the knowledge-spillover effects of email use by firms on their productivity and innovation. In particular, they stress that these spillovers are stronger for the most productive firms, but also for firms that commonly engage less in innovation: non-exporters, single-plant firms, and firms located in smaller locations. Clarke et al. (2015) find similar evidence that Internet use increases firms' growth and labour productivity, particularly in the case of small firms.

Taken together, these studies emphasize the large returns to ICTs adoption and diffusion, driven by network effects and ICT-induced productivity spillovers, especially for smaller firms, not necessarily oriented towards world markets and innovation.

2.2. ICTs and trade

Another strand of the literature also looked at the effect of ICT on trade and FDI, generally arguing that the reduction of related communication and transport costs induced by their diffusion would

[1] For a review of these contributions, see Goldfarb and Tucker (2019).

accelerate the “death of distance” between countries and between economic agents (Brun et al, 2005).

While studies generally agree on the positive effect of ICTs on foreign exchanges, some findings question the relevance of the distance channel. Freund and Weinhold (2004, 2002) show that Internet stimulates bilateral exports, including service exports (see also Choi, 2010). Clarke and Wallsten (2006) find that this positive effect holds for developing countries only, in particular for exports from developing countries to developed countries. This positive effect does not however seem to be explained by the decrease in the distance, but rather by the increased competition from foreign markets (Freund and Weinhold, 2004). Fink et al (2005) nuance this view by showing that international variations in communication costs stimulate trade, but in a greater extent for differentiated products than for homogenous products. Therefore, this finding suggests that the expected effect of reduced distance on trade induced by ICT adoption depends on exported products’ characteristics, as also evidenced by Blum and Goldfarb (2007).

2.3 ICTs and employment

Recent papers (Michaels et al, 2014; Akerman et al, 2015; Acemoglu and Restrepo, 2016) support that the introduction and penetration of new ICTs in industrialized countries contribute to polarizing the labor market, increasing the demand for skilled and educated workers, at the expense of less skilled and less educated workers. Empirical evidence in developing countries are however scarcer.

To our knowledge, the strongest evidence on the ICT-employment nexus in developing countries has been provided by Hjort and Poulsen (2019), who show that increased broadband Internet penetration following the laying of SMCs in Africa has stimulated job creation by increasing firms’ entry, exports and productivity in the continent. Their results confirm the Internet’s potential for job creation, but also point to the risk of widening economic inequalities, induced by unequal access to employment for skilled and unskilled workers.

This literature review highlights that ICT adoption and diffusion in developed and developing countries has yielded important dividends in terms of growth, productivity, foreign market access and employment. Our results broadly confirm theses evidence, but point that the strength of these relationships may have been underestimated by overlooking how within-country spatial heterogeneities in ICT access and usage affect firm outcomes (Röller & Waverman, 2001; Dicken & Malmberg, 2001; Bjorkegren, 2019; Goldfarb and Tucker, 2019). In the next three sections, we present our model, the instrumental variable setup, and our results.

3- Model and Data

3.1. Model

To study the effect of email use by firms on their performance, we estimate the following general model:

$$Y_{j,l,i,t} = \alpha_0 + \alpha_1 Email_{j,l,i,t} + \alpha_2 X_{j,l,i,t} + \delta_t + (\gamma_j \times \delta_t) + \mu_s + \theta_l + \varepsilon_{j,l,i,t} \quad (1a)$$

Where the subscripts *j, l, i, t* respectively refer to the country, the location (municipality or province), the firm, and the year of survey. $Y_{j,l,i,t}$, $Email_{j,l,i,t}$, and $X_{j,l,i,t}$ are respectively firm's performance variables, a dummy variable of email use for business purposes (the variable of interest), and firm's characteristics. This equation also includes year (δ_t), country-year ($\gamma_j \times \delta_t$), industry (μ_s) and location (θ_l) fixed effects. $\varepsilon_{j,l,i,t}$ is a random error term.

To conduct within fixed-effect estimations, we follow Deaton (1985) and build a pseudo panel by averaging repeated cross-sectional firm-level data at the location level. This approach has a number of benefits:

i) it lowers the risk of omitted variable bias, including bias induced by firm-level time-invariant unobserved heterogeneity;

ii) it accounts for network externalities and/or spatial spillovers among firms located in the same place (Röller & Waverman, 2001; Dicken & Malmberg, 2001; Bjorkegren, 2019; Goldfarb and Tucker, 2019);

iii) it exploits information from firms that would have been excluded from pooled firm-level regressions because of missing information; and

iv) it transforms the dichotomous endogenous variable, $email_{j,l,i,t}$, into a continuous variable lying between 0 and 1, which reduces the concern for the efficiency of IV estimates using a linear probability model.

Moreover, because our IV varies at the location level, aggregating data by locations is also a way to neutralize spatial interactions between firms which would raise problems of over-identification if IV estimations were conducted with firm-level data. [2]

Therefore, after data aggregation, locations become observation units, the email-use and the industry dummies become incidence variables (lying between 0 and 1), while firms' outcomes and characteristics are location-level averages, yielding the following econometric specification:

$$Y_{j,l,t} = \alpha_0 + \alpha_1 Email_{j,l,t} + \alpha_2 X_{j,l,t} + \delta_t + (\gamma_j \times \delta_t) + \theta_l + \varepsilon_{j,l,t} \quad (1b)$$

[2] See sub-section 4.3 and Appendix C.

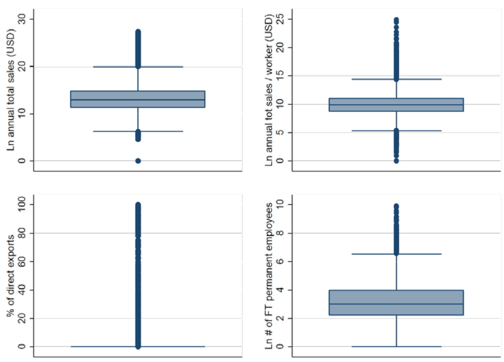
3.2. The data

All variables used in our model are drawn from the World Bank Enterprise Survey (WBES) harmonized cross-sectional dataset. These surveys cover a representative sample of a formal economy's manufacturing and service sectors. In each country, data were gathered by an extensive and internationally comparable questionnaire administered by face-to-face interviews with business owners and senior managers. The design of the survey is initially not suited for panel data analysis because of missing panel identification number. Data is therefore aggregated at the municipality level (province-level in low population density areas), using the sample weights[3], and yielding a pseudo-panel baseline sample of 257 observations, based on surveys conducted over more than 32,328 firms located in 128 locations, from 39 developing and transition economies. Appendix A reports information on variables' summary statistics and on the sample composition.

3.2.1. Dependent variables ($Y_{j,l,t}$)

We measure firm performance using alternatively four main outcomes variables: the logarithms of firm's total annual sales and sales per full-time employee (both converted into USD), the share of direct exports in total sales, and the logarithm of the number of permanent full-time employees.[4] Graphs 2 and 3 plot these variables' distribution at the firm and location levels, respectively. More detailed data on manufactures' workforce is used to deepen the analysis of the email-employment relationship: the manufactures' production and non-production workers, and on their skilled and unskilled production workers. Appendix A.1 reports summary statistics of these variables.

Graph 2. Distributions of firm outcomes, firm-level data before aggregation

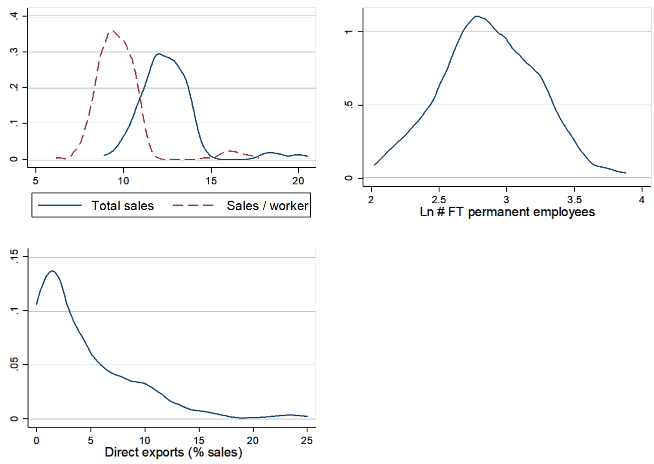


Data: World Bank Enterprise Surveys. Sample: 32,328 firm-level observations from 128 locations, 39 countries.

[3] The World Bank uses stratified random sampling, based on firm size, sector and location. The majority of firms tend to be small and medium sized, so that large firms are over-weighted because of their importance within most national economies.

[4] The variable sales per worker is the ratio of the firm's total sales over the number of FT permanent employees (adjusted for the number of temporary workers), transformed in logarithm.

Graph 3. Kernel densities of firm outcomes, location-level data after aggregation



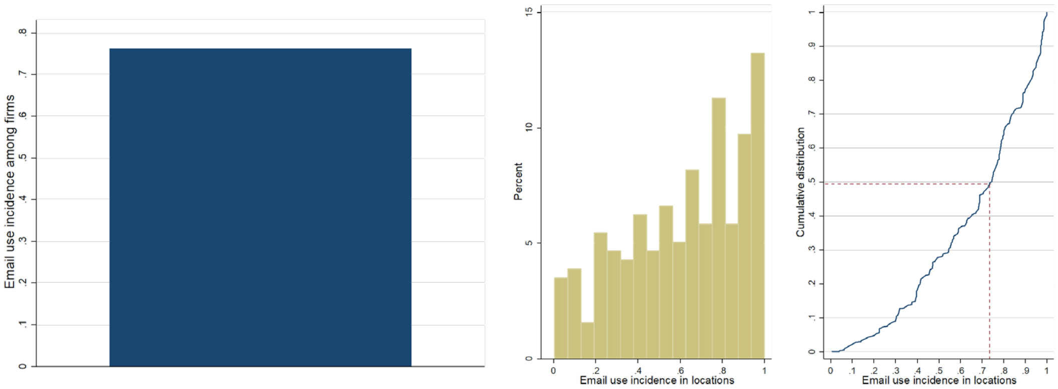
Data: World Bank Enterprise Survey. Sample: 257 location-level observations: 128 locations, 39 countries, based on data from 32,328 surveyed firms.

3.2.2. Variable of interest ($Email_{j,l,t}$)

Our variable of interest is a dummy variable equal to one if the firm i declares using emails to communicate with its clients and suppliers at the time of the survey, drawn from the World Bank Enterprise Surveys. At the location level, this variable represents the incidence of email use among firms, expressed as a share of firms. This variable is used as variable of interest because email is the most basic way to use Internet, associated to simple and more complex usages of the Internet by firms, and therefore the most relevant for an heterogeneous sample of developing and transition countries.

Graph 4 below gives some insights into firms' reported experience of email use. The left-hand side graph shows that around 70% of firms have reported using email during their operations. After data aggregation at the location level, we get a continuous variable which distributions (middle and right-hand side graphs) exhibits higher variability, and stress that half of locations display email-use penetration rates lower than 70%. Therefore, while a large majority of firms use email for their operations, data suggests a much greater spatial heterogeneity, depending on where firms are located.

Graph 4. Global and local incidence of email use among firms



Data: World Bank Enterprise Surveys. Baseline sample: 32,328 firms from 128 locations, 39 countries. Firm-level observations in the left-hand side graph. Data aggregated at the location-level in the middle and right-hand side graphs (257 observations). Data has been aggregated using sample weights.

3.2.3. Control variables ($X_{j,l,t}$)

We control for the following determinants of firm performance (Dollar et al, 2006a, Paunov & Rollo, 2016) averaged by location: the number of full time permanent employees when the firm has started operations, the firm's age, top manager experience (in years), ownership structure (public and foreign ownership), the share of direct and indirect exports in total sales (only indirect exports when the dependent variable is the share of direct exports), the firm's experience of corruption[5], the location distance to international connectivity infrastructure (in km, logarithmic), firms' industry of activity (expressed in share of firms), and the number of power outages the firm's has been subject to.

3.3. Descriptive statistics

Our baseline sample consist of firms located in Eastern-Europe and Central Asia (47% of locations representing 34% of firms), sub-Saharan Africa (23% of locations and 23.5% of firms), Latin-America (19% of locations representing 35% of firms), Middle East and North-Africa (3% of locations and 1.4% of firms), and East-Asia and Pacific (2% of locations and 1.1% of firms).[6] The sample consists of service firms (49.7%) and manufactures (50,3%), mostly operating in retail and wholesale trade (28%), in the food industry (12.3%), in other manufacturing (9.99%) and other services industries (8.6%). Appendix A provides further statistics on the sample composition.

Graphs 5 and 6 provide preliminary insights into the relationship between email use incidence at the location level, and average firm outcomes and characteristics. In fact, firms using email seem

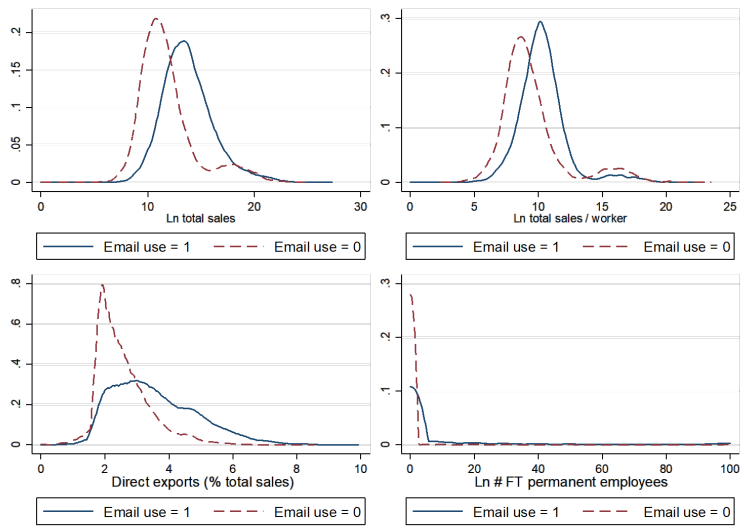
---

[5] A five-point scale ordered categorical variable reflecting the degree to which corruption is an obstacle to current operations.

[6] Thereby covering very few Asian countries (due to few repeated cross-sections in this areas), which are particularly exposed to seismic risk. Therefore, results should not be driven by this area's seaquake exposure.

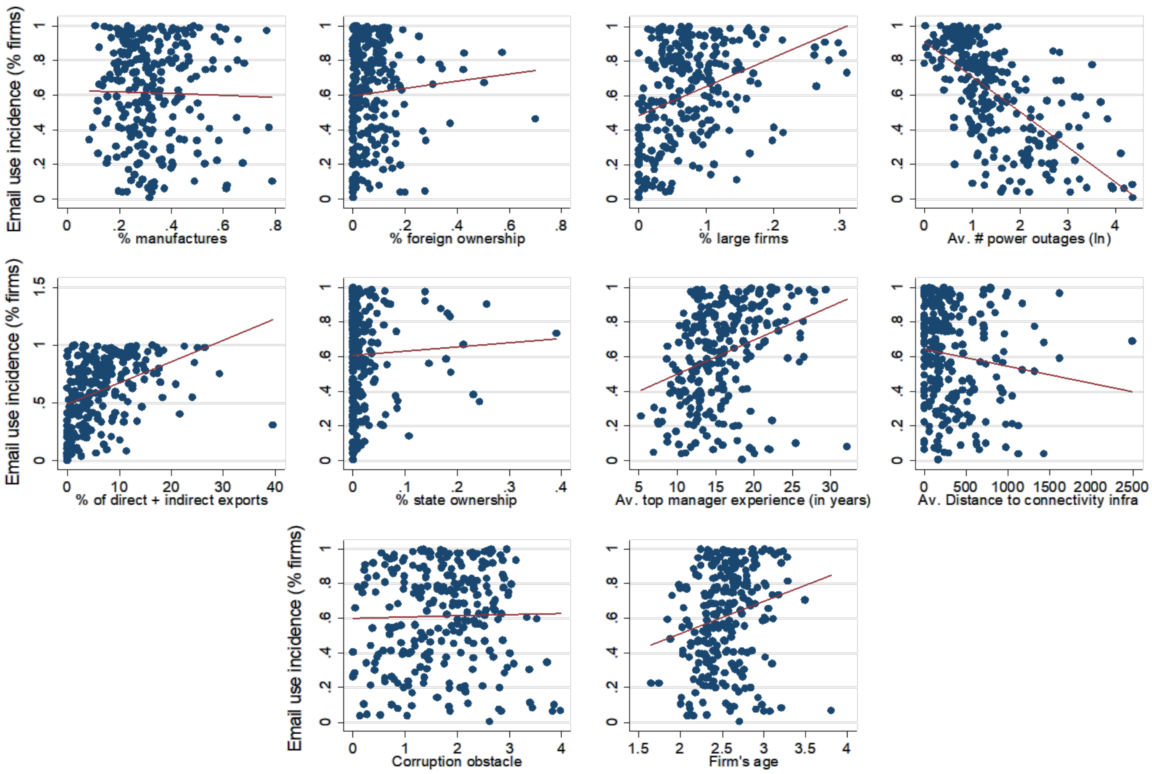
to have higher annual sales, higher labour productivity, and higher exports, on average. Looking at the firm determinants of email use, data aggregated at the location level provides us with interesting insights into simple relationships between firm characteristics and email-use incidence at the location level. In fact, graphical correlations depicted in graph 6 suggest that locations with a higher proportion of large firms, of firm (indirect and direct) exports, of older firms, and with a greater experience of managers, are associated with a higher incidence of email use during operations. By contrast, locations subject to power outages and remote from connectivity infrastructures are associated with a lower incidence of email use. A positive but weaker relationship between state or foreign ownership on the one hand, and email use by firms on the other, is also observable. We see, at first sight, no clear relationship between the share of manufactures or the firm's experience of corruption, and email use incidence at the location level.

Graph 5. Firm outcomes & email use.



Data: World Bank Enterprise Surveys. Sample: 32,328 firm-level observations from 128 locations, 39 countries.

Graph 6. Firm characteristics & email use.



Data: World Bank Enterprise Surveys. Sample: 257 observations (128 locations, 39 countries) from an original sample of 32,328 firms. Data has been aggregated using sample weights.

4- Identification strategy

To identify an eventual impact of email use on the performance of firms, we need to ensure that estimates are not biased by measurement errors, omitted variables, and reverse causality. A first step to address the omitted variable bias is to use the within fixed-effect estimator which allows controlling for location time-invariant unobserved heterogeneity. In fact, time-invariant location's characteristics, such as the proximity to the coast or to the capital, could explain firms' or infrastructure location choice. For similar concerns, we also control for higher-level unobserved characteristics by including year, country-by-year dummies. Second, to address an eventual simultaneity bias between email use and firm performance, we adopt an IV approach, and apply the within fixed-effect two-stage least-square estimator (FE-2SLS) to equation (1b) and the following first-stage equation:

$$Email_{j,l,t} = \beta_0 + \beta_1 Instruments_{l,t} + \beta_2 X_{j,l,t} + \delta_t + (\gamma_j \times \delta_t) + \theta_l + \varepsilon_{j,l,t} \quad (2)$$

where the subscripts have the same meaning as in eq. (1). Control variables and fixed effects are the same as those specified in eq. (1b). The $Instruments_{j,l,t}$ vector corresponds to the instrument set described in the next sub-sections. Standard errors are clustered by country and survey year and are robust to heteroscedasticity. Moreover, because our pseudo-panel is constructed from cohorts of heterogeneous size, we weight the standard errors by the square root of the cohort size, i.e. number of surveyed firms in each location.[7] Appendixes A.2 and A.3 provide information on the sample structure and average cohorts size.

The identification strategy underlying the IV approach described below provides an answer to the following question: what happens to firms when the SMC network's integrity is threatened? For this purpose, it combines an external source with an internal source of digital vulnerability, that is, the SMC network exposure to seismic risk with the location distance to international connectivity infrastructures.

4.1. Seismic risk and digital vulnerability

SMC faults induce large costs for operators owning these cables, amounting to millions of dollars, related to repair and insurance costs; but also large indirect costs related to i) the reporting of repair and insurance costs on Internet tariffs; ii) the rerouting of Internet traffic towards more expensive cable paths, with limited available bandwidth; and iii) the disorganization of global manufacturing chains and ICT-based service provision (Widmer et al., 2010; Clark, 2016; Aceto et al., 2018). Last but not least, these costs are amplified by delays necessary for cable repairs (Palmer-Felgate et al., 2013).

By generating turbidity currents, landslides, and in the most extreme case, tsunamis, maritime seismic events represent a major cause of direct or indirect cable breaks (Soh et al., 2004; Carter et al., 2009; Clark, 2016, Aceto et al., 2018; Carter et al., 2014). Beyond the spectacular large-magnitude seaquakes that repeatedly affected Eastern and Southern Asia's coasts, the redundancy of lower-magnitude seismic activity may break, damage or erode entire sections of the cable network by destabilize the seabed and shaking sediments (Pope et al, 2017). For this same reason, these events also increase the likelihood of future faults.

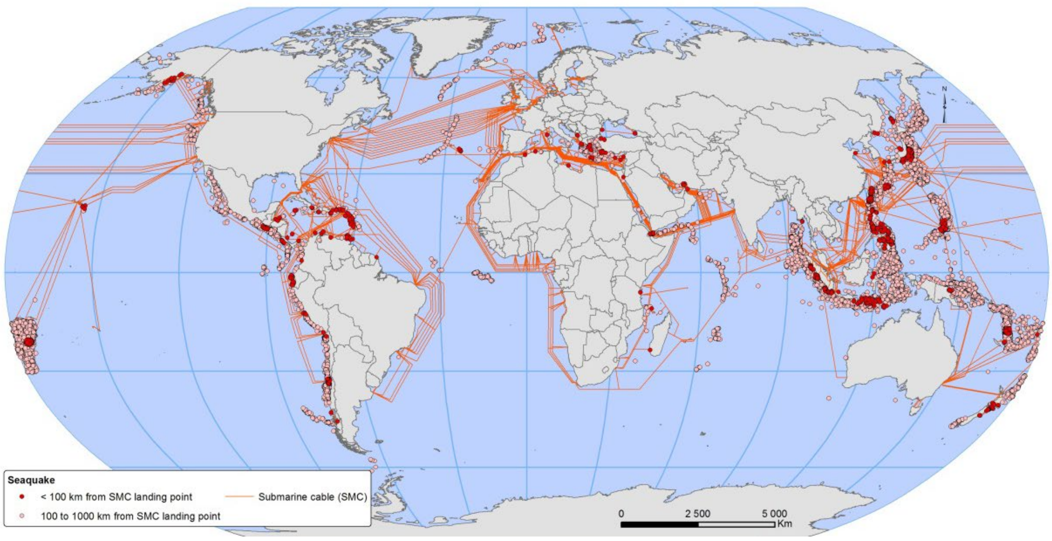
We exploit information on the location, timing, frequency, and magnitude of seaquakes to build, at the country level, a variable of *seismic risk exposure*: the annual frequency of seaquakes likely to affect the functioning of SMCs. In fact, because seismic activity may cause damages to the whole economy and not only to the SMC network, our identification strategy focuses on the occurrence of seaquakes and exclude earthquakes from the analysis. Moreover, to respect identification restrictions, we only consider medium-magnitude seaquakes, whose epicentres are located within a 100-1000km radius from the SMC landing station. First, we only count seaquakes with

[7] Consisting in multiplying left-hand and right-hand variables, and the constant by $\sqrt{n_{l,t}}$.

intensity above 5 on the Richter scale, and drop observations associated with seaquakes whose magnitude is above 6.5: below 5, the seaquake might have little effect on the SMC infrastructure, while above 6.5 the seaquake may damage coastal areas.[8] Second, we do not account for seaquakes located within a 100km radius and those beyond a 1000km radius from SMC landing stations. The former, despite their medium magnitude, could eventually damage the littoral, while the latter may have little effect on SMCs[9]. This instrument calibration is therefore meant to respect under and over-identification restrictions.

To illustrate the world infrastructure's exposure to seismic risk, Figure 1 below maps worldwide seismic events whose epicentre is located within a 100 and 1000km radius from SMC landing stations, from 2005 to 2017. Seismic events considered in the analysis are those depicted in light-pink dots.

Figure 1. International seismic activity - between 5 and 6.5 on Richter scale within a 100 or 1000km radius from SMC landing stations, 2005-2017.



Sources: Authors. Raw data: Telegeography and Northern California Earthquake Data Center of the University of California, Berkeley.

4.2. Digital isolation and digital vulnerability.

Grubesic and Murray (2006) and Grubesic et al (2003) stress that, when telecommunication assets are geographically concentrated, locations distant from telecommunication nodes, i.e. inland, rural and isolated locations, are particularly exposed to telecommunication disruptions,

[8] The lower bound has been chosen according to the work of Soh et al. (2004), who find that cable breaks occurred in the eastern part of Taiwan following earthquakes ranging from 5.0 to 6.0 on the Richter scale. The upper bound is based on interviews with Dr Raphaël Paris, Research Officer in volcanology at CNRS and Laboratoire Magmas et Volcans (LMV) (Observatoire de Physique du Globe de Clermont-Ferrand, Clermont-Auvergne University), who pointed out that the risk of tsunami becomes significant with seismic activity above 6.5 on the Richter scale.
[9] Precise information on SMC maritime paths is not available.

and are slower to recover after telecommunication shutdowns (see also Gorman & Malecki, 2000; Gorman et al., 2004). In support of this finding, Cariolle (2018) shows that, in a panel of sub-Saharan African countries, the average population distance to SMC landing stations increases fixed phone-line fault frequency, and reduces Internet penetration rates.

To get location-level variation and considering the fact that being geographically remote from backbone infrastructures increases the likelihood of telecommunications disruptions, we weight our seaquake variable by the geographical distance of firms' locations to the closest SMC landing station or Internet Exchange Point (IXP).[10] With SMC landing stations, IXPs represent key elements of the international connectivity infrastructure and an important source of network efficiency and Internet bandwidth (Weller & Woodcock, 2013; OECD, 2014; Towela & Tesfaye, 2015). IXPs are indeed physical telecommunication hubs favouring direct interconnections between countries, and enhancing the telecommunication network efficiency and capacity (Weller & Woodcock, 2013; OECD, 2014).

4.3. Instrument set

Our instrument therefore reflects the firm's digital vulnerability to seismic shocks upon the SMC network. Appendix B details the IV's data collection and treatment process. It is made of two components: i) a variable of seismic risk exposure, measured by the annual frequency of seaquakes respecting the constraints specified in subsection 4.1; and ii) a variable of firm's digital isolation, measured by the location distance (in km, logarithmic) to the closest connectivity infrastructures (either SMC landing station or IXPs):

$$Instrument_{j,l,t} = Seaquake\ frequency_{j,t} \times (ln) location\ distance\ to\ connectivity\ infrastructure_{j,l,t} \quad (3)$$

This instrument hence exhibits time and location-level variability. To account for possible nonlinearities in the impact of seaquakes depending on their distance to SMC landing stations, and to conduct instrument validity tests, this instrument is split into two sub-instruments, using two different radiuses of the seaquake frequency variable – 100-500km and 500-1000km from SMC landing stations.

[10] While information on terrestrial backbone infrastructure deployment is not available for all developing and transition countries, the location of SMC landing stations and Internet Exchange Points (IXPs) is available for most of them. The distance variable is therefore the gap between SMC landing station GPS coordinates and the municipality or province's geographic centroid coordinates where the firm has declared its activity. Special cases and data treatment are explained in Appendix B.

5- Results

5.1. Main results

5.1.1. Baseline estimations

Within-FE 2SLS estimates of equations (1b) and (2) are reported in columns (8) and (9) of table 1. They point to a large effect of email use on firms' revenue, as a 10% increase in email use incidence in locations is found to increase by 37-38% average firms' total sales. To ensure that our model is correctly specified and to properly interpret this first evidence, we proceed to various benchmark estimations (columns (1) to (7)). First, we compare firm-level pooled estimations (columns (1) to (4)) to location-level within-FE estimations (columns (5) to (10)), to identify eventual spatial spillovers or network effects of Internet use reflected through data aggregation. Second, we compare pooled OLS (columns (1), (2)) and within FE estimations (columns (5), (6)) to IV estimations (columns (3), (4), (7) to (9)), to assess how our IV approach affects estimated relationships. Last, for each estimator, we compare estimates obtained from the baseline sample resulting from within-FE estimations (columns (4), (6), (8),(9)) to those obtained from an unrestricted sample of firms (columns (2), (4), (6), (8),(9)), to see if results are affected by sample attrition.[11]

Estimations suggest that the large effect of email use on firm revenue results from our IV setup rather than data aggregation or sample attrition. In fact, while data aggregation tends to slightly lowers estimated relationships (columns (2) versus (6), columns (4) versus (8)), sample attrition generates a similar magnitude increase in estimated coefficients (columns (1) versus (2), columns (3) versus (4)). However, pooled IV firm-level estimations (column (4)) and within-FE IV pseudo-panel estimations (columns (7) to (9)) all point to a much larger effect of email use than pooled OLS (columns (1), (2)) and within-FE estimations (columns (5), (6)).

[Table 1]

FE-IV pseudo-panel estimations using the additional outcome variables are reported in Table 2 and stress positive and more-than-proportional effects of email use on firms' labour productivity and, albeit in a less significant way, on employment. According to these estimates, a 10% increase in email-use incidence generates a 22-23% increase in average labour productivity and a 12-14% increase in the average number of full-time permanent employees. However, we do not find any significant effect of email use on firm exports, possibly because of the heterogeneous effects of reduced communication costs on exporting firms (Melitz, 2003; Clarke & Wallsten, 2006; Fink et

[11] Conducting within-FE estimations induce large sample attrition explained by the exclusion of countries that have been once surveyed.

al, 2005). 2SLS pooled estimations conducted on firm-level data are reported in Appendix C and support a strong and positive effect of email use on sales per worker. Additional estimations in Appendix D show that the significance and strength of relationships hold after removing outliers from the sample.

[Table 2]

Whatever the model considered, identification statistics support that our instrument is relevant and valid. First-stage estimates point to a negative and significant effect of digital vulnerability variables on email use. In particular, the effect of firms exposure to seaquakes within a 500-1000km radius from SMC landing stations is quite consistent across IV estimations, which corroborates the findings of Carter et al (2009, 2014) according to which seismic events located in deeper sea water are more detrimental to SMC integrity than those closer to the coast.

In table 3, we aggregate data on manufacturing and service firms separately to study whether previous relationships depend on firms' sector of activity. Results show that the effect of email use on total sales, labour productivity, and employment is driven by the service sector. They also show that an increased use of email by service firms leads to a strong 10%-significant reduction in their direct exports. However, first-stage estimates stress that manufactures are not digitally vulnerable. This additional evidence is supported by previous findings of researches highlighting the heterogeneous effects of ICTs on firm outcomes. First, it has been shown that service firms are generally more affected by ICTs diffusion than manufactures (Stiroh, 2002; Aboal & Tacsir, 2017). Second, the literature on trade liberalization has put in evidence the heterogeneous effects of trade and communication costs reduction on firm outcomes, and shows that an increased exposure to trade through reduced trade barriers may provoke the less productive exporters exit from international markets (Melitz, 2003; Bustos, 2011).

[Table 3]

Despite inconclusive evidence on manufactures, we hereafter exploit detailed information on manufactures' employment occupation types to study an eventual heterogeneous effect of email use incidence among firms, including service firms, on manufactures' workforce composition.

5.1.2. Email use and manufactures' employment outcomes

The literature on the ICT-employment nexus has stressed that technological change induced by the introduction of ICTs is biased in favour of an educated and skilled workforce (Michaels et al, 2014; Akerman et al, 2015; Acemoglu & Restrepo, 2018; Hjort & Poulsen, 2019). To test this "skill-biased technological change" (SBTC) hypothesis, we exploit information reported by manufactures on their workforce's type of occupation: the (logarithmic) number of non-production and production workers, and the (logarithmic) number of skilled and unskilled production workers.[12] The analysis focuses on employment outcomes in manufactures, keeping right-hand side explanatory variables averaged over the whole sample of firms, i.e. both service and manufacture firms. We are therefore interested in studying an eventual indirect or spillover effect of email use incidence on manufactures' workforce composition, which could be explained by the digitization of local ecosystems.

Results, reported in table 4, support the existence of employment spillovers from ICT diffusion in locations, since a greater local incidence of email use is found to increase number of non-production workers (column (1)) and unskilled production workers (columns (4)) in manufactures. Therefore, beyond the employment dividends of email use evidenced in table 2, we find heterogeneous patterns of employment in manufactures: while the SBTC hypothesis seems confirmed by the positive effect of email use on manufacture jobs associated with non-production tasks (e.g. management, auditing, accounting, etc.), this hypothesis is questioned by the evidence of a stronger effect on the number of unskilled production workers.

[Table 4]

To conclude this subsection, one key implication of this first series of estimations is that spatial heterogeneities underlying our IV approach are at the core of the Internet-production nexus. Digital vulnerability appears to be a good predictor of email use at the local level, with a large indirect impact on revenue, labour productivity and employment, especially for service firms.

5.3. Robustness checks

In this subsection, we proceed to a range of robustness checks that are aimed at challenging the validity of our identification strategy. To keep this article to a reasonable length, we perform these checks with our four main outcome variables.[13]

5.3.1. Robustness 1: Does increased firm vulnerability to seismic shocks around SMC actually increase the telecommunication constraint?

In a first robustness check, we make sure that the negative effects of firm's exposure to seismic shocks upon the SMC network on email use, evidenced in first stage estimations, actually result in a stronger telecommunication constraint. Although the WBES does not provide objective information on the quality and affordability of firms' access to Internet, the survey includes an

---

[12] Such information for service firms was not available. Summary statistics on these dependent variables are provided in Appendix A.

[13] Robustness tests applied to manufactures' workforce variables can be sent upon request.

ordered categorical variable[14] reflecting the extent to which access to telecommunications represents an obstacle to the firm's operations. We therefore replace the instrumented variable of firm email use by this proxy of the telecommunication constraint reported by the firm. Results are reported in table 5. First-stage estimates confirm our interpretation of first-stage estimates, as seismic shocks around the SMC network are found to increase the average telecommunication constraint in locations remote from international connectivity infrastructures. Moreover, results strongly support that an increased telecommunication obstacle results into lower sales, lower sales per worker and reduces employment, in line with previously evidenced relationships. Last, estimates also provide additional evidence on the negative effect of an improved access to telecommunication on firms' exports.

[Table 5]

5.3.2. Robustness 2: Are results affected by the presence of landlocked countries in the sample?

In a second check, we exclude landlocked countries from the sample because they cannot directly host SMCs, but can be indirectly connected to them via the terrestrial cable network. Although the presence of fixed effects control for the consequences of this geographic feature on telecommunication outcomes, landlocked countries are particularly dependent on neighbouring coastal countries hosting SMCs, so that the non-treatment might act in a heterogeneous way for these countries because of unobserved information on the terrestrial infrastructure network deployment and cross-border connectivity. However, results, reported in table 6, confirm baseline estimations. First-stage estimates are indeed almost unaffected by this sample restriction, suggesting that the hypothesis of non-treatment of landlocked countries is realistic. Moreover, second-stage ones lie within the same range as those reported in tables 1 and 2, except for labour productivity, suggesting that keeping landlocked countries in the sample has only a small (downward) effect on estimated relationships.

[Table 6]

5.3.1. Robustness 3: Is there a location selection bias?

The following next robustness checks are meant to address location selection bias that can be induced by i) firms' decision to locate their activity next to connectivity infrastructures, and ii) the deployment of telecommunications infrastructures where the most performing firms are located. So, in a third check, we exclude large and foreign firms from the sample, which are known to be

[14] Firms were asked: "Is Telecommunications No Obstacle, a Minor Obstacle, a Moderate Obstacle, a Major Obstacle, or a Very Severe Obstacle to the current operations of this establishment?". The resulting ordered categorical variable becomes a continuous variable after aggregation at the location level.

more geographically mobile (Baldwin & Okubo, 2006; Dollar et al., 2006ab), and aggregate the data on a sub-sample of domestic small and medium enterprises (SMEs). Moreover, large and foreign firms could be the main drivers of sales and productivity improvements, as evidenced by Van Biesebroeck (2005) in the African context. Last, we are also interested in the effect of email use on the performance of SMEs because their performances are expected to be particularly responsive to ICT use (Dholakia & Kshetri, 2004; Clarke et al, 2011; Paunov & Rollo, 2015). Results are presented in Table 7 and support a positive, significant, but slightly softer effects of email use on domestic SMEs' sales, sales per worker, and full-time employment. Interestingly, results support a very large, 1%-significant and positive effect of email use on the share of direct exports in SMEs' total sales, a 10% increase in email use incidence yielding a doubling of the average share of direct exports in locations.

[Table 7]

We also make sur that results are not driven by the presence in our sample of capital cities – where telecommunication infrastructures and firms are expected to be the most performing – and provinces – where telecommunication infrastructures are often missing, firms are sparsely distributed and often less performing. We therefore re-run previous estimations confining the sample to intermediary cities, that is, excluding capital cities and provinces from the sample, and report estimates in table 8. Estimates remain robust to this restriction and consistent with baseline estimations, the difference lying in the effect of email use on firm exports, which are found to rise in the same proportion as with the sample of SMEs. Therefore, the use of email, in addition to be beneficial to labour productivity, employment and total revenue, is found to stimulate exports of smaller firms in smaller cities.

[Table 8]

Finally, we exclude both capital cities and locations situated outside a 100km radius from international connectivity infrastructures (Table 9). Because the resulting estimates and associated identification statistics could be biased by an eventual weak-instrument bias, we perform the "continuously-updated"-GMM estimator of Hansen et al (1996), which is a GMM version of the Limited Information Maximum Likelihood estimator (Davidson and MacKinnon, 1993) robust to non-spherical errors. We also proceed to various IV calibrations, aimed at minimizing problems of weak and under-identification. Based on identification statistics and coefficient strength, and in light of first-stage statistics from previous estimations, we focus on estimates resulting from the single seaquake frequency instrument calibrated over the [500km; 1000km] radius (columns, (2), (5), and (8)). Doing this yields significant and much stronger relationships, to be considered with caution but which make sense, since this restricted sample is made of locations more distant from core connectivity infrastructures and therefore more exposed to telecommunication disruptions (Gorman & Malecki, 2000; Gorman et al., 2004).

[Table 9]

5.3.4. Check 4: Could over-identification restrictions by violated by seaquakes close to the coast?

Our last robustness check is aimed at ensuring that our instrument set is not invalid by considering seaquakes with epicentres close the coast. Such seaquakes could indirectly affect firm outcomes by provoking physical or human casualties and thereby make our model over-identified. While this concern is lowered by the parametrization of the IV setup, which leads to take into account medium size seaquakes that are located beyond 100km from the SMC landing stations, we impose an alternative and more conservative constraint consisting in dropping observations associated with at least one seaquake 100 km-close to the coast. Because this calibration induces strong observation attrition for the instrument set (more than 60% of instrument observations consider at least one seaquake within 100km from the coast), we apply this sample restriction using a single instrument equal to the annual frequency of seaquakes located within a [0-1000km] radius from SMC landing station. Results are reported in table 10. Under-identification statistic point to a somewhat weaker but still relevant instrument, which is probably explained by the attrition of observations associated to the treatment. Nevertheless, estimates confirm the relevance of our IV setup and the positive effect of email use on firms' sales, and sales per worker.

[Table 10]

6- Discussion and concluding remarks

In this paper, we provide evidence of a large and positive impact of ICT use and diffusion on firms' performance at the location level. A higher prevalence of email use at the location level is found to boost firm sales, labour productivity, and employment. According to baseline FE-IV estimations, a 10% increase in local email use incidence raises by 37-38% the firms' average annual sales, by 22-23% their average sales per worker, by 12-14% the average number of full-time permanent workers. Evidence on the effect of email use on firm exports is more mixed, and somewhat weaker, since the direction and significance of the coefficient depends on firms' size, location size, and sector of activity. While the use of email is found to boost exports of domestic SME's and firms located in intermediary cities, it has a detrimental effect on exports in the service sector, probably because of an increased exposure to international competition. These findings are found to be driven by the service sector, but a further analysis reveals the existence of positive spillovers of email use on manufactures' non-production workers and unskilled production workers. This last evidence also nuances the conclusions of recent researches supporting the hypothesis of a skilled-biased technological change in developing economies, since the adoption and use of ICT are found to benefit to the unskilled production workforce. All in all, this paper is, to our knowledge, among the few ones providing evidence on the benefits of email use for firms from large sample of developing and transition countries.

Through the IV approach followed in this paper, we also pointed out new digital vulnerabilities countries may be subject to. In fact, the arrival of SMCs has boosted the Internet economy as a whole, but has also increased economies' exposure to fibre cable faults and Internet shutdowns, especially for peripheral populations remote from key international connectivity infrastructures. Our results indeed stress the importance of the SMC network exposure to seismic risk for the development of the telecommunication sector and for the performance of firms, but the underlying mechanisms could be extended to other previously sources of cable faults, such as maritime activities, piracy, or other natural hazards.

A last key finding is that within-country spatial heterogeneities are found to be a central element of the ICT-growth nexus. In fact, we show that the location distance to IXPs or SMC landing stations is a critical transmission factor of the effect of seismic shocks on Internet diffusion among firms. In other words, firms that are remote from connectivity infrastructures face geographical handicaps that increase their exposure to telecommunication disruptions, which may deprive them of the potential benefits of the digital revolution.

Tables

Table 1. Baseline estimations (1/2)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Var dep: (ln) Total sales	**Pooled firm-level framework**				**Within FE location-level framework** [a]				
	Pooled OLS		**Pooled 2SLS**		**Within-FE**		**Within FE-2SLS**		
2nd stage estimates		Baseline sample		Baseline sample		Baseline sample		Baseline sample	
Email use	**0.994***** **(0.087)**	**1.284***** **(0.096)**	**4.440***** **(1.689)**	**5.490***** **(2.030)**	**0.782*** **(0.415)**	**0.664** **(0.625)**	**9.763**** **(4.665)**	**3.795**** **(1.310)**	**3.683***** **(1.294)**
State-owned	-0.031 (0.623)	0.422 (0.714)	0.175 (0.835)	0.613 (0.670)	-0.458 (1.628)	-0.263 (2.041)		1.974 (1.389)	1.942 (1.370)
Foreign	0.464*** (0.135)	0.844*** (0.109)	0.507*** (0.198)	0.549*** (0.215)	1.441 (1.039)	0.783 (1.563)		-0.585 (1.633)	-0.206 (1.630)
Age	0.368*** (0.067)	0.415*** (0.072)	0.247*** (0.0864)	0.309*** (0.111)	-0.213 (0.368)	-0.371 (0.449)		-0.869* (0.456)	-0.856* (0.454)
# power outages	-0.038 (0.067)	-0.026 (0.040)	-0.017 (0.094)	-0.021 (0.067)	0.032 (0.160)	0.101 (0.184)		0.196 (0.186)	0.191 (0.186)
% of exports	0.008*** (0.001)	0.008*** (0.002)	0.003 (0.004)	-0.0007 (0.005)	0.001 (0.008)	0.009 (0.014)		-0.009 (0.011)	-0.009 (0.011)
Initial # of FT employee	0.649*** (0.047)	0.477*** (0.055)	0.492*** (0.098)	0.336*** (0.084)	0.831*** (0.284)	0.830** (0.326)		0.724** (0.331)	0.731** (0.327)
Experience	0.006 (0.004)	-0.006 (0.005)	0.007 (0.005)	-0.006 (0.007)	0.043** (0.016)	0.041* (0.023)		0.016 (0.023)	0.016 (0.022)
Obstacle: corruption	-0.027 (0.042)	-0.022 (0.034)	-0.115** (0.051)	-0.061 (0.043)	0.100 (0.084)	0.049 (0.117)		0.024 (0.094)	0.024 (0.093)
Distance to telecom infra	-0.033 (0.027)	0.209* (0.126)	-0.245*** (0.054)	-0.247*** (0.071)	-0.132* (0.033)	-0.130 (0.081)		-0.198** (0.081)	-0.194** (0.079)
1st stage est. (eq. (2))									
Seaquake freq 100-500km x Ln dist infra			-0.0009** (0.0004)	-0.0008** (0.0003)			0.000 (0.0007)	0.0007 (0.001)	
Seaquake freq 500-1000km x Ln dist			-0.006***	-0.005***			-0.004***	-0.006***	-0.006***

infra			(0.002)	(0.001)			(0.001)	(0.001)	(0.001)
Controls			Yes	Yes			No	Yes	Yes
Fixed effects	Location, Year, Sector, Country×Year				Year; Country×Year; Location				
Observation units:	Firms				Locations				
# countries/locations/obs. treated			33/143/13,759	7/23/2,781			12/30/43	10/27/37	
Hansen test (p. value)			0.34	0.46			0.47	0.46	-
Weak-ident. SW F-test			6.65***	5.97***			4.80**	17.81***	29.43***
Under-ident. SW Chi2.			13.53***	12.17***			12.71***	55.14***	45.31***
R^2	0.59	0.63			0.86	0.87			
N	39,398	13,331	34,184	13,311	305	257	313	257	257
# locations	517	128	431	128	152	128	156	128	128
# countries	112	39	103	39	46	39	45	39	39
Original sample size (# firms)	-	-	-	-	40,304	32,328	43,539	32,328	32,328

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity. Standard errors are clustered by location and survey year in columns (1)-(2), and by country and survey year in columns (5) to (7). a: in columns (5) to (9), sector dummies become control variables reflecting the sector share in total firms by location.

Table 2. Baseline estimations (2/2)

		(4)		(6)	(8)	
Var dep:	**(ln) Sales per worker**		**(ln) # FT employees**		**% direct exports**	
Email use	2.197**	2.303**	1.399**	1.247*	-1.338	0.028
	(0.961)	(0.980)	(0.682)	(0.653)	(6.571)	(5.370)
Controls	Yes		Yes		Yes[a]	
First stage estimates						
Seaquake freq 100-500km x Ln dist infra	0.0007		0.0007		0.0005	
	(0.001)		(0.001)		(0.001)	
Seaquake freq 500-1000km x Ln dist infra	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed effects	Year; Country×Year; Sector; Location					
Observation unit	Location					
Hansen test (p. value)	0.47	-	0.17	-	0.29	-
Weak-identification SW F-test	17.81***	29.43***	14.14***	28.19***	14.31 ***	26.83***
Under-identification SW Chi-sq.	55.14***	45.31***	43.21***	42.86***	43.72***	40.79***
N	257	257	283	283	283	283
# locations	128	128	141	141	141	141
# countries	39	39	43	43	43	43

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country and survey year. **a:** controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

Table 3. Service firms vs. manufactures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Var dep:	**Sales**		**Sales/worker**		**# of FT employees**		**Direct exports[a]**	
	Manuf	Services	Manuf	Services	Manuf	Services	Manuf	Services
Email use	21.20	5.351***	10.23	4.011***	4.496	1.158**	-5.926	-9.903*
	(32.73)	(1.274)	(17.69)	(0.924)	(4.565)	(0.512)	(50.30)	(5.571)
First stage estimates:								
Seaquake freq 100-500km x Ln dist infra	0.0001	0.0016	0.0001	0.0016	-0.0007	0.0014	-0.0005	0.0013
	(0.001)	(0.0013)	(0.001)	(0.0013)	(0.001)	(0.0011)	(0.001)	(0.0011)
Seaquake freq 500-1000km x Ln dist infra	- 0.0007	-0.008***	-0.0007	-0.008***	-0.0008	-0.009***	-0.0009	-0.009***
	(0.001)	(0.0016)	(0.001)	(0.0016)	(0.001)	(0.0015)	(0.002)	(0.0016)
Controls	Yes							
Fixed effects	Year; Country×year; Location							
Observation unit	Location							
Hansen test (p. value)	0.33	0.28	0.21	0.32	0.56	0.26	0.26	0.28
Under-ident.. SW F-test	0.15	15.45***	0.15	9.84***	0.74	15.95***	0.74	14.80***
Weak indent. SW Chi-sq	0.47	47.33***	0.47	26.73***	0.98	48.26***	0.98	44.78***
N	249	257	249	257	275	283	275	283
# Locations	124	128	124	128	137	141	137	141
# Countries	39	39	39	39	43	43	43	43

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country and survey year. **a:** controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

Table 4. Manufactures' workforce composition.

	(1)	(2)	(3)	(4)
	All workers		Prod. Workers	
Var dep:	Non prod.	Prod.	skilled	unskilled
Email use	1.274***	0.506	-0.780	1.822**
	(0.585)	(0.458)	(0.586)	(0.823)
1st stage est.				
Seaquake freq 100-500km x Ln dist infra	0.0002	0.0002	0.0002	0.0002
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Seaquake freq 500-1000km x Ln dist infra	-0.005***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Fixed effects	Year; Country-year; Sector; Location			
Controls	Yes			
Hansen test p-value	0.86	0.33	0.12	0.95
Under-ident.. SW F-test	14.95***	14.95***	14.95***	14.95***
Weak indent. SW Chi-sq	46.51***	46.51***	46.51***	46.51***
N	277	277	277	277
# locations	138	138	138	138
# Countries	41	41	41	41

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors in parentheses are robust to heteroscedasticity and clustered by country and survey year.

Table 5. Telecommunication obstacle and firm performance.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Var dep:	**Sales**		**Sales/worker**		**# of FT employees**		**Direct exports[a]**	
Telecom obstacle	-1.030***	-1.226**	-0.743***	-0.766***	-0.237*	-0.425**	5.313*	5.773*
	(0.354)	(0.483)	(0.306)	(0.360)	(0.126)	(0.483)	(3.035)	(3.290)
First stage estimates:								
Seaquake freq 100-500km x Ln dist infra	0.004		0.004		0.004		0.002	
	(0.006)		(0.006)		(0.006)		(0.005)	
Seaquake freq 500-1000km x Ln dist infra	0.019***	0.018***	0.019***	0.018***	0.017***	0.017***	0.031**	0.033**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.0.13)	(0.013)
Controls	Yes							
Fixed effects	Year; Country×year; Location							
Hansen test (p. value)	0.22	-	0.81	-	0.15		0.26	
Under-ident.. SW F-test	7.64**	11.62**	7.64**	11.62**	6.55**	7.65**	3.55*	6.79**
Weak indent. SW Chi-sq	23.60***	17.85***	23.60***	17.85***	19.96***	11.60***	10.83***	10.30***
N	255	255	255	255	281	281	281	281
# Locations	127	127	127	127	140	140	140	140
# Countries	38	38	38	38	42	42	42	42

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country and survey year. **a:** controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

Table 6. IV estimations, excluding landlocked countries.

	(1)	(2)	(3)	(4)
Var dep:	**Total sales**	**Sales / worker**	**# of FT employees**	**Direct exports[a]**
Email use	**3.976*****	**3.195*****	**1.628*****	**-7.959**
	(0.816)	**(0.956)**	**(0.425)**	**(0.255)**
1st stage est.				
Seaquake freq 100-500km x Ln dist infra	-0.002	-0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Seaquake freq 500-1000km x Ln dist infra	-0.007***	-0.007***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Fixed effects	Year; Country×year; Location			
Controls	Yes			
Hansen test p-value	0.29	0.26	0.20	0.32
Under-ident.. SW F-test	15.18***	15.18***	53.76***	40.37***
Weak indent. SW Chi-sq	60.90***	60.90***	203.47***	152.78***
N	108	108	124	124
# locations	54	54	62	62
# Countries	18	18	20	20

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors in parentheses are robust to heteroscedasticity and clustered by country. a: controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

Table 7. IV estimations, domestic SMEs.

	(1)	(2)	(3)	(4)
Var dep:	**Total sales**	**Sales / worker**	**# of FT employees**	**Direct exports[a]**
Email use	3.230***	1.967***	0.679*	11.29***
	(0.582)	(0.532)	(0.375)	(4.165)
1st stage est.				
Seaquake freq 100-500km x Ln dist infra	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Seaquake freq 500-1000km x Ln dist infra	-0.008***	-0.008***	-0.008***	-0.008***
	(0.0013)	(0.0016)	(0.0014)	(0.0015)
Fixed effects	Year; Country×year; Location			
Controls	Yes			
Hansen test p-value	0.65	0.33	0.10	0.14
Under-ident.. SW F-test	20.31***	20.31***	17.37**	17.89**
Weak indent. SW Chi-sq	62.22***	62.22***	52.57***	54.13***
N	257	257	283	283
# locations	128	128	141	141
# Countries	39	39	43	43

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country and survey year. **a:** controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

Table 8. IV estimations, firms located in intermediary cities.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Var dep:	**Sales**		**Sales/worker**		**# of FT employees**		**Direct exports**[a]	
Email use	4.062***	4.166***	4.911***	4.438**	0.874**	1.055**	9.901*	9.349
	(1.444)	(1.578)	(1.728)	(1.929)	(0.405)	(0.521)	(5.892)	(7.449)
First stage estimates:								
Seaquake freq 100-500km x Ln dist infra	-0.0016 (0.0006)		-0.0016 (0.0006)		-0.002 (0.001)		-0.002 (0.001)	
Seaquake freq 500-1000km x Ln dist infra	-0.013**	-0.011**	-0.013**	-0.011**	-0.016***	-0.014***	-0.017***	-0.015***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Controls	Yes							
Fixed effects	Year; Country×year; Location							
Hansen test (p. value)	0.85	-	0.66		0.50		0.86	
Under-ident.. SW F-test	4.14*	4.82*	4.14*	4.82*	7.99**	8.90**	6.53***	8.38**
Weak indent. SW Chi-sq	15.77***	9.07***	15.77***	9.07***	30.37***	16.74***	22.38***	14.20***
N	141	141	141	141	149	149	149	149
# Locations	70	70	70	70	74	74	74	74
# Countries	31	31	31	31	34	34	34	34

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country and survey year. **a:** controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

Table 9. IV estimations, distance to connectivity infrastructures > 100km (excluding capital cities).

	(1)	(2)	(3)	(4)	(5)	(6)
Var dep:	**Total sales**[a]			**Sales / worker** [a]		
Email use	10.14***	9.906***	5.706***	7.956***	7.648***	4.387***
	(3.340)	(4.462)	(1.985)	(2.835)	(3.774)	(1.602)
1st stage est.						
Seaquake freq 100-500km x Ln dist infra	-0.034			-0.034		
	(0.044)			(0.044)		
Seaquake freq 500-1000km x Ln dist infra	-0.186	-0.260**		-0.186	-0.260**	
	(0.133)	(0.100)		(0.133)	(0.100)	
Seaquake freq 100-1000km x Ln dist infra			-0.057**			-0.057**
			(0.026)			(0.026)
Fixed effects	Year; Country×year; Location					
Controls	Yes					
Hansen test p-value	0.31	-	-	0.31	-	-
Under-ident.. SW F-test	4.90*	6.81**	4.97*	4.90*	6.81**	4.97*
Weak indent. SW Chi-2	19.23***	13.20***	9.65***	19.23***	13.20***	9.65***
N	159	159	159	159	159	159
# locations	79	79	79	79	79	79
# Countries	36	36	36	36	36	36

	(7)	(8)	(9)	(10)	(11)	(12)
Var dep:	**# FT employees**[a]			**Direct exports**[ab]		
Email use	2.150***	2.832**	1.543***	-72.43**	-31.26	-67.55**
	(0.674)	(1.329)	(0.460)	(36.96)	(26.61)	(34.82)
1st stage est.						
Seaquake freq 100-500km x Ln dist infra	-0.041			-0.022		
	(0.031)			(0.029)		
Seaquake freq 500-1000km x Ln dist infra	-0.125	-0.181*		-0.111	-0.142	
	(0.092)	(0.097)		(0.109)	(0.106)	
Seaquake freq 100-1000km x Ln dist infra			-0.051**			-0.032
			(0.024)			(0.025)
Fixed effects	Year; Country×year; Location					
Controls	Yes					

Hansen test p-value	0.36	-	-	0.37	-	-
Under-ident.. SW F-test	3.05†	3.51†	4.52*	1.28	1.80	1.72
Weak indent. SW Chi-2	11.82***	6.74**	8.68***	4.96*	3.46*	3.31*
N	169	169	169	169	169	169
# locations	84	84	84	84	84	84
# Countries	39	39	39	39	39	39

Note: † significant at 15%, * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country and survey year. a: "continuously-updated" GMM estimator. **b:** controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

Table 10. IV estimations[a], excluding seaquakes 50km to the coast.

	(1)	(2)	(3)	(4)
Var dep:	**Total sales**	**Sales / worker**	**# of FT employees**	**Direct exports[b]**
Email use	7.645**	6.075**	1.885***	2.115
	(3.054)	(2.593)	(0.660)	(8.914)
1st stage est.				
Seaquake freq 1000km > 100km from the coast × Ln dist infra	-0.014*	-0.014*	-0.011†	-0.011*
	(0.007)	(0.007)	(0.007)	(0.007)
Fixed effects	Year; Country×year; Location			
Controls	Yes			
Under-ident.. SW F-test	3.84*	3.84*	2.70†	3.03†
Weak indent. SW Chi-sq	6.02**	6.02**	4.21**	4.73**
N	211	211	221	221
# locations	105	105	110	110
# Countries	30	30	32	32

Note: † significant at 15%, * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country and survey year. a: "continuously-updated" GMM estimator. **b:** controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

References

Aboal, D., & Tacsir, E. (2017). Innovation and productivity in services and manufacturing: the role of ICT. *Industrial and Corporate Change*, 27(2): 221-241.

Acemoglu, D., & Restrepo, P. (2018). The race between machine and man: Implications of technology for growth, factor shares and employment, *American Economic Review*, 108(6): 1488-1542.

Aceto, G., Botta, A., Marchetta, P., Persico, V., & Pescapé, A. (2018). A comprehensive survey on internet outages. *Journal of Network and Computer Applications*, In Press.

Aker; J. (2017) "Using digital Technology for Public service provision in developing countries: Potential and pitfalls", in Eds. Sanjeev Gupta, Michael Keen, Alpa Shah, & Geneviève Verdier, *Digital Revolutions in Public Finance*, IMF, 363p, Washington DC.

Akerman, A. , Gaarder, I., & Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *Quarterly Journal of Economics*, 130: 1781-1824.

Baldwin, R. E., & Okubo, T. (2005). Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. *Journal of Economic Geography*, 6(3): 323-346.

Björkegren, D. (2018) The Adoption of Network Goods: Evidence from the Spread of Mobile Phones in Rwanda, *The Review of Economic Studies*, published online.

Brun, J. F., Carrère, C., Guillaumont, P., & De Melo, J. (2005). Has distance died? Evidence from a panel gravity model. *The World Bank Economic Review*, *19*(1): 99-120.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms. *American Economic Review*, 101(1), 304-40.

Buys, P., Dasgupta, S., Thomas, T. S., & Wheeler, D. (2009). Determinants of a digital divide in Sub-Saharan Africa: A spatial econometric analysis of cell phone coverage. *World Development*, 37(9): 1494-1505.

Cariolle, J. (2018). Telecommunication Submarine-Cable Deployment and the Digital Divide in Sub-Saharan Africa. CESifo Working Paper No. 7415, December.

Carter, L., Burnett, D., Drew, S., Marle, G., Hagadorn, L., Bartlett-McNeil, D., & Irvine, N. (2009). *Submarine cables and the oceans – Connecting the world* (UNEP-WCMC Biodiversity Series No. 31). Cambridge: UNEP-WCMC.

Carter, L., R. Gavey, P.J. Talling, & J.T. Liu. 2014. Insights into submarine geohazards from breaks in subsea telecommunication cables. *Oceanography* 27(2): 58–67.

Choi, C. (2010). The effect of the Internet on service trade. *Economics Letters*, *109*(2): 102–104.

Choi, C., & Yi, M. H. (2009). The effect of the Internet on economic growth: Evidence from cross-country panel data. *Economics Letters*, *105*(1): 39–41.

Clark, B. (2016). Undersea cables and the future of submarine competition. *Bulletin of the Atomic Scientists*, *72*(4): 234–237.

Clarke, G. (2008). Has the Internet increased exports for firms from low and middle-income countries? *Information Economics and Policy*, *20*(1): 16–37.

Clarke, G., & Wallsten, S. J. (2006). Has the Internet increased trade? Developed and developing country evidence. *Economic Inquiry*, *44*(3): 465–484.

Clarke, G., Qiang, C. Z., & Xu, L. C. (2015). The Internet as a general purpose technology: Firm-level evidence from around the world. *Economics Letters*, *135*(C): 24–27.

Czernich, N., Falck, O., Kretschner, T., & Woessmann, L. (2011). Broadband infrastructure and economic growth. *The Economic Journal*, *121*(552): 205–532.

Davidson, R. & MacKinnon, J. *Estimation and Inference in Econometrics*. New York: Oxford University Press, 1993.

Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of econometrics*, 30(1-2): 109-126.

Dholakia, R. R., & Kshetri, N. (2004). Factors impacting the adoption of the Internet among SMEs. *Small Business Economics*, *23*(4): 311-322.

Dicken, P., & Malmberg, A. (2001). Firms in territories: a relational perspective. *Economic Geography*, 77(4): 345-363.

Dollar, D., Hallward-Driemeier, M., & Mengistae, T. (2006a). Investment Climate and firm Performance. *Economic Development and Cultural Change*, *54(1*): 1-31.

Dollar, D., Hallward-Driemeier, M., & Mengistae, T. (2006b). Investment climate and international integration. *World Development*, *34*(9): 1498–1516.

Fink, C., Mattoo, A. & Neagu, I.C. (2005) Assessing the impact of communication costs on international trade, *Journal of International Economics* , 67(2): 428-445.

Freund, C., & Weinhold, D. (2002). The Internet and international trade in services. *American Economic Review*, *92*(2): 236-240.

Freund, C., & Weinhold, D. (2004). The Effect of the Internet on international trade. *Journal of International Economics*, *62*(1): 171–189.

Giuri, P., Torrisi, S., & Zinovyeva, N. (2008). ICT, skills, and organizational change: evidence from Italian manufacturing firms. *Industrial and Corporate change*, *17*(1): 29-64.

Goldfarb, A., & Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1): 3-43.

Gorman, S. P., & Malecki, E. J. (2000). The networks of the Internet: an analysis of provider networks in the USA. *Telecommunications Policy*, 24(2): 113-134.

Gorman, S. P., Schintler, L., Kulkarni, R., & Stough, R. (2004). The revenge of distance: Vulnerability analysis of critical information infrastructure. *Journal of Contingencies and Crisis Managemen*t, 12(2): 48–63.

Grubesic, T. H. & Murray, A. T. (2006). Vital nodes, interconnected infrastructures, and the geographies of network survivability. *Annals of the Association of American Geographers*, 96(1): 64-83.

Grubesic, T. H., O'Kelly, M. E., & Murray, A. T. (2003). A geographic perspective on commercial Internet survivability. *Telematics and informatics*, 20(1): 51-69.

Hansen, L.P., Heaton, J., & Yaron, A. 1996. Finite Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics*, 14(3): 262-280.

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

Litan, R. E., & Rivlin, A. M. (2001). Projecting the economic impact of the Internet. *American Economic Review*, *91*(2): 313-317.

Malecki, E. J. (2002). The economic geography of the Internet's infrastructure. *Economic Geography*, 78(4): 399–424.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6): 1695-1725.

Michaels, G., Natraj, A., & Van Reenen, J. (2014). Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years. *The Review of Economics and Statistics*, 96(1): 60-77.

Niebel, T. (2018). ICT and economic growth–Comparing developing, emerging and developed countries. *World Development*, *104*:197-211.

Palmer-Felgate, A., Irvine, N., Ratcliffe, S., & Bah, S. S. (2013, April 22–25). *Marine maintenance in the zones – A global comparison of repair commencement times*. Paper presented at Suboptic Conference: From Ocean to Cloud*,* Paris*.*

Paunov, C., & Rollo, V. (2015). Overcoming Obstacles: The Internet's Contribution to Firm Development. *The World Bank Economic Review*, *29*(suppl_1): S192-S204.

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

Pope, E. L., Talling, P. J., & Carter, L. (2017). Which earthquakes trigger damaging submarine mass movements: Insights from a global record of submarine cable breaks?. *Marine Geology*, 384: 131-146.

Röller, L-H, & Waverman, L. (2001). Telecommunications infrastructure and economic development: A simultaneous approach. *American Economic Review*, *91*(4): 909–923.

Schumann, R., & Kende, M. (2013). Lifting barriers to Internet development in Africa: Suggestions for improving connectivity (Report for the Internet Society). London: Analysys Mason.

Stiroh, K. J. (2002). Are ICT spillovers driving the new economy?. *Review of Income and Wealth*, *48*(1), 33-57.

Soh, W., Machiyama, H., Shirasaki, Y., & Kasahara, J. (2004). Deep-sea floor instability as a cause of deepwater cable fault, off eastern part of Taiwan. *Frontier Research of Earth Evolution*, *2*: 1–8.

Towela, N-J., & Tesfaye, B. (2015, May 22). *Internet development and Internet governance in Africa*. Internet Society.

TeleGeography. (2016). *Submarine cable map*. https://www.submarinecablemap.com/

Van Biesebroeck, J. (2005). Firm Size Matters: Growth and Productivity Growth in African Manufacturing. *Economic Development and Cultural Change*, 53(3): 545-583.

Wallsten, S. (2005). Regulation and internet use in developing countries. *Economic Development and Cultural Change*, 53(2), 501-523.

Weller, D., & Woodcock, B. (2013). Internet traffic exchange: Market developments and policy challenges, OECD Digital Economy Papers, No. 207, Paris: OECD Publishing.

Widmer, G., Wolff, M. (Producers), & Dotzer, U. (Director). (2010). *Les câbles Internet sous-marins: L'autoroute mondiale de l'information* [motion picture]. Arte reportage, Fact+ film production.

Yincan, Y., Xinmin, J., Guofu, P., & Wei, J. Submarine Optical Cable Engineering. Elsevier Academic Press, July 2018, 374 p.

Acknowledgements and funding

This work was supported by the Agence Nationale de la Recherche of the French government through the program 'Investissements d'avenir' (ANR-10-LABX-14-01), through the IDGM+ initiative led by Ferdi (Fondation pour les études et recherches sur le développement international).

We would like to thank Dr. Raphaël Paris (Laboratoire Magma Volcan – UCA, CNRS) for the insightful discussions we had on seismic risk, as well as Jaime de Melo and Florian Leon for their particularly helpful remarks. We also thank Simone Bertoli, Michaël Goujon, Michele Imbruno, Luc Jacolin, Mathilde Maurel, Jaime de Melo, and participants at the Ferdi-Banque de France research seminar, at the GRETHA research seminar, the 2017 DIAL, the 2017 AFSE, the 2018 DEGIT XXIII, and the RES 2019 conferences, for their valuable comments. We are also grateful to Martine Bouchut (CERDI-UCA) for helping us in the programming of some variable calculation.

Appendix A. Summary statistics

A.1. Summary statistics (location-level)

	# locations	# firms	Mean	Std. Dev.	Min	Max
Dep. var:						
Ln Total sales (USD)	257	27,139	12.76	1.82	8.89	20.50
Ln Sales per worker (USD)	257	27,024	10.00	1.71	6.16	17.75
Ln # fulltime employees	257	32,067	2.92	0.32	2.02	3.88
Ln # production workers	254	15,199	2.76	0.39	1.72	4.41
Ln # non-production workers	254	15,158	1.83	0.43	0.34	3.26
Ln # skilled production workers	254	14,957	2.27	0.43	0.42	4.34
Ln # unskilled production workers	254	14,767	1.39	0.48	0.00	2.93
Direct export (% sales)	257	32,001	4.23	3.94	0.00	25.00
Explanatory. var:						
Dummy email use	257	32,244	0.67	0.25	0.01	1.00
Telecommunication obstacle	256	23,167	1.09	0.59	0.00	3.28
Dummy state-owned	257	32,328	0.01	0.03	0.00	0.39
Dummy foreign-owned	257	32,328	0.08	0.08	0.00	0.70
Ln Firm age (in years)	257	32,319	2.58	0.32	1.65	3.81
Ln # power outages	257	17,140	1.30	0.82	0.00	4.37
% of exports (direct + indirect)	257	31,976	6.54	5.31	0.00	39.48
Ln # initial FT employees	257	29,328	2.13	0.31	1.47	3.64
Top-manager experience (in years)	257	31,717	16.85	5.11	5.31	32.13
Corruption obstacle	257	31,177	1.85	0.77	0.00	3.99
Ln firm distance (km) to infrastructures (A)	257	32,328	3.63	2.65	0.00	7.82
Seaquake annual freq., [100; 1000] km rad. (B)	257	-	6.10	13.35	0.00	44.00

Note: Averages are weighted by sample weights, while standard deviations are adjusted cohort's size, that is, std dev = $\{n/[W(n-1)]\} \Sigma w_i (x_i - x^{bar})^2$ where W is the sum of the raw weights w_i (cohort size), xi the data, x^{bar} the population mean, and n is the population size.

A.2. Baseline firm-level sample composition

39 countries, 32,328 firms

iso	2006	2007	2008	2009	2010	2011	2013	2014	Total
AFG	0	0	436	0	0	0	0	410	846
AGO	425	0	0	0	360	0	0	0	785
ARG	1,063	0	0	0	982	0	0	0	2,045
ARM	0	0	0	354	0	0	343	0	697
AZE	0	0	0	380	0	0	390	0	770
BIH	0	0	0	361	0	0	311	0	672
BLR	0	0	273	0	0	0	360	0	633
BOL	613	0	0	0	362	0	0	0	975
BWA	342	0	0	0	268	0	0	0	610
COL	1,000	0	0	0	942	0	0	0	1,942
GEO	0	0	210	0	0	0	204	0	414
GHA	0	345	0	0	0	0	419	0	764
GTM	522	0	0	0	590	0	0	0	1,112
HND	436	0	0	0	360	0	0	0	796
HUN	0	0	0	161	0	0	208	0	369
KAZ	0	0	0	544	0	0	600	0	1,144
KGZ	0	0	0	188	0	0	146	0	334
LTU	0	0	0	276	0	0	270	0	546
MDA	0	0	0	363	0	0	360	0	723
MDG	0	0	0	14	0	0	30	0	44
MKD	0	0	0	366	0	0	360	0	726
MLI	0	490	0	0	360	0	0	0	850
MNG	0	0	0	362	0	0	360	0	722
NIC	478	0	0	0	336	0	0	0	814
NPL	0	0	0	368	0	0	482	0	850
PAN	540	0	0	0	258	0	0	0	798
PRY	613	0	0	0	361	0	0	0	974
ROM	0	0	0	348	0	0	354	0	702
RWA	212	0	0	0	0	241	0	0	453
SLV	401	0	0	0	223	0	0	0	624
SRB	0	0	0	242	0	0	209	0	451
TJK	0	0	261	0	0	0	239	0	500
TZA	355	0	0	0	0	0	601	0	956
UGA	563	0	0	0	0	0	688	0	1,251
UKR	0	0	851	0	0	0	1,002	0	1,853
URY	621	0	0	0	607	0	0	0	1,228
YEM	0	0	0	0	301	0	151	0	452
ZAR	340	0	0	0	359	0	0	0	699
ZMB	0	484	0	0	0	0	720	0	1,204
Total	**8,524**	**1,319**	**2,031**	**4,327**	**6,669**	**241**	**8,807**	**410**	**32,328**

A.3. Baseline location-level sample composition

39 countries, 128 locations, 257 observations

iso	2006	2007	2008	2009	2010	2011	2013	2014	Total	Average cohort size
AFG	0	0	5	0	0	0	0	5	10	85
AGO	3	0	0	0	3	0	0	0	6	131
ARG	4	0	0	0	4	0	0	0	8	256
ARM	0	0	0	3	0	0	3	0	6	116
AZE	0	0	0	4	0	0	4	0	8	96
BIH	0	0	0	4	0	0	4	0	8	84
BLR	0	0	6	0	0	0	6	0	12	53
BOL	3	0	0	0	3	0	0	0	6	163
BWA	2	0	0	0	2	0	0	0	4	153
COL	4	0	0	0	4	0	0	0	8	243
GEO	0	0	3	0	0	0	3	0	6	69
GHA	0	2	0	0	0	0	2	0	4	191
GTM	2	0	0	0	2	0	0	0	4	278
HND	3	0	0	0	3	0	0	0	6	133
HUN	0	0	0	2	0	0	2	0	4	92
KAZ	0	0	0	5	0	0	5	0	10	114
KGZ	0	0	0	3	0	0	2	0	5	67
LTU	0	0	0	4	0	0	4	0	8	68
MDA	0	0	0	4	0	0	4	0	8	90
MDG	0	0	0	1	0	0	1	0	2	22
MKD	0	0	0	4	0	0	4	0	8	91
MLI	0	4	0	0	4	0	0	0	8	106
MNG	0	0	0	5	0	0	5	0	10	72
NIC	2	0	0	0	2	0	0	0	4	204
NPL	0	0	0	3	0	0	3	0	6	142
PAN	1	0	0	0	1	0	0	0	2	399
PRY	2	0	0	0	2	0	0	0	4	244
ROM	0	0	0	6	0	0	6	0	12	59
RWA	2	0	0	0	0	2	0	0	4	113
SLV	1	0	0	0	1	0	0	0	2	312
SRB	0	0	0	2	0	0	2	0	4	113
TJK	0	0	3	0	0	0	3	0	6	83
TZA	3	0	0	0	0	0	3	0	6	159
UGA	5	0	0	0	0	0	5	0	10	125
UKR	0	0	5	0	0	0	5	0	10	185
URY	2	0	0	0	2	0	0	0	4	307
YEM	0	0	0	0	4	0	4	0	8	57
ZAR	4	0	0	0	4	0	0	0	8	87
ZMB	0	4	0	0	0	0	4	0	8	151
Total	**43**	**10**	**22**	**50**	**41**	**2**	**84**	**5**	**257**	126

A.5. Baseline sample composition, by region

	Firm-level data		Location-level data	
Region	**Obs.**	**Percent**	**Obs.**	**Percent**
AFR	7,616	23.56	60	23.35
EAP	360	1.11	5	1.95
ECA	10,896	33.7	120	46.69
LAC	11,308	34.98	48	18.68
MNA	452	1.4	8	3.11
SAR	1,696	5.25	16	6.23
Total	**32,328**	**100**	**257**	**100**

A.6. Baseline sample composition, by sector

Sector	# firms	% firm
Textiles	1,140	3.53
Leather	303	0.94
Garments	2,403	7.43
Food	3,989	12.3
Metals and machinery	1,949	6.03
Electronics	208	0.64
Chemicals and pharmaceuticals	1,247	3.86
Wood and furniture	447	1.38
Non-metallic and plastic materials	1,288	3.98
Auto and auto components	26	0.08
Other manufacturing	3,354	10.4
Retail and wholesale trade	9,059	28
Hotels and restaurants	1,503	4.65
Other services	3,231	9.99
Other: Construction, Transportation, et	2,154	6.66
Other	27	0.08
Total	**32,328**	**100**

Appendix B. Instrument set: data collection and treatment

B.1. Infrastructure deployment variables

Raw data on SMCs are drawn from Telegeography:

- All cables with date of commissioning
- All the landing stations of cables and their GPS coordinates

Raw data on Internet Exchange Points are drawn from Telegeography and completed by the *Packet Clearing House* and *Peering DB* databases:

- All IXPs with their status (active/inactive/project)
- their year of activation
- their GPS coordinates

After a conversion into polygons (disk with 5 km diameter) to avoid topological inaccuracies, the SMC landing stations and IXPs from each country are identified and located, and counted.

B.2. firm distance to infrastructure nodes

A distance raster map is defined from all coordinate points, which gives the distance of each firm to the nearest Internet Exchange Point.

Statistical inputs: SMC landing station coordinates, IXPs' coordinates, firm's city location coordinates.

Firms' location centroids (WBES), SMC landing station coordinates (Telegeography), and IXPs coordinates (Telegeography, peering DB) give points for which the distances to SMCs are calculated using the previously calculated distance raster. In some countries, firms may have been pooled and interviewed by province rather than municipality of location. In this case, we take the province's centroid as firm's location coordinates. The firm distance to infrastructure variable is the minimum distance for the firm to reach the closest infrastructure node: either a landing station or an IXP.

B.3. Exposure to seaquake-induced cable faults

The Northern California Earthquake Data Center of the University of Berkeley provides a global database of earthquakes. For each country, we get for each year the number, the location, and the average magnitude of epicenters of occurring seaquakes, and are therefore able to compute the annual frequency of seaquakes near the stations according a 1000 km radius.

To ensure that we do not take into account seaquakes that could induce tsunamis, which would hence violate restriction identification conditions, we drop observations when seaquake magnitude exceeds 6.5 on the Richter scale. To ensure that we do take into account seaquakes that are strong enough to induce cable faults, we only count seaquakes which magnitudes exceeds 5 on the Richter scale. All in all, seaquakes considered for the empirical analysis are those occurring within a 1000 km radius from SMC landing station.

Appendix C. 2SLS pooled firm-level estimations

Extended sample versus baseline sample

	(1)	(2)	(3)	(4)	(5)	(6)
Var dep:	**Sales/worker**		**# of FT employees**		**Direct exports[a]**	
		Baseline sample		Baseline sample		Baseline sample
Email use	3.048***	3.666***	0.996	1.709	-34.10	-36.46
	(1.212)	(1.356)	(1.182)	(1.238)	(24.93)	(32.88)
First stage estimates:						
Seaquake freq 100-500km x Ln dist infra	-0.001***	-0.0008***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Seaquake freq 500-1000km x Ln dist infra	-0.007***	-0.006***	-0.005**	-0.004**	-0.004**	-0.003*
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Controls	Yes					
Fixed effects	Location, Year, Sector, Country×Year					
Hansen test (p. value)	0.32	0.40	0.74	0.67	0.11	0.07
Under-ident.. SW F-test	7.92***	7.30***	5.69***	6.20***	5.18***	6.04***
Weak indent. SW Chi-sq	16.10***	14.88***	11.56***	12.60***	10.52***	12.28***
N	34,131	13,276	38,805	15,064	38,882	15,110
# Locations	430	128	451	141	451	141
# Countries	103	39	109	43	109	43

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors, in parentheses, are robust to heteroscedasticity and clustered by country and survey year. **a:** controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.

Appendix D. Removing outlier influence

Baseline estimations, excluding outliers detected by the Grubbs test.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Var dep:	**Sales**		**Sales/worker**		**# of FT employees**		**Direct exports[a]**	
Email use	4.467***	4.273***	3.114***	3.101***	1.400**	1.248*	-0.028	0.218
	(0. 855)	(0.483)	(0.691)	(0.693)	(0.682)	(0.652)	(5.178)	(4.982)
First stage estimates:								
Seaquake freq 100-500km x Ln dist infra	-0.0006		-0.0006		0.0005		-0.0001	
	(0.001)		(0.001)		(0.006)		(0.001)	
Seaquake freq 500-1000km x Ln dist infra	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	0.031**	-0.006***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0.13)	(0.0011)
Controls	Yes							
Fixed effects	Year; Country×year; Location							

Hansen test (p. value)	0.28	-	0.92	-	0.17		0.27	
Under-ident.. SW F-test	21.35***	28.92***	21.35***	28.92***	14.14***	26.19***	17.95***	26.34***
Weak indent. SW Chi-sq	66.72***	44.94***	66.72***	44.94***	43.21***	42.86***	55.09***	40.23***
N	243	243	243	243	283	283	279	279
# Locations	121	121	121	121	141	141	139	139
# Countries	37	37	37	37	42	42	43	43

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country. a: controls include the share of indirect exports, instead of the share of direct and indirect exports used in other regressions.