

Digital vulnerability and the local performance of firms in developing and transition countries

Joël Cariolle, Research Officer, Fondation for International Development Studies and Researches (FerdDi), joel.cariolle@ferdi.fr

With Maëlan Le Goff (Banque de France) and Olivier Santoni (FerdDi-CERDI)

DEGIT XXIII, September 6-7, 2018

Moscow, Russia

- **Motivation**
- **Literature Review**
- **Our contribution**
- **Infrastructure deployment and digital vulnerability**
- **Empirical Strategy**
- **Results**
- **Robustness checks**
- **Concluding Remarks**

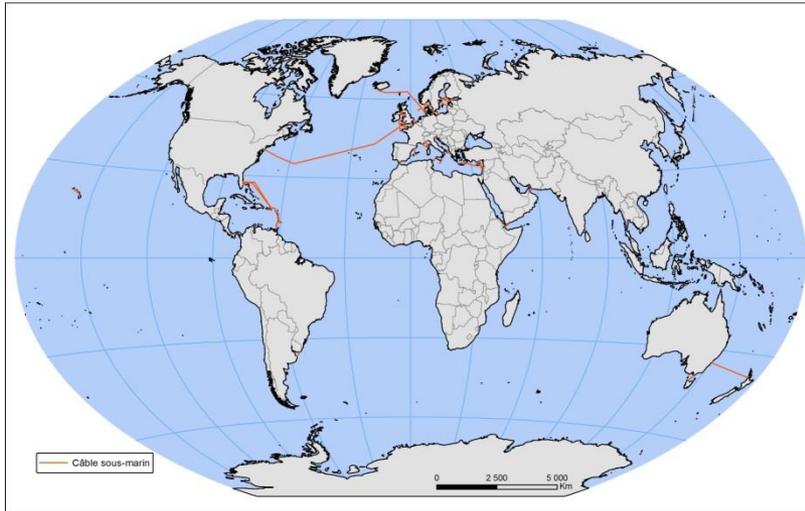
Over the last decades, international connectivity underwent a dramatic improvement, promoted by the laying of **more than 320 fiber-optic submarine cables (SMC)** over 1990-2015...

- irrigating a USD 20.4 trillion industry,
- connecting 3 billion Internet users worldwide (Internet Society 2015).
- In 2013, “20 households with average broadband usage generate as much traffic as the entire Internet carried in 1995” (OECD, 2013)
- In 2016, more than 99% of the world telecommunications passes through SMCs.

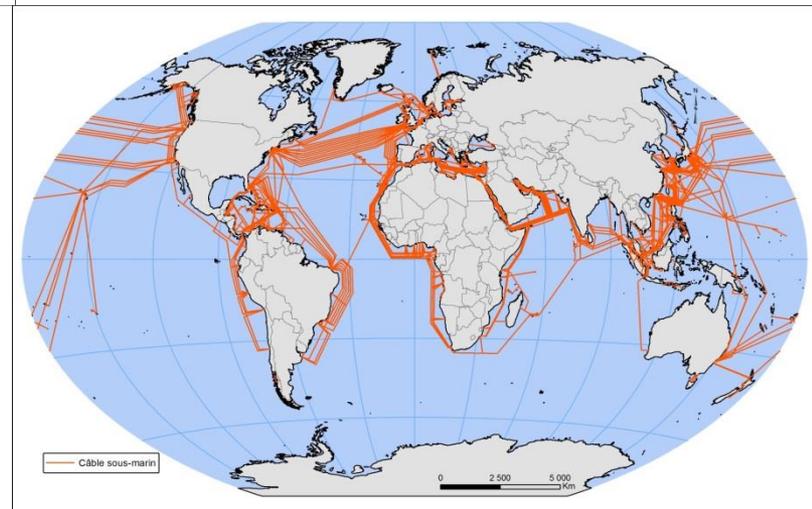
The submarine telecom infrastructures are now one of the mainstays of the global economy

Motivation

1995



2015



Gordon (2012): the impact of the third industrial revolution (penetration of computers, internet and mobile phones) on the American labor productivity, is very limited compared to what happened in the wake of the first and 2nd industrial revolution.

However, a growing literature evidences a significant effect of a better access to ITCs on economic activity in developing countries.

The digital revolution seems to be rather a developing countries' revolution.

What are the consequences of the increasing role of ICTs for the development process?

A positive effect of ICTs is found on:

- **Economic growth** (Roller and Waverman, 2001; Choi et Yi, 2009; Andrianaivo and Kpodar, 2011), **employment** (Hjort and Poulsen, 2016) and **labor productivity** (Clarke et al., 2015; Paunov and Rollo, 2015; Menon, 2011; Cette et al, 2016)
- **Trade** (Freund and Weinhold, 2004; Clarke and Wallsten, 2006), and **exports** (Clarke, 2008; Hjort and Poulsen, 2016)
- **Attractiveness** (Choi, 2003)
- **Agricultural development** (Jansen, 2007; Eygir et al. , 2011; Subervie, 2011; Aker and Fafchamps, 2013)
- **Governance** (Andersen et al., 2011; Asongu and Nwachukwu, 2016), **political stability** (Stodden et Meier, 2009)

Our contribution

- This paper is among the first studies of the csq of broadband infrastructure deployment on firm outcomes, with Hjort and Poulsen (2017)
- It differs from this last study by **enlarging the scope of the sample to other non-african developing and transition areas** (60-70 countries), and by controlling in our baseline model for a **wider set of broadband infrastructure variables**.
- Our contribution also lies in our **identification strategy**. To control for a potential endogeneity bias, we instrument firm access to Internet by **structural sources of digital vulnerability, related to infrastructure deployment**:
 - The **exposure to telecommunication submarine cable (SMC) outages**
 - The **digital isolation**.

We find that an easier access to internet during firms' operations, induced by lower digital isolation and exposure to SMC outages, boosts their sales and their productivity, and to a lesser extent, temporary employment.

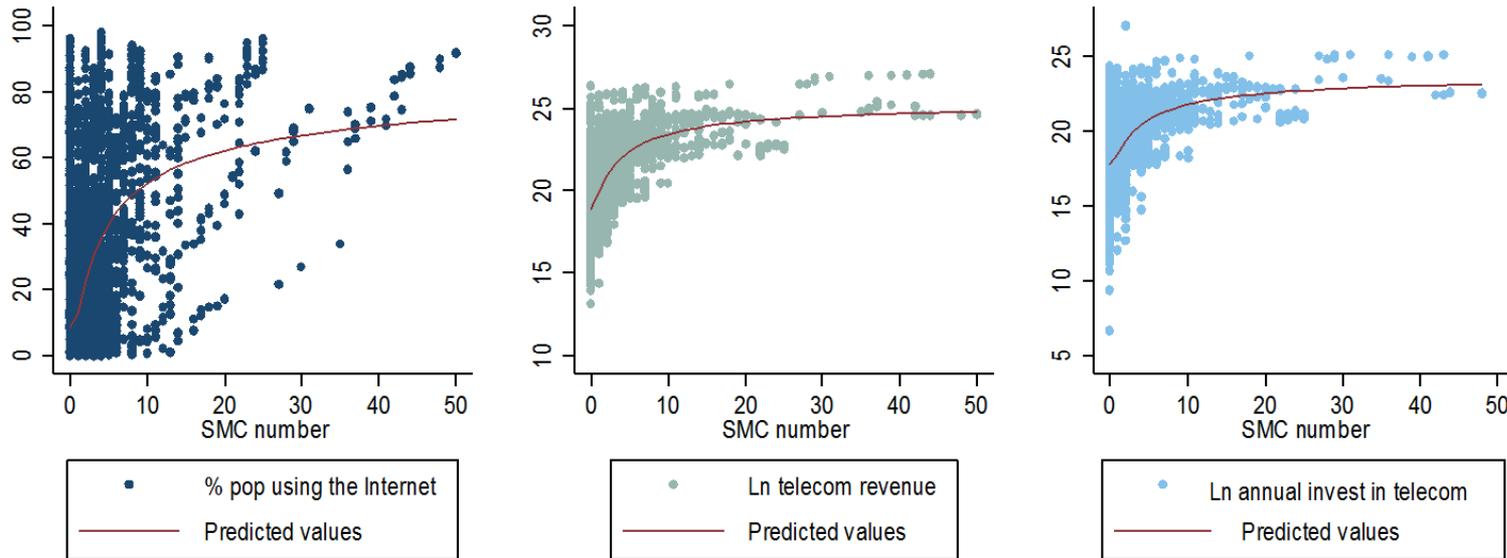
Infrastructure deployment and digital vulnerability

Infrastructure deployment and digital vulnerability

- SMC network is the first component of the global broadband Internet access value chain
- When SMCs are missing, international connectivity is ensured by:
 - **Satellites** (expensive, limited bandwidth, slow)
 - **Buying bandwidth** to a SMC-connected neighboring countries (expensive, limited bandwidth)
 - **Internet exchange points**, confining internet communications at the local or regional level.
- The number of SMCs plugging countries to the global Internet is expected to boost the Internet economy by:
 - Widening the bandwidth, and fastening the internet speed;
 - Shortening the distance between internet users, and lowering the cost of internet access;
 - Increasing the competition between cable operators and ISPs;
 - Creating scale economies, and triggering terrestrial infrastructures investments;
 - Increasing the redundancy, and therefore the resilience of communication networks to cable faults and internet disruptions;

Infrastructure deployment and digital vulnerability

SMC deployment and the internet economy



Notes: world evidence, 1990-2014. Raw data from ITU (2016) and Telegeography (2016).

Infrastructure deployment and digital vulnerability

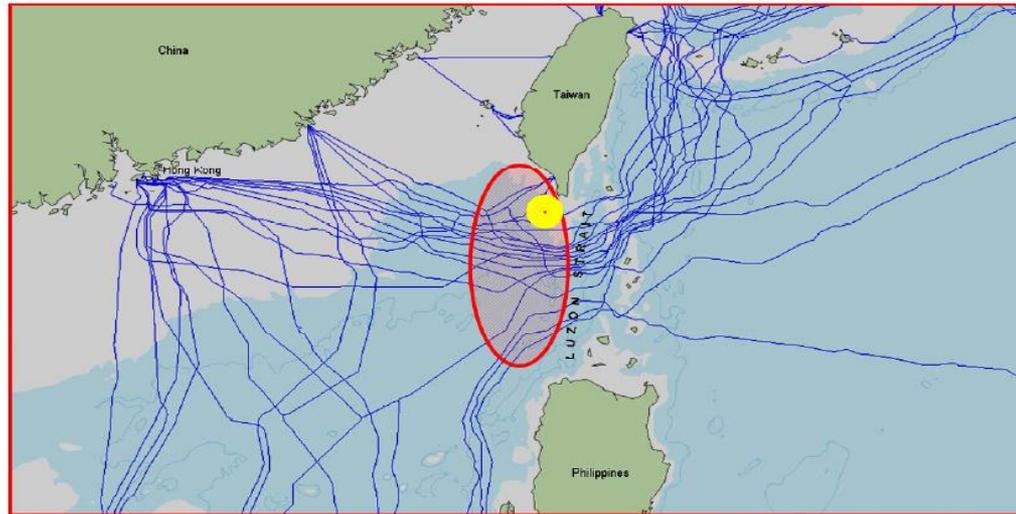
- As the growth of the digital economy strongly relies on SMCs, new vulnerabilities arose from their deployment.
- We define **digital vulnerability** as the risk of a country and its population to see their access to the global Internet disrupted by failures in their telecommunications network.
- These failures may result from the **under-capacity** or **gradual obsolescence** of the telecommunications infrastructure network, as well as its **exposure to recurring external shocks and internal failures** (breakdowns of server, SMC-outs, closing data-centers of Internet exchange points), power outages, and cyber-attacks.

The paper focuses primarily on digital vulnerabilities related to the deployment of SMCs

- Once a country is plugged to the world Internet by SMC, countries are exposed to **two structural interrelated sources of digital vulnerability**, independent from policy:
 - i) **their exposure to SMC outages**, resulting from human activities (shipping, piracy, sabotage) or natural hazards (earthquakes, typhoons, floods, etc.) (Palmer-Felgate et al, 2013; UNEP, 2009; Clark, 2016).
 - ii) **the digital isolation** resulting from the geographical distance to SMC landing stations, and making isolated countries and remote populations even more exposed to telecommunication failures (Grubestic and Murray, 2006; Grubestic et al. 2003; OECD, 2013; ITU, 2015).

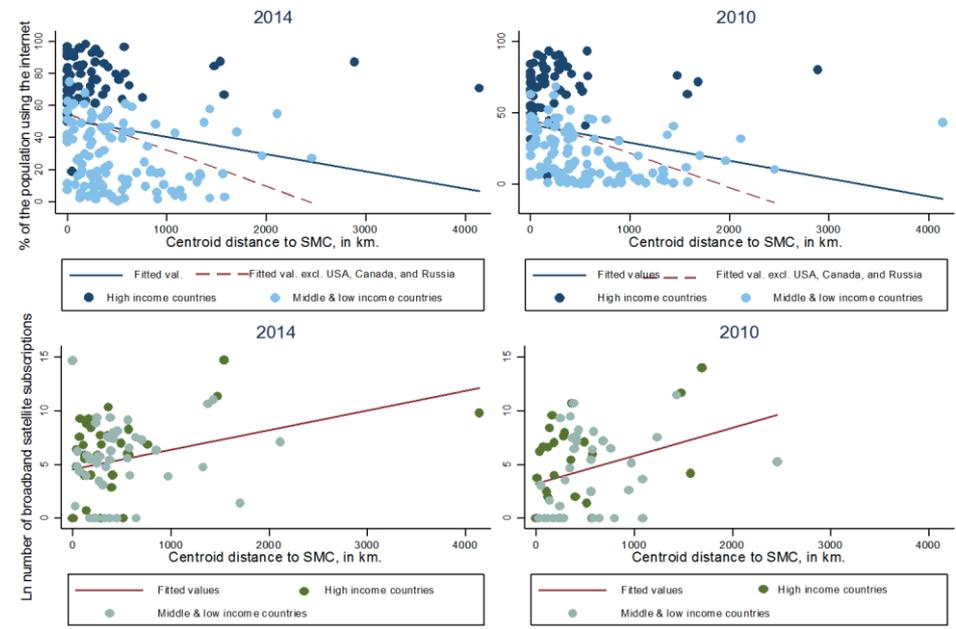
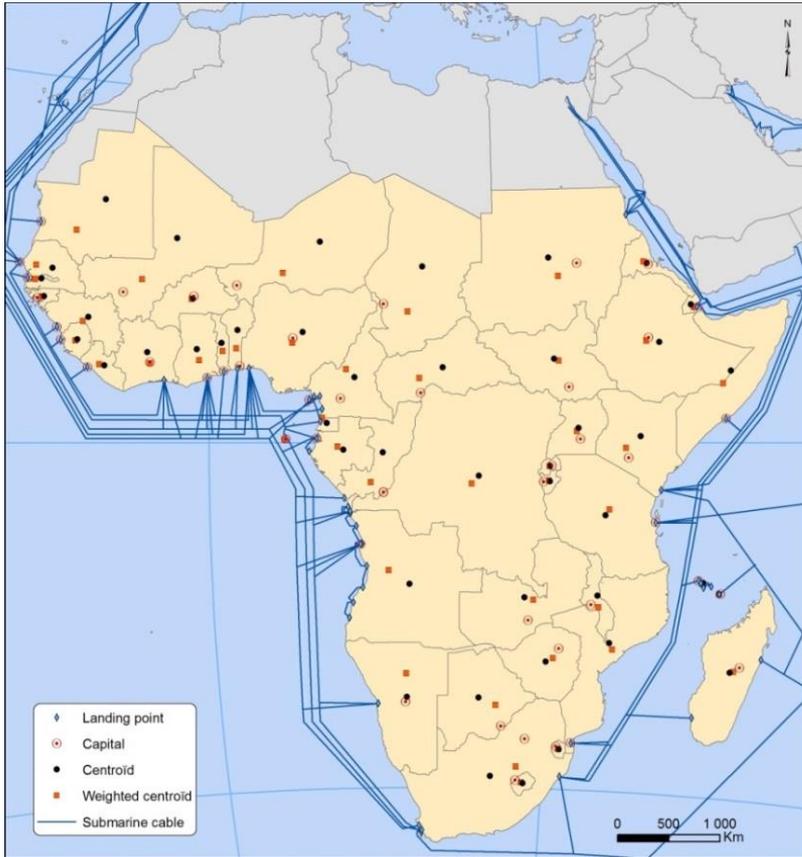
SMC deployment and digital vulnerability

“On 26th December, 2006, a powerful earthquake shook the seabed off southern Taiwan [...] followed by one of the largest disruptions of modern telecommunications systems [...] disabling vital connections between SE Asia and the rest of the world” (ICPC press release, 2007)



Earthquake epicentre (yellow dot) and the general area of the disruption of the submarine cables (blue lines). Courtesy Global Marine Systems Ltd

Digital isolation



Notes: High-income (dark-coloured) vs. middle- and low-income (light-coloured) countries. Top graphs: 192 countries (2014) and 188 countries (2010). Bottom graphs: 82 countries (2014), 66 countries (2010). The dashed fitted correlation line has been drawn to exclude outliers USA, Canada, and Russia. Data on Internet access proxies drawn from the International Telecommunication Union.

Source: Cariolle (2018)

The exposure to SMC outages

- The recent laying of SMCs in developing countries has also increased their vulnerability to SMC outages (especially countries relying on few cables).
- **471 cable repairs worldwide over 2008-2012** (Palmer-Felgate et al, 2013).
- There is a **direct cost of SMC outages**, related to **cable repairs**, depending on the number of cable outages/breaks, their location, the delay in repair-permit acquisition, or other operational issues.
- ...and also **indirect costs** for the whole economy, amplified by repair delays, are related to (Widmer et al, 2010; Clark, 2016):
 - The **reporting of repair/insurance costs on internet tariffs** and its consequences on internet penetration;
 - The **rerouting of internet traffic** towards more expensive cable paths and its consequences on internet capacity and tariffs;
 - The **telecommunication network instability** and the **disorganization of global manufacturing** chains and internet-related service provision.

Digital isolation

The laying of SMCs has also increased the **spatial digital divide** between:

- **coastal or urban populations (the core)** close to SMC landing stations and key other backbone infrastructures, benefitting from a faster and more stable telecommunication network,
- and **isolated inland or rural populations (the periphery)** with low infrastructure coverage and more exposed to telecommunication network failures.

Two mechanisms (Malecki, 2002; Grubestic, O’Kelly, and Murray, 2003; Gorman et al, 2004; Grubestic and Murray, 2006; OECD, 2013; Cariolle, 2018):

1. ***communication costs = speed x distance traveled*** : the distance between the user and key telecommunications infrastructure nodes a critical determinant of Internet service quality.
2. **spatial hierarchy in telecommunications nodes** favoring urban/coastal areas when disruptions occur: lower (stronger) impact of, and faster (slower) recovery in the core (periphery) after telecommunication network failures.

Empirical strategy

IV-2SLS and IV-multilevel pooled estimations the effect of internet use on firm performance are conducted, based on firm data **aggregated at the location level** :

$$y_{l,s,t,j} = \delta_0 + \delta_1 \mathbf{Internet}_{l,s,t,j} + \delta_2 X_{l,s,t,j} + \delta_3 V_{t,j} + \delta_4 \mathbf{BBI}_{l,s,t,j} + \pi_s + \mu_r + \tau_t + \varepsilon_{2l,s,t,j} \quad (2)$$

With the first-stage equation:

$$\mathbf{Internet}_{l,s,t,j} = \alpha_0 + \alpha_1 \mathbf{Instruments}_{l,s,t,j/t,j} + \alpha_2 X_{l,s,t,j} + \alpha_3 V_{t,j} + \alpha_4 \mathbf{BBI}_{t,j} + \pi_s + \mu_r + \tau_t + \varepsilon_{1l,s,t,j} \quad (3)$$

Where l, s, t, j subscripts respectively refer to the firm's location l , sector s , survey year t , and country j . $\varepsilon_{1l,s,t,j}$ and $\varepsilon_{2l,s,t,j}$ are respectively the error terms of the first-stage and second-stage equations.

$y_{l,s,t,j}$: the firm's **annual total sales, sales per worker, and number of temporary full-time employees (WBES)**

$\mathbf{Internet}_{l,s,t,j}$: **email use, or website use** during firm operations (dummies) (WBES)

$X_{l,s,t,j}$: firm's size, age, and property status; the size of the city where the firm operates; the firm's share of exports; the share of external funding; and the firm's electricity constraint (WBES)

$V_{t,j}$: the population logarithm, the GDP per capita logarithm, a landlockedness dummy variable, democracy, and primary school enrolment (WDI, Polity IV, Ferdi)

$\mathbf{BBI}_{t,j}$: the number of SMCs, the number of IXPs, and the number of years since the first SMC laying. (Telegeography)

π_s, μ_r, δ_t : Region and year-of-interview fixed-effects.

Original sample

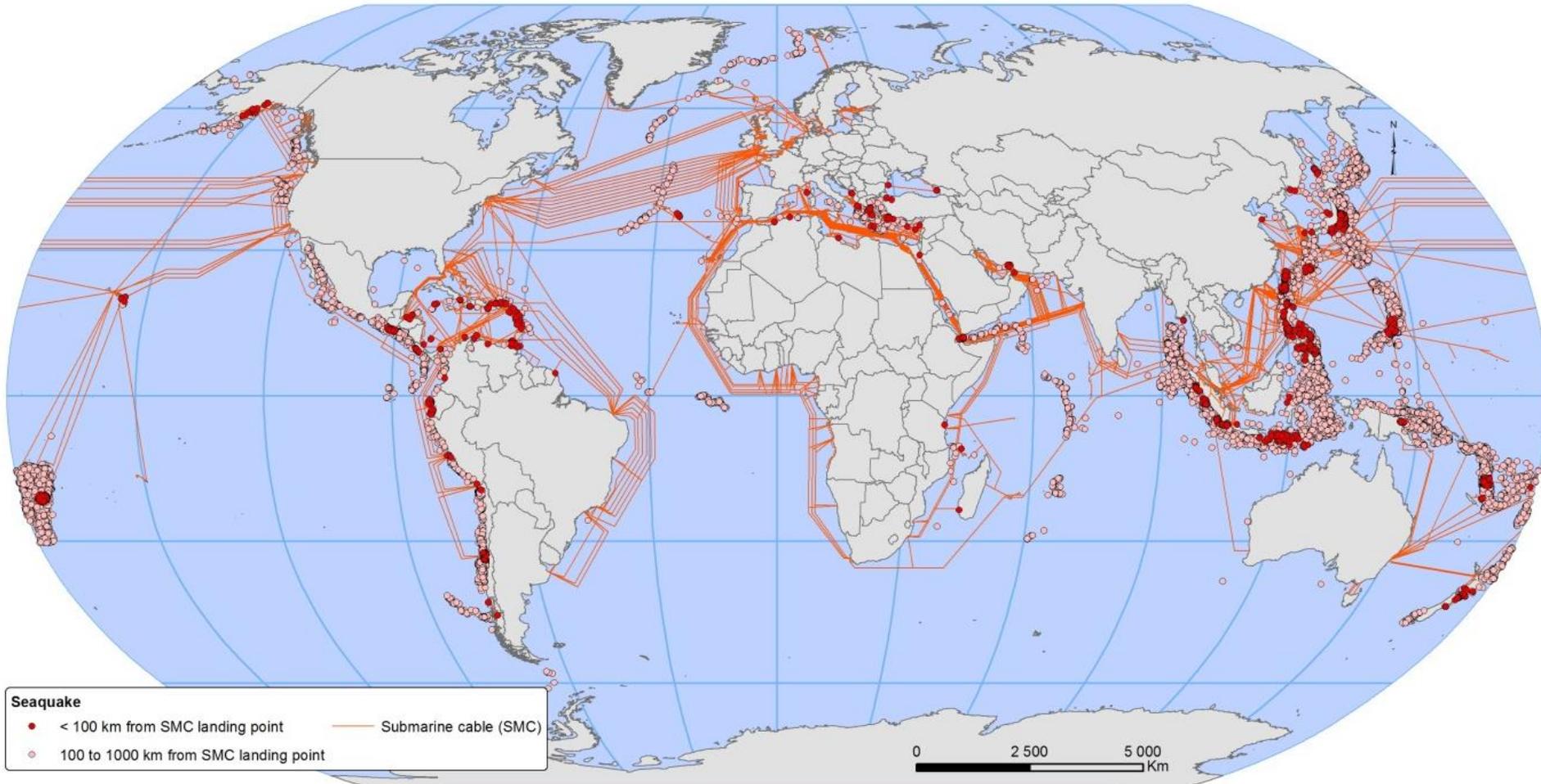
- Around 25-30,000 firms surveyed in some 60 developing/transition countries

Instrumental variables of Internet

- 3 IVs reflecting country-level and local-level vulnerabilities:
 - the **exposure to cable faults** is proxied the **annual frequency of seaquakes** which epicenters are located **within a 1000km radius from SMC landing stations (country-level)**.
 - **Digital isolation** is proxied by
 - i) the **geographical centroid distance** to the closest SMC landing station **(country-level)**.
 - ii) the **firm's distance to the closest infrastructure nodes** – SMC landing stations or Internet exchange point (IXP) **(local-level)**.

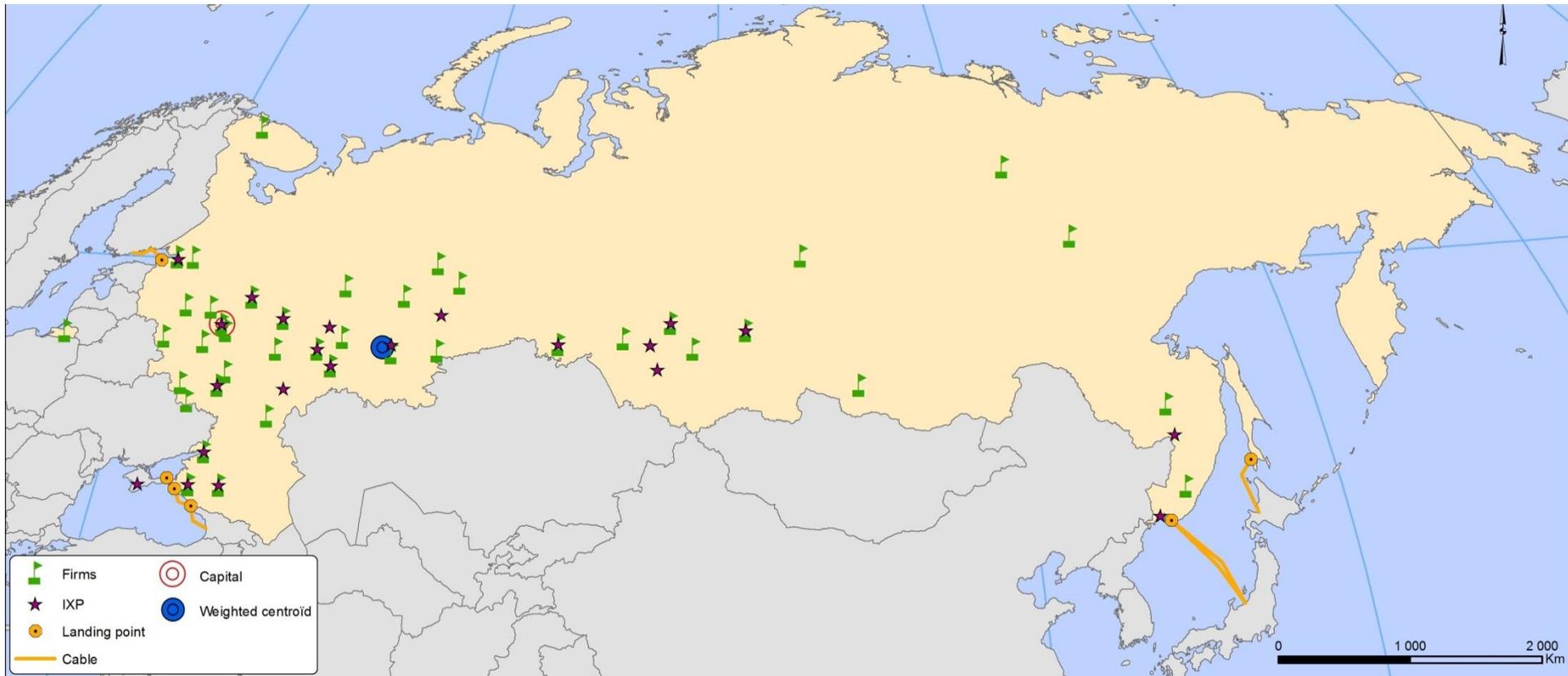
Main variables and data

International seismic activity around SMC landing stations, 1990-2016



Main variables and data

Infrastructure deployment, weighted centroid and firms location in Russia



Data aggregation

IV estimations using firm-level data would lead to important biases caused by:

1. **network externalities** (Röller & Waverman, 2001) by which firm-level decisions (including decisions to use the Internet or not) may be affected locally by behaviours of other firms from the same sector and/or location.
2. **IV regressions with a firm-level dichotomous endogenous variable** - Internet use - may cause inference problems, and displays lower variability among firms

To address these issues, we conduct IV pooled estimations **averaging firm data at the country>survey-year>sector>location level**, so that

- i. the dummy variable becomes a continuous variable lying between 0 and 1
- ii. we neutralize network effects of Internet use among firms sharing geographical and sectorial characteristics.
- iii. we keep the hierarchy in the data, which makes multi-level analysis possible (Hox, 2010; Stoker, 2010).

Multi-level modelling

- the eventuality of **intra-class correlation** at various levels of the data (country, sector, or location levels) casts doubt on the unbiasedness of single-level estimations (Moulton, 1990; Hox, 2010), especially if we are interested in the local effect of aggregated variables.
- We therefore control for **intra-class correlation** by estimating equations (2) and (3) through a multi-level estimation framework:

$$y_{l,s,t,j} = A_0 + A_1 Internet_{l,s,t,j} + \alpha_2 X_{l,s,t,j} + \alpha_3 V_j + \alpha_4 BBI_{t,j} + \pi_s + \mu_r + \tau_t + \varepsilon_{l,s,t,j} \quad (3')$$

$$Internet_{l,s,t,j} = \Delta_0 + \Delta_1 Instruments_{l,s,t,j/t,j} + \delta_2 X_{l,s,t,j} + \delta_3 V_j + \delta_4 BBI_j + \pi_s + \mu_r + \tau_t + \varepsilon_{l,s,t,j} \quad (2')$$

with the following **random components (0-mean, constant var)**: $\Delta_0 = \delta'_0 + \delta''_{0,j} + \delta'''_{0,s,j}$, $A_0 = \alpha'_0 + \alpha''_{0,j} + \alpha'''_{0,s,j}$, $\Delta_1 = \delta'_1 + \delta''_{1,j}$, and $A_1 = \alpha'_1 + \alpha''_{1,j}$.

We implement a multilevel IV setup, with **100-replications non-parametric double-bootstrap of the first-stage and second-stage equations**, to correct the standard errors of the second-stage equation

Results

IV-2SLS single-level estimations

	(1)	(3)	(6)
	Total sales	Sales/worker	Temporary employment
2nd stage estimates			
Email use	2.394**	2.258**	1.746**
	(0.964)	(0.917)	(0.878)
BBI, macro, and micro controls	Yes	Yes	Yes
1st stage estimates			
Instruments:			
Firm distance to SMC/IXP	-0.015**	-0.015**	-0.012**
	(0.005)	(0.005)	(0.005)
Centroid distance to SMC	-0.060***	-0.060***	-0.043**
	(0.019)	(0.019)	(0.017)
Seaquake freq., 1000km rad.	-0.003*	-0.003**	-0.003
	(0.0018)	(0.0018)	(0.0022)
BBI, macro, and micro controls	Yes	Yes	Yes
Dummies	Region, survey year, sector, firm size		
N	2633	2633	2797
Countries	62	62	66
Hansen test (p. value)	0.75	0.34	0.10
Weak ident. test (F-stat)	11.54***	11.50***	7.16***
Under-ident. test (LM-stat)	27.38***	27.32***	18.73***

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses. Standard errors are robust to heteroscedasticity and are clustered by country-year-sector. Firm-level variables are averaged at the country-year-sector-location level.

IV multi-level estimations

	(2)	(4)	(6)
	Total sales	Sales / worker	Temp employment
	Bootstrapped 3-level IV	Bootstrapped 3-level IV	Bootstrapped 3-level IV
Email use	2.639^{***}	2.061^{***}	0.636^{***}
	(0.259)	(0.229)	(0.160)
BBI, macro, and micro controls	Yes	Yes	Yes
1st stage estimates			
Instruments:			
Firm distance to SMC/IXP	-0.00805 ^{***}	-0.00793 ^{***}	-0.00702 ^{***}
	(0.00232)	(0.00227)	(0.00267)
Centroid distance to SMC	0.00405	0.00424	0.00917
	(0.0109)	(0.00923)	(0.00805)
Sequake freq., 1000km rad.	-0.00350 ^{***}	-0.00344 ^{***}	-0.00333 ^{***}
	(0.00123)	(0.0009)	(0.0012)
BBI, macro, and micro controls	Yes	Yes	Yes
Observation unit	Location	Location	Location
Dummies	Region, survey year, sector, firm size		
Random intercepts	Country, sector	Country, sector	Country, sector
Random slope [email use]	Country, sector	Country, sector	Country, sector
N	2,637	2,630	2,592
Countries	62	62	66
Completed replications	100	100	100

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Control estimates not reported. Standard errors are presented in parentheses. Standard errors are robust to heteroscedasticity and are clustered by country-year-sector. Firm-level variables are averaged at the country-year-sector-location level.

Robustness

Robustness 1

To address an eventual **location selectivity bias** - firm location decisions influenced by the presence of infrastructures - we follow a similar approach to Dollar et al. (2006) and **exclude from the sample large firms (1.), exporters and foreign firms (2.), and firms created after SMC arrival (3.).**

1. Sample: SMEs

	(1) Total sales	(2) Sales/ worker	(3) Temp. emp.	(4) Total sales	(5) Sales/ worker	(6) Temp. emp.
	IV-2SLS			3-level IV		
<i>Email use</i>	<i>2.229**</i> (0.937)	<i>1.843**</i> (0.864)	<i>1.347*</i> (0.800)	<i>3.873***</i> (0.255)	<i>3.118***</i> (0.196)	<i>0.511***</i> (0.177)
First stage estimates:						
Firm distance to SMC/IXP	-0.015*** (0.005)	-0.015*** (0.005)	-0.014*** (0.005)	-0.0102*** (0.0027)	-0.0100*** (0.0033)	-0.0083*** (0.0026)
Centroid distance to SMC	-0.058*** (0.018)	-0.057*** (0.018)	-0.046*** (0.017)	0.0056 (0.011)	0.0058 (0.011)	0.0135 (0.0097)
Seaquake freq., 1000km rad.	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004*** (0.0014)	-0.004*** (0.0012)	-0.0036*** (0.0013)
Hansen test (p. value)	0.96	0.47	0.13			
Weak ident. test (F-stat)	9.73***	9.70***	6.99***			
Under-ident. test (LM-stat)	23.74***	23.68***	18.06***			
				Random effect parameters		
Random intercepts				Country, sector	Country, sector	Country, sector
Random slope [email use]				Country, sector	Country, sector	Country, sector
BBI, macro, micro controls Dummies	Yes	Yes Region, sector, year of interview, firm size	Yes	Yes	Yes	Yes
<i>N</i>	2527	2524	2682	2527	2591	2484
<i>Countries</i>	62	62	66	62	62	66

Robustness 1

2. Sample: domestic non-exporter SMEs

	(1)	(2)	(3)	(4)	(5)	(6)
	Domestic non-exporter SMEs					
	IV 2SLS			3-level IV		
	Total sales	Sales/ worker	Temp. employment	Total sales	Sales/ worker	Temp. employment
<i>Email use</i>	1.082 (1.075)	1.168 (0.935)	1.568 (1.070)	4.152*** (0.274)	3.416*** (0.232)	0.140 (0.173)
<i>1st stage estimates:</i>						
Firm distance to SMC/IXP	-0.0135** (0.005)	-0.0135** (0.005)	-0.0113** (0.005)	-0.0101*** (0.0026)	-0.0100*** (0.0027)	-0.0086*** (0.0027)
Centroid distance to SMC	-0.059** (0.020)	-0.059** (0.020)	-0.044** (0.020)	0.0132 (0.0109)	-0.0132 (0.0115)	0.0180* (0.009)
Seaquake freq., 1000km rad.	-0.003 (0.002)	-0.003 (0.002)	-0.0025 (0.002)	-0.0036*** (0.0012)	-0.0036*** (0.0012)	-0.0033** (0.0014)
Hansen test (p. value)	0.93	0.80	0.17			
Weak ident. test (F-stat)	6.81***	6.80***	4.81***			
Under-ident. test (LM-stat)	17.42***	17.40***	13.09***			
Random intercepts				Country, sector	Country, sector	Country, sector
Random slope [email use]				Country, sector	Country, sector	Country, sector
BBI, macro, micro controls	Yes	Yes	Yes	Yes	Yes	Yes
Dummies		Region, sector, year of interview, firm size				
<i>N</i>	2574	2571	2559	2344	2341	2470
<i>Countries</i>	62	62	66	62	62	66

Robustness 1

3. Sample: firms created before SMC arrival

	(1)	(2)	(3)	(4)	(5)	(6)
	Total sales	IV 2SLS Sales/ worker	Temp. employment	Total sales	3-level estimations Sales/ worker	Temp. employment
Email use	2.202*	2.605**	1.794	3.933***	3.124***	0.493**
	(1.168)	(1.091)	(1.329)	(0.240)	(0.291)	(0.243)
<i>1st stage estimates:</i>						
Firm distance to SMC/IXP	-0.012*** (0.005)	-0.012*** (0.005)	-0.009** (0.004)	-0.0067** (0.00294)	-0.0068** (0.0029)	-0.0057** (0.00283)
Centroid distance to SMC	-0.062*** (0.021)	-0.063*** (0.021)	-0.053*** (0.019)	-0.0031 (0.0119)	-0.0031 (0.0095)	0.0038 (0.0096)
Seaquake freq., 1000km rad.	-0.003 (0.002)	-0.003 (0.002)	-0.003 (0.002)	-0.0040*** (0.0011)	-0.0040*** (0.0013)	-0.0036*** (0.0012)
Hansen test (p. value)	0.86	0.40	0.62			
Weak ident. test (F-stat)	6.18***	6.17***	4.73***			
Under-ident. test (LM-stat)	16.32***	16.30***	12.67***			
Random intercepts				Country, sector	Country, sector	Country, sector
Random slope [email use]				Country, sector	Country, sector	Country, sector
BBI, macro, micro controls Dummies	Yes	Yes	Yes	Yes	Yes	Yes
	Region, sector, year of interview, firm size					
<i>N</i>	2431	2428	2559	2431	2428	2561
<i>Countries</i>	62	62	66	62	62	66

Robustness 2

We also use another proxy for Internet use: **the use of website during operations**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	IV-2SLS				3-level IV			
	Total sales	Sales/worker	Employment		Total sales	Sales/worker	Employment	
<i>Website use</i>	2.639** (1.149)	2.607** (1.051)	2.721* (1.472)	2.585** (1.286)	3.185*** (0.342)	2.734*** (0.299)	-0.113 (0.237)	
1st stage estimates:								
Firm distance to SMC/IXP	-0.014*** (0.005)	-0.014*** (0.005)	-0.013*** (0.005)	0.014*** (0.005)	-0.011*** (0.002)	-0.011*** (0.002)	-0.010*** (0.002)	
Centroid distance to SMC	-0.051*** (0.023)	-0.051*** (0.023)	-0.024 (0.020)	-	-0.0138 (0.017)	-0.0140 (0.017)	-0.0003 (0.018)	
Seaquake freq., 1000km rad.	-0.002 (0.002)	-0.002 (0.002)	-0.0016 (0.0017)	-	-0.003*** (0.001)	-0.003*** (0.001)	-0.0026** (0.001)	
Hansen test (p. value)	0.76	0.50	0.28	-				
Weak ident. test (F-stat)	5.27***	5.27***	3.12**	8.55***				
Under-ident. test (LM-stat)	13.92***	13.92***	9.04**	8.06***				
Random intercepts					Country, sector	Country, sector	Country, sector	
Random slope [website use]					Country, sector	Country, sector	Country, sector	
BBI, macro, micro controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Dummies			Region, sector, year of interview, firm size					
<i>N</i>	2631	2628	2795	2795	2631	2632	2798	
<i>Countries</i>	62	62	66	66	62	62	66	

All estimations (main and robustness) are compared to a benchmark estimation framework, including fixed effect dummies but using the location-level digital isolation variable as one single IV.

Estimates are robust this alternative estimation framework are consistent with main and robustness estimations.

Concluding remarks

Concluding remarks

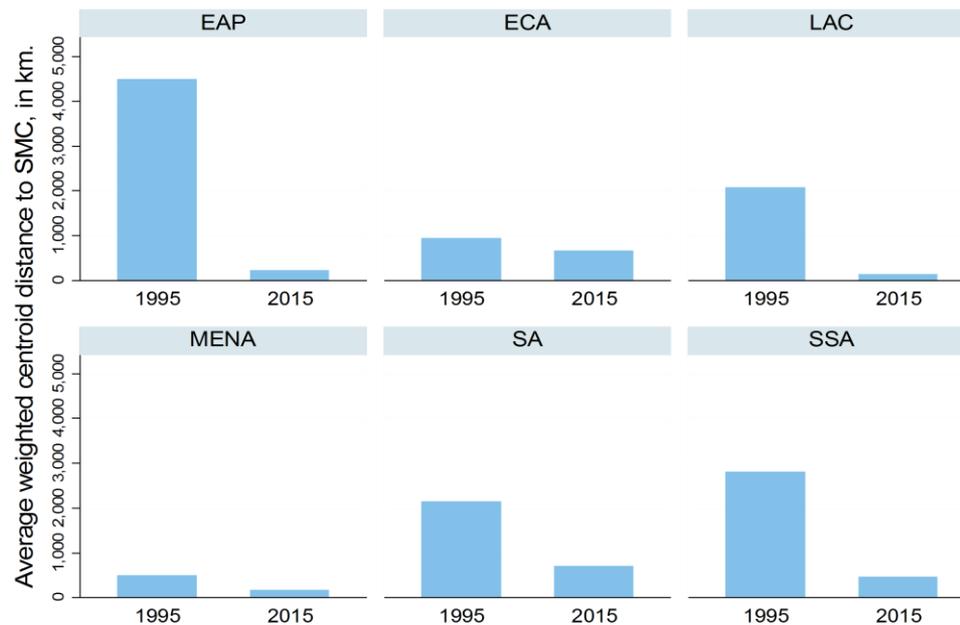
- IV estimations put in evidence the positive contribution of ICTs to firms' performances in developing countries.
- An increased used of e-mail, website are found in a consistent way to stimulate sales, productivity, and to a lower extent, temporary employment.
- These results point out the importance of the broadband infrastructure to the economy, by highlighting the concomitant digital vulnerabilities some developing countries may be subject to:
 - in fact, in some countries, the arrival of fiber boosts the internet economy
 - but may have enlarged the digital divide between coastal and inland areas, and increased the economy's exposure to fiber cable breaks and internet shutdowns

Thank you!

Main variables and data

Digital isolation

Graph 2. Average regional weighted centroid distances to SMC landing stations, 1995 and 2015.



Source: authors. Raw data on SMC landing station coordinates drawn from Telegeography. EAP: East-Asia and Pacific; SSA: Sub-Saharan Africa; MENA: Middle-East and North Africa; LAC: Latin America and Caribe; ECA: (Eastern)-Europe and Central Asia; SA: South-Asia.