



**HAL**  
open science

## Spatial internet spillovers in manufacturing

Joël Cariolle, Maëlan Le Goff

► **To cite this version:**

| Joël Cariolle, Maëlan Le Goff. Spatial internet spillovers in manufacturing. 2021. hal-03201735v3

**HAL Id: hal-03201735**

**<https://hal.science/hal-03201735v3>**

Preprint submitted on 1 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# Spatial internet spillovers in manufacturing

JOËL CARIOLLE

MAËLAN LE GOFF\*

➤ JOËL CARIOLLE, Research Officer, FERDI. Corresponding author.  
Contact: [joel.cariolle@ferdi.fr](mailto:joel.cariolle@ferdi.fr)

➤ MAËLAN LE GOFF, Agence française de développement (AFD), Paris, France.  
Banque de France, Paris, France.

## Abstract

Does local internet diffusion spur manufacturing firm performance in developing countries? To answer this question, we conduct instrumental variable estimations, using repeated cross-section data on 40,154 manufacturing firms from 91 developing and transition economies, and find that a 10 percentage-point increase in the incidence of email use in locations where firms operate, raises by 36% their sales and sales per worker. This evidence turns out to be driven by the local dissemination of email technology within industries, rather than across industries. This higher performance in manufacturing is also found to be accompanied by output diversification, driven by inter-industry spillovers, and workforce contraction, driven by intra-industry ones. However, we further provide evidence of U-shaped inter-industry spillovers, i.e. negative inter-industry spillovers turning positive once the email incidence threshold reaches 50% of the local universe of firms, suggesting that network effects are at play. Last, these threshold effects seem related to the presence of outward-oriented firms, which tend to operate in places where internet is more diffused. Overall, this paper shows that local industrialisation paces may diverge between poorly and highly digitalised environments.

**Keywords:** Connectivity, internet, spillovers, manufactures, industrialisation.

**JEL classification:** F61, L25, O33, O14, O18.

\* † This work was started when Maëlan Le Goff was affiliated to the Banque de France.

# 1 Introduction

Constraints on economic interactions are particularly acute in developing countries, where hard infrastructures are often missing and market imperfections are widespread. By reducing transaction costs and improving firms' production and organisation processes (Aker, 2017; Goldfarb & Tucker, 2019), digital technologies have become essential to the conduct of business in both high- and lower-income countries.<sup>1</sup> Beyond the benefits derived by the firms from their own adoption of such technologies, their dissemination in their proximate environment may also indirectly affect their activity.

In fact, a greater diffusion of digital technologies, such as the internet, in a geographical area, within or across industries, can generate positive spillover effects on firms' activity via network effects or digital knowledge spillovers (Marsh, Rincon-Aznar, Vecchi, & Venturini, 2017; Paunov & Rollo, 2015, 2016). Such externalities are inherent to digital technologies being general purpose technologies and network goods, since their benefits tend to spill over industries and to increase with the size of the user network (Björkegren, 2019; Bresnahan & Trajtenberg, 1995; Crémer, Rey, & Tirole, 2000; Katz & Shapiro, 1985). However there is also a risk that digitalisation<sup>2</sup> primarily benefits first adopters or dominant firms, at the expense of less performing firms with limited absorptive capacity (Görg & Greenaway, 2004; Marsh et al., 2017), and that it pushes forward the most advanced sectors of the economy to the detriment of the more traditional ones (Choi, Dutz, & Usman, 2020; Hjort & Poulsen, 2019; Rodrik, 2018). The digitalisation process could hence accelerate the decline of certain industries, widen regional economic disparities, and even assist in the de-industrialisation process in poorly digitalised places (Rodrik, 2016a, 2016b, 2018; Wu, Wang, & Sun, 2021). Therefore, the net economic benefits drawn from the diffusion of digital technologies are likely but not guaranteed.

The uncertainty over the digital dividends is particularly salient for firms operating in the manufacturing sector, especially those located in developing countries. While there is strong evidence of the positive effect of digitalisation on the service sector, at both the micro-level (Kneller & Timmis, 2016) and the macro-level (Freund & Weinhold, 2002), it is much less clear whether this process has benefited the manufacturing sector (Stiroh, 2002). Yet, some argue that knowledge spillovers should be greater in manufacturing than in services because technological progress should be higher in the former (Glaeser, Kallal, Scheinkman, & Shleifer, 1992). Second, evidence on the consequences of the manufacturing sector's digitalisation is widely documented for industrialised countries, but rather scarce in the case of developing ones.<sup>3</sup> Yet, the manufacturing sector is at the core of the industrialisation process, and thereby, identified as a critical source of income and job creation in developing countries (Tybout, 2000; Rodrik, 2016a, 2018). A better understanding of the contribution of internet technologies to manufacturing firms' performance is therefore of utmost importance for developing countries, especially those caught in an under-industrialisation trap or suffering from premature de-industrialisation (Diao, McMillan, & Rodrik, 2019; Rodrik, 2016b). Third, while recent studies have addressed the direct impact of firms' adoption of internet technology on their performance in low-income countries (Cariolle, Le Goff, & Santoni, 2019; Hjort & Poulsen, 2019; IMF, 2020), less is known about the spillovers stemming from the local diffusion of internet technologies in developing areas. The issue of spatial spillovers in the manufacturing sector of industrialised economies has been, however, extensively covered (Cingano & Schivardi, 2004; Martin, Mayer, & Mayneris, 2011; Wixe, 2015). This paper fills this gap by estimating spatial internet

spillover effects – resulting from the local diffusion of email technology – on the performance of firms in the manufacturing sector, in a large sample of developing and transition economies. The decomposition of local intra- and inter-industry spillovers, as well as the identification of network effects, also represent a significant add.

Another contribution of this paper comes from our identification strategy, which is meant to address two statistical challenges. First, an individual firm’s performance may affect overall local economic activity, and hence, neighbouring firms’ inclination to adopt emails, through for instance imitation behaviours. Such a mechanism could be a source of reverse causality bias. Second, omitted variables, especially unobserved local conditions, may influence both firm’s performance and the local incidence of internet technologies. To address these sources of endogeneity, the local incidence of email use among firms, our measure of local internet spillovers, is instrumented by a set of shift-share type instrumental variables (IVs). In our setting, identification follows from the quasi-random assignment of infrastructure-related aggregate connectivity shocks (Borusyak & Hull, 2020; Borusyak, Hull, & Jaravel, 2021), weighted by the firm’s location distance to connectivity infrastructures. This IV setting incorporates country-year, location, and stratification sector fixed-effects, thus reducing the concern for eventual omitted variable bias that could make connectivity shocks and exposure to them non-randomly assigned. We also perform a range of robustness, sensitivity, pre-trend and falsification tests that support the relevance of our identifying assumptions.

Our empirical analysis combines firm-level data, drawn from the World Bank Enterprise Surveys (WBES), with data on the roll-out of telecommunication submarine cable (SMC) infrastructures from Telegeography and an original data-set on internet disruptions caused by SMCs’ faults. Our baseline estimation sample comprises 40,154 manufacturing firms, surveyed between 2006 and 2018 through 11 survey waves, located in 521 cities or regions, and spread across 91 developing and transition economies. Our results support that local diffusion of email technology in the manufacturing sector generates substantive net positive spillovers in terms of revenue and labour productivity. Moreover, distinguishing between intra-industry and inter-industry local spillovers, we stress that this evidence is driven by the local dissemination of the email technology within rather than outside the industry. However, we also show that local inter-industry spillovers are U-shaped, i.e. that negative inter-industry spillovers turn positive once reaching a local email incidence threshold of approximately 50%, suggesting that network effects are at play. Going further, we find that these positive (negative) inter-industry spillovers on the firm’s output are actually linked with its outward (inward) orientation, namely, with its exporter, foreign, and multi-plant status. Last, spillover effects are also found to translate into a smaller permanent workforce size and a greater product diversification. Overall, by estimating and disentangling spatial internet spillovers in manufacturing, our analysis highlights the contribution of digitalisation to industrialisation, and probably structural change, in developing economies.

The next section sets out our analytical framework and the related literature review. The third section exposes our empirical framework, while the fourth section presents our main findings. Robustness checks are performed in the fifth section. The sixth section concludes.

## 2 Local internet spillovers and industrialisation: analytical framework

The study of local internet spillovers on developing countries' manufacturing sector is a critical dimension of the digitalisation-industrialisation nexus which has been so far overlooked by the literature. However, because of the general purpose nature of digital technologies (Bresnahan & Trajtenberg, 1995) and their network-based functioning (Katz & Shapiro, 1985), their diffusion across places and industries may spur economic transformations that go beyond the service sector, and that exceed the individual and direct benefits of internet technology adoption.

Industry-level internet spillovers have been examined by Paunov and Rollo (2015, 2016) in the context of developing countries, and by Marsh et al. (2017) in the case of the United-States. However, as suggested by studies on agglomeration economies (Caragliu et al., 2016; Frenken et al., 2007; Glaeser et al., 1992; Jacobs, 1969; Marshall, 1890), spillovers from ICTs may materialize inside a delimited geographical area. In fact, while internet diffusion makes it possible to overcome physical distances and to share knowledge and information with very distant entities, this is especially valid for codified information, the exchange of non-codified information requiring physical encounters (Storper & Venables, 2004). Thus, knowledge and information spillovers made possible by the diffusion of internet should take place locally, between proximate firms, since the transmission of a certain type of information has to be supplemented by face-to-face interactions which are easier in delimited geographical areas, especially urban centers (Gaspar & Glaeser, 1998; Wu et al., 2021).

Spatial information spillovers are a critical "engine of growth", generating *intra-industry* and *inter-industry* externalities (Glaeser et al., 1992; Jacobs, 1969; Lucas, 1988; Marshall, 1890; Romer, 1986). On the one hand, the diffusion of information and knowledge within proximate industries is considered as a source of localisation/specialisation externalities (Marshall-Arrow-Romer externalities), while on the other hand, its diffusion across industries is considered as a source of diversification externalities (Jacob externalities).<sup>4</sup> Getting back to digitalisation, spatial intra-industry internet spillovers hence refer to the transmission of knowledge, ideas and information related to internet technologies, between firms operating in a same industry and located in the same place. They occur, for instance, when firms spy or imitate proximate first movers, or through inter-firm movements of the workforce endowed with digital skills. By contrast, local inter-industry internet spillovers are linked to the diffusion of internet technologies across proximate industries, facilitating upstream and downstream industries communication and exchange ideas, new processes and practices, or the creation of new inter-industry linkages.<sup>5</sup>

In Marshall and Jacob externality models, spillovers stem from the local diffusion of knowledge within and across industries, respectively. However, internet technologies are not only general purpose technologies whose related knowledge spill over firms and industries, they are also network goods (Crémer et al., 2000; Katz & Shapiro, 1985). The benefits derived from ICTs' adoption indeed depend on the users' network size: the greater the number of users of a digital technology in a given location or industry, the greater the benefits derived from its adoption (Katz & Shapiro, 1985). In a developing context, network effects of digital technologies have been empirically evidenced in Rwanda and China by BJORKEGREN (2019) and Wu et al. (2021), respectively. Therefore, on the one hand, the local diffusion of internet may induce *knowledge spillovers* resulting from the sharing of information, the imitation of good practices and processes, and the dissemination of innovations related to internet technologies. On the other hand, it may

generate *network effects*, leading to the multiplication and acceleration of interactions and information exchanges between adopters in a same network (Grace et al., 2003; Stiroh, 2002; Wu et al., 2021).

However, internet technologies diffusion will not be equally absorbed by manufacturing firms and industries. In fact, negative internet spillovers may prevail if the increased use of related digital technologies by some firms translates into greater competition, which in turn leads to revenue losses for firms with a limited technology absorptive capacity (Görg & Greenaway, 2004; Marsh et al., 2017). According to Marsh et al. (2017), this limited absorptive capacity can be explained by a lack of digital skills within the firm, by the delayed diffusion of digital technologies within industries, by a low exposure to international competition, or by insufficient research and development (*R&D*) activities. On the longer run, the diffusion of digital technologies and related knowledge across industries may also spur structural change, increasing disparities between cities with low and high internet penetration rates (Wu et al, 2021), causing the decline of traditional industries using obsolete technologies or made obsolete by technological shift (McMillan, Rodrik, & Sepulveda, 2016; Choi et al., 2020; Diao et al., 2019).

In this paper, we focus on and disentangle the spillover effects generated by the local diffusion of email technology, both within and across industries of the manufacturing sector. We also highlight certain characteristics of companies that allow them to take advantage of internet spillover effects to a greater or lesser extent.

### 3 Empirical framework

#### 3.1 Baseline model

To estimate the causal effect of the local diffusion of internet technologies on manufacturing firms' performance, our IV approach consists in estimating the following second-stage and first-stage equations:

$$Y_i = \alpha_0 + \alpha_1 \cdot \overline{Internet}_{l(i)t(i)} + \alpha_2 \cdot \mathbf{X}_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + \varepsilon_{1i} \quad (1a)$$

$$\overline{Internet}_{l(i)t(i)} = \beta_0 + \beta_1 \cdot Z_{l(i)t(i)} + \beta_2 \cdot \mathbf{X}_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + \varepsilon_{2l(i)t(i)} \quad (1b)$$

Where the subscripts  $i$ ,  $j$ ,  $s$ ,  $l$ , and  $t$  respectively refer to the firm, country, stratification sector, location, and fiscal year.  $Y_i$  measures the performance of the firm,  $\overline{Internet}_{l(i)t(i)}$  is the incidence of email use in location  $l$  at time  $t$ , which measures the spatial internet spillovers.  $Z_{l(i)t(i)}$  represents the set of instruments and  $X_i$  is a vector of firm-level characteristics. These equations also include country-year ( $d_{j(i)t(i)}$ ), sector ( $d_{s(i)}$ ), location dummies ( $d_{l(i)}$ ), and random error terms ( $\varepsilon_i$ ). Standard errors are robust to heteroscedasticity and clustered at the country-year level.

#### 3.2 Data and descriptive statistics

The variables used in our baseline model are drawn from the standardised World Bank Enterprise Surveys (WBES). They provide repeated cross-sectional data, covering an original representative sample (random stratified sampling) of small, medium and large-size firms, operating in formal, non-agricultural sectors and urban areas of developing and transition countries. Our baseline sample covers 40,154 manufacturing firms from 521 locations (cities or regions) in 91 developing and transition countries, surveyed over the

period 2006-2018 (taking associated fiscal years as time reference)<sup>6</sup>. In each country, data were gathered using an extensive and internationally comparable questionnaire administered via face-to-face interviews with business owners and senior managers. Appendix A presents the sources and definitions of variables used in this study while Appendix B provides descriptive statistics.

### 3.2.1 Firms' performance ( $Y_i$ ).

To measure the performance of a firm we use two output variables: the logarithm of the firm's total annual sales (in USD) and the logarithm of the firm's total annual sales per full-time permanent worker adjusted for temporary workers (in USD) to measure labour productivity, as commonly used in the literature (Chemin, 2020; Léon, 2020). Estimations using additional outcome variables – permanent/temporary workforce size and diversification variables – are also conducted.

### 3.2.2 The spatial diffusion of email technology ( $\overline{Internet}_{l(i)t(i)}$ ).

Internet diffusion is proxied by the incidence of email use among manufacturing and service firms in the location where the firm operates. We focus on email use rather than website ownership (but controlling for it), the other internet-related variable included in the WBES. Indeed, while a company may have a website without using internet on a daily basis or even having an internet connection (Fabiani et al., 2005), the use of emails reflects a cheaper, multipurpose (internal or external communications), less strategic and less skill-intensive use of internet (Sadowski et al., 2002). Moreover, emails can be accessed via a slow, fixed or mobile, internet connection. For these reasons, we consider that the incidence of email incidence among firms is a better indicator of internet diffusion than website incidence.

Therefore, we follow Acs et al. (1994), Haskel, et al. (2007) and Paunov and Rollo (2016), and compute the incidence rate of email use at the location level :

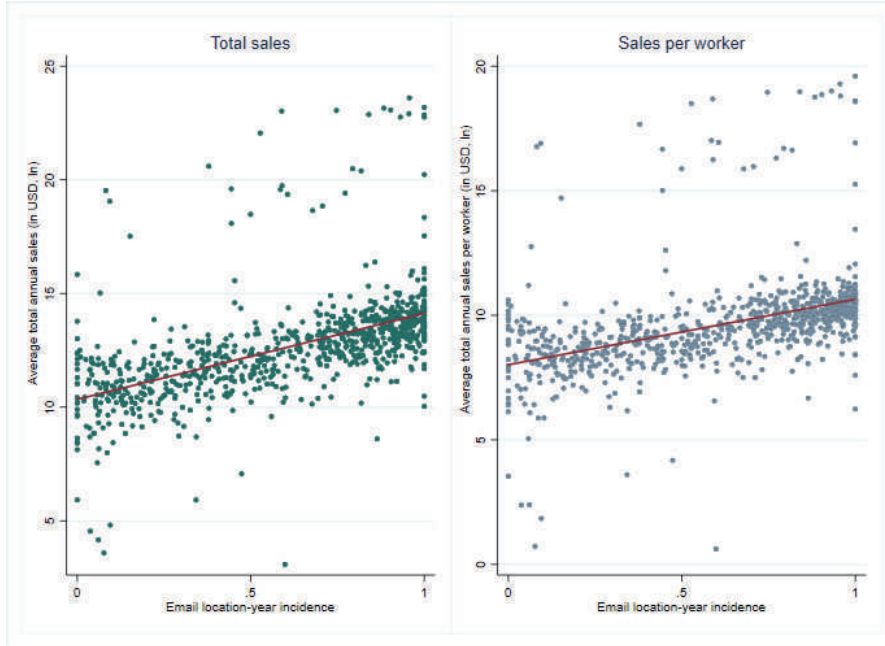
$$\overline{Internet}_{l(i)t(i)} = \frac{1}{N_l - 1} \left( \sum_{f \in l} Email_{f,l} - Email_{i,l} \right) \forall l \in L, i \neq f \quad (2)$$

Where *Email* is a dichotomous variable indicating whether firms *i* and *f* in location *l* use email for the conduct of business - or not -. *L* refers to the set of locations *l* where firms operate, and  $N_l$  refers to the respective number of firms in each location *l*. This incidence variable is computed imposing a minimum of 10 observations by location and excluding manufacturing firm *i*'s own adoption of email technology to address eventual reverse causality bias. This approach is also justified by the necessity to separate the direct effect of a firm's individual decision to adopt email technology (included as a control variable), from the spillover effects of its diffusion at the location level.

Figure 1 graphically represents the simple correlation between email incidence and firms' average performances. It highlights a strong and positive relationship between email incidence and firms' output and productivity.



Figure 1: Email incidence in locations and manufacturing firms' performance.



Source: WBES data and authors' calculation based on 868 pooled observations (650 locations from 120 countries).

### 3.2.3 Control variables

We control for a set of firm-level characteristics whose impact on performance has been evidenced in the literature (Dollar, Hallward-Driemeier, & Mengistae, 2005; Paunov & Rollo, 2015, 2016). To separate the effect of the individual adoption of email technology from spillover effects caused by its local diffusion, we control for email and website adoption at the firm level. We also control for the size and composition of the workforce, measured by the number of full-time permanent employees when the firm started operations, the share of non-production workers in the total workforce, and the share of skilled workers among production workers. Since management quality may be an important determinant of the firm's absorptive capacity, we control for the firm's maturity, proxied by the firm's age (in years) and its top manager's experience (in years). We control for determinants of the firm's performance such as its public and foreign ownership structure, its degree of export orientation, and its financial liabilities (measured by a dummy equal to one if the firm has a credit line or a loan from a financial institution). Since access to internet relies on access to energy, we also take into account the firm's electricity constraint, as reported by the firm.<sup>7</sup> Last, our set of control variables also comprises the geographical distance to the closest international connectivity infrastructure – i.e. the closest submarine cable (SMC) landing stations or Internet Exchange Point (IXP) in the country where the firm operates – which also serves as weight in our instrumental variable setup, hereafter described. Summary statistics of standard WBES variables used in our empirical analysis are reported in Table 1.



Table 1: Summary statistics of WBES variables.

Variables	Mean	Std. Dev.	Min	Max
<b>Firm outcomes (<math>Y_i</math>)</b>				
Real annual sales (USD, ln)	13.30	2.56	0	27.19
Real annual sales / worker(USD, ln)	9.87	1.95	0	21.90
<b>Internet spillovers (<math>Internet_{l(i)t(i)}</math>)</b>				
Email location incidence [0;1]	0.72	0.25	0	1
<b>Control variables (<math>X_i</math>)</b>				
Email adoption (0 or 1)	0.72	0.45	0	1
Website adoption (0 or 1)	0.45	0.50	0	1
% of state ownership	0.51	5.77	0	100
Distance to connectivity infra (km, ln)	4.65	2.00	0.69	8.35
% of foreign ownership	6.55	22.9	0	100
% dir. indir. Exports in sales	12.46	27.50	0	100
Firm's age (years, ln)	2.74	0.70	0	4.76
Initial # perm. FT employees(ln)	2.69	1.25	0	13.82
% non-production workers in total workforce	0.70	0.31	0	1
% skilled workers in production workforce	0.25	0.17	0	1
Manager experience (years, ln)	2.66	0.75	0	4.31
Bank loan (0 or 1)	0.38	0.48	0	1
Electricity obstacle (ordered, 0 to 4)	1.79	1.49	0	4

Sample: 40,154 manufactures from 521 locations (cities/regions) in 91 developing and transition countries.

### 3.3 Instrumental variable approach

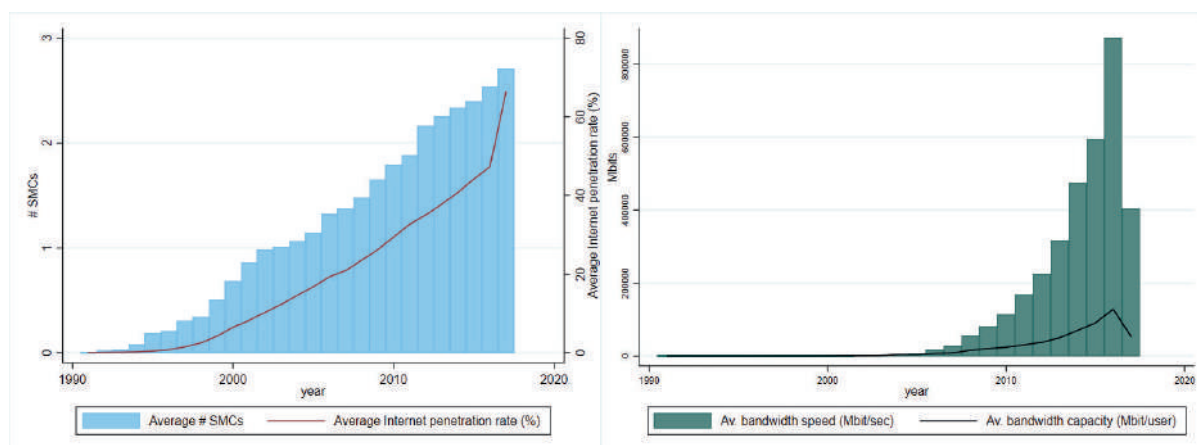
To estimate the causal effect of email incidence on manufacturing firms' performance, we adopt a shift-share instrumental variable (SSIV) approach, formalized by Borusyak et al. (2021), consisting in weighing random aggregate connectivity shocks by a plausibly endogenous factor reflecting firms' exposure to them.

#### 3.3.1 Aggregate connectivity shocks

Our instrument set combines two complementary sources of random variation in aggregate connectivity, related to the deployment of telecommunication submarine cables (SMCs): i) the variation in the number of SMCs laid in a country, and ii) the country's experience of SMC outages, more specifically, the duration of their associated repairs. Because the risk of cable outages naturally arises from the laying of SMCs and increases with their number, considering together these two nested sources of aggregate connectivity reduces the concern for under-identification and sheds light on the mechanisms tying SMC connectivity and internet diffusion.<sup>8</sup> Moreover, having two instruments allows instrumenting two endogenous variables, which is be done when estimating inter- and intra-spillover effects and identifying thresholds in spillovers effects.

**The number of SMCs.** SMCs are the cornerstone of the worldwide telecommunications network, whereby more than 95% of world telecommunications are carried. The exponential rise in their deployment over time has led to a dramatic increase in the worldwide telecommunications network's size, capacity, and redundancy. Figure 2 illustrates this trend by representing the strong increase in internet connectivity – measured by the internet penetration rate and the average international internet bandwidth per user – that has followed the connection of developing countries to the global SMCs' network. While SMCs'

Figure 2: Internet connectivity and international bandwidth in developing countries (1990-2016).



(a) Internet connectivity.

(b) International bandwidth.

Source: Authors' calculation, based on data from ITU (2019) and Telegeography's submarine cable map:

<https://www.submarinecablemap.com/>.

roll-out leads to an increased internet capacity and digitalisation (Cariolle, 2021), it also generates a higher exposure to outage risks, which can be mitigated only if there is enough alternative SMC paths to reroute internet traffic.

One attribute of SMCs, relevant to our empirical strategy, is that their laying faces strong geographical and geological constraints (Carter, 2010; Eichengreen, et al, 2016; World Bank 2018; Cariolle 2021; Simone & Li, 2021), and therefore very high fixed costs. As a result, SMCs are usually managed by large consortia of public and private operators and investors, deployed at the regional or intercontinental levels, and laid in various countries located along their path, to take advantage of scale economies. Nevertheless, (non-) treatment could be subject to a selection bias, i.e. more likely in more (less) economically dynamic countries. That is why our causality claim is subject to fully controlling for cross-country-year characteristics, which is done through the inclusion of country-year fixed-effects.

**SMCs' outages.** The SMCs' infrastructure integrity is critical for international exchanges and threats upon it are a key concern for governments, companies, and international agencies.<sup>9</sup> SMCs' outages, caused by humans (sabotage, maintenance), maritime activities (anchors and fishing nets) or natural hazards (typhoons, earthquakes, turbidity currents), are a critical source of telecommunication shut- or slowdowns (Carter et al., 2010; Carter, 2014; Pope, Talling, & Carter, 2017; Aceto, Botta, Marchetta, Persico, & Pescapé, 2018).<sup>10</sup> Moreover, in addition to the direct economic and welfare costs related to telecommunication disruptions, SMCs' outages induce expensive repairs, higher insurance costs, and additional costs related to the re-routing of internet traffic towards more expensive and lower-capacity cable paths. Importantly, these direct and indirect costs are amplified by the time needed to repair cables (Carter et al., 2010; Palmer-Felgate, Irvine, Ratcliffe, & Bah, 2013; OECD, 2014). While SMCs' outages increase the network's vulnerability and undermines internet stability, the number of SMCs can limit outages' consequences by providing alternative networks in the event of failure of one or various SMCs.

Our IV therefore also exploits country-level variations in cumulative annual time to repair cables, reflecting both the recurrence and duration of adverse shocks incurred by the SMCs' network. This

Table 2: Internet disruptions and repair duration caused by SMCs’ faults over  $(t; t-4)$ , estimation sample.

iso	$\Sigma$ disruptions	$\Sigma$ repairs days	iso	$\Sigma$ disruptions	$\Sigma$ repairs days
ARM	1	1	KEN	3	26
BDI	2	16	LBN	2	13
BEN	1	15	LBR	1	2
BGD	3	9	LKA	2	56
CHN	3	5	MMR	1	20
CMR	3	10	MYS	1	1
COL	1	2	NGA	3	37
DJI	2	22	PAK	3	29
ECU	1	1	PHL	3	53
ETH	1	7	SLE	2	8
GEO	1	1	TZA	1	10
IDN	1	49	YEM	2	12
IND	4	33	ZMB	1	12
JOR	1	1	<b>Total</b>	<b>50</b>	<b>451</b>

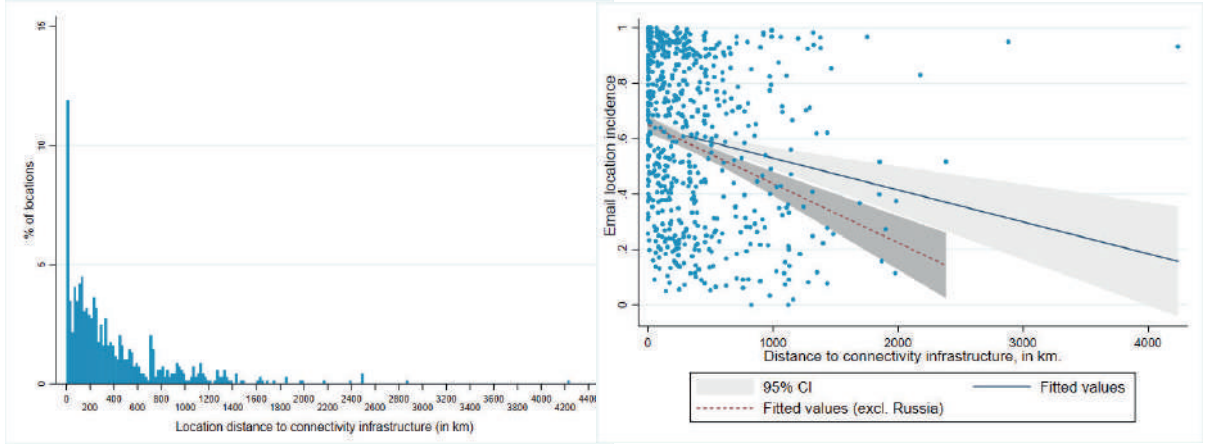
Source: Authors’ calculation. Data drawn from the Subtel forum <http://subtelforum.com/category/cable-faults-maintenance/>, Akamai’s reports on the “State of Internet Connectivity”, and completed by manual internet searches.

has been done by building an original database documenting the occurrence of SMC-induced internet disruptions by country and year, the cause of these outages, as well as the duration (in days) of cable repairs, over the 2005-2020 period.<sup>11</sup> It is worth noting that only cable faults that have caused internet disruptions are considered. Due to possible endogeneity concerns, we drop all observations from the sample that include a cable outage induced by natural hazards (earthquake, hurricanes, floods). In fact, natural shocks could directly affect firms’ performance, regardless of their use of internet, by causing material and human casualties. For similar concern, we also drop observations for which internet disruptions are suspected to have been caused by a government intervention.<sup>12</sup> In our estimation sample, 30 out of 91 countries have experienced internet disruptions caused by SMCs’ faults – induced by ship anchoring, fishing nets, sabotage, or unidentified causes – during the current and four years preceding the reference fiscal year (Table 2). Among these countries, India’s SMCs’ network has been hit four times over  $(t; t-4)$  and has undergone 33 accumulated days of repair, while Sri-Lanka experienced two internet disruptions associated with 56 days of repair.

**Weighting factor: distance to connectivity infrastructure.** Our IV design assumes that firms’ exposure to SMCs’ connectivity shocks depends on their distance to the closest international connectivity infrastructure. In a core-periphery setting, populations remote from (close to) connectivity infrastructures are more (less) exposed to telecommunication network disruptions and suffer (benefit) from worse (better) connectivity than close (remote) ones. This spatial hierarchy in internet connectivity is explained by the necessity to maintain internet access in most important economic and population centres if the whole network capacity is undermined. As a result, urban and coastal areas are favoured, compared to rural and inland ones, when national capacity increases but also when disruptions occur (Gorman & Malecki, 2000; Malecki, 2002; Grubestic et al., 2003; Gorman et al, 2004; Grubestic & Murray, 2006).

Using information on the location of observational units (city or region identification) available in the WBES, we use the distance between the firm’s location and the closest SMC landing station or IXPs as a weight placed upon aggregate shocks. SMC landing stations and IXPs indeed represent key backbone

Figure 3: Location distances to international connectivity infrastructures (SMCs or IXPs): sample distribution and correlation with email incidence.



(a) Sample distribution.

(b) Correlation with location email incidence.

infrastructure nodes and an important source of network efficiency and internet bandwidth (Weller & Woodcock, 2013; OECD, 2014).<sup>13</sup> This distance variable is included alone in the econometric equation to control for the direct effect of infrastructure proximity on local internet diffusion and firm performance. Figure 3a shows the distribution of this distance variable across locations in our baseline estimation sample. Figure 3b plots a negative correlation between the sample locations' distance to connectivity infrastructures and the incidence of email use at the location level.

### 3.3.2 Instrumental variables

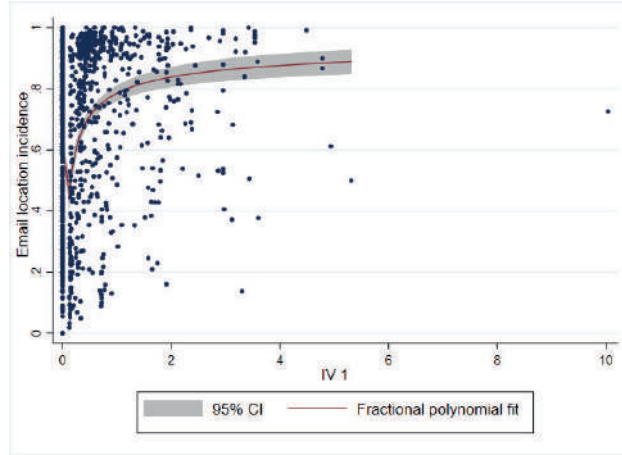
Our IV strategy consists in weighting two main sources of exogenous aggregate connectivity shocks – the number of SMCs laid in a given country and the annual cumulative duration of outage repairs – by the distance from the location to the closest internet connectivity infrastructure. The number of SMCs is divided by the distance to connectivity infrastructure (the closer, the greater the connectivity) while SMCs' outages are multiplied by this distance (the closer, the lower the vulnerability). As a result, we obtain two IVs, exhibiting location-year variability after the inclusion of country-year and location fixed-effects.

The first instrument (*IV1*) is calculated as follows:

$$IV1_{j(i)l(i)t(i)} = SMCnum_{j(i)} \times \frac{1}{1 + \ln(1 + Distance_{l(i)t(i)})} \quad (3)$$

Where  $SMCnum_{j(i)}$  is the number of SMCs laid in country  $j$  and  $Distance_{l(i)t(i)}$  is the distance from the firm's location  $l$  to the closest connectivity infrastructure at time  $t$ . Figure 4 below represents graphically the correlation between this instrument and the incidence of email use at the location level. As expected, the weighted number of SMCs (*IV1*), associated with a greater local internet capacity and stability, increases email incidence at the location level.

Figure 4: Correlation between IV1 and email incidence in locations.



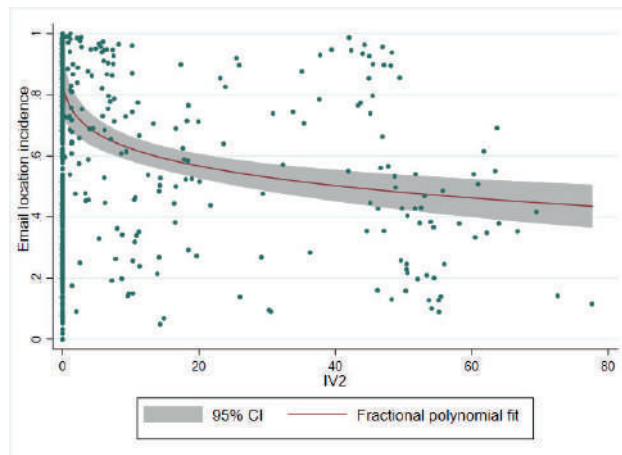
Note: Fitted values have been plotted excluding the outlier (IV1=10)

The second ( $IV_2$ ) reflects firms' exposure to SMC-induced internet disruptions and associated cable repair duration. Because shocks can have a lagging effect on the penetration of internet through a decrease in internet service quality and an increase in internet tariffs, this IV is calculated over various time-windows going up to four years before the fiscal year, as expressed in Eq.(4):

$$IV2_{j(i)l(i)t(i)} = \frac{1}{1 + \tau_2 - \tau_1} \sum_{\tau=t-\tau_2}^{t-\tau_1} (Repair\_days_{j(i)\tau(i)} \times [1 + \ln(1 + Distance_{l(i)\tau(i)})]) \quad (4)$$

Where  $\tau_1 = [0; 3] \leq \tau_2 = [1; 4]$ ,  $Repair\_days_{j(i)\tau(i)}$  is the cumulative number of days necessary to repair damaged cables in country  $j$  in year  $(t - \tau)$ , and  $Distance_{l(i)\tau(i)}$  is the same weighting distance variable as in Eq.(3). In baseline estimations,  $IV_2$ 's time window is set to  $[t; t - 4]$  in order to optimise first-stage statistics. Below, Figure 5 illustrates the negative association between  $IV_2$ , computed over  $[t; t - 4]$ , and the incidence of email use at the location level. IV estimations using different time-windows are reported in Online Appendix A.2.

Figure 5: Correlation between IV2 and email incidence in locations.



To ensure our instrumental variable strategy do not violate exclusion restrictions, we control for unob-

served country-year level and local-level factors that could affect both connectivity and firms' performance by including country-year and location fixed-effects (Henderson, 2003). We also control for the distance to connectivity infrastructure, used as a weighting factor in our IV set-up, because broadband infrastructures may be located in ports or in large urban centres, whose proximity could directly affect manufacturers' performance. Identifying assumptions are further discussed in section 5 and tested in Online Appendixes A and B, proceeding to various calibrations of the IV set-up, falsification tests, and to sample restrictions to avoid eventual weak-instrument, omitted variable and location selection biases.

## 4 Empirical results

### 4.1 Baseline estimations

Table 3 reports the results of IV-2SLS pooled estimations of the baseline econometric model (Equation (1)) on a baseline sample of 40,154 firms located in 91 developing and transition economies. We start by estimating our model without including control variables (column (1)), then including the distance to infrastructure with email and website adoption variables (column (2)), and finally including all control variables (columns(5)-(8)). To check whether estimated spillover effects are directly related to firm's internet adoption and not to indirect channels, we interact  $\overline{Internet}_{l(i)t(i)}$  with email and website adoption variables, separately (columns (5)-(8)).

First-stage test statistics support the validity of the IV set-up. As expected, the instrument based on the SMC network size (*IV1*) has a significant positive effect on email location incidence, while the instrument based on adverse connectivity shocks (*IV2*) has a significant negative effect.<sup>14</sup> Second-stage estimations first indicate a positive and significant effect of email location incidence on manufacturing firms' total sales and sales per worker, independent of the firms' use of email or website. According to our estimates, total sales and sales per worker are found to increase by 3.6 percentage points following a 1 percentage point increase in the local incidence of email. Second, estimates in columns (5) to (8) also show that local internet spillovers are mediated by the firm's own use of the internet, captured by the email and website adoption dummy variables. They suggest that if firms do not adopt internet technologies such as email or website, their revenue would suffer from the local diffusion of emails, probably due to an increased competition from proximate firms adopting emails and/or local economic transformations.

Therefore, unlike studies that do not find evidence of ICT spillovers in industrialized countries (Stiroh, 2002; Cardona, Kretschmer, & Strobel, 2013; Acharya, 2016), and as opposed to the study conducted by Marsh et al. (2017) which stresses negative (intra-industry) ICT spillovers on US companies' productivity, we highlight substantial positive spatial internet spillovers on manufacturing firms' performance in a large sample of developing and transition economies. However, we cannot yet determine the nature of these spatial spillovers, that is, whether they are driven by intra-industry or inter-industry diffusion of emails. The next sub-section addresses this question.

Table 3: Baseline estimations.

Dep. Var:	(1) Sales	(2) Sales	(3) Sales	(4) Sales / work.	(5) Sales	(6) Sales / work.	(7) Sales	(8) Sales / work.
(A) $\overline{Internet}_{t(i)t(i)}$	4.738** (2.113)	6.613*** (2.217)	3.620*** (0.712)	3.655*** (0.668)	-12.73** (5.523)	-10.76** (4.511)	-8.369*** (2.054)	-6.916*** (1.749)
(B) $\overline{Internet}_{t(i)t(i)}$ x email					40.76*** (13.67)	35.94*** (10.49)		
(C) $\overline{Internet}_{t(i)t(i)}$ x website							39.32*** (3.674)	34.68*** (2.819)
Email adopt		1.123*** (0.165)	0.647*** (0.0987)	0.357*** (0.0730)	-28.83** (11.52)	-25.64*** (9.010)	2.489*** (0.399)	1.982*** (0.394)
Website adopt		1.017*** (0.0724)	0.655*** (0.0562)	0.296*** (0.0336)	0.689*** (0.246)	0.326 (0.241)	-31.88*** (3.379)	-28.39*** (2.494)
Dist. to connect. infra (km, ln)		-0.00256 (0.0280)	-0.0563*** (0.0207)	-0.0405*** (0.0124)	-0.0733 (0.0607)	-0.0555 (0.0443)	-0.0723* (0.0435)	-0.0546* (0.0300)
Exports (% sales)			0.00626*** (0.0015)	0.00116 (0.0013)	0.0119*** (0.0024)	0.00617*** (0.00196)	0.0104*** (0.00148)	0.00486*** (0.00132)
Initial # perm. FT employees (ln)			0.653*** (0.0375)	0.0367 (0.0363)	0.827*** (0.0858)	0.190** (0.0950)	0.858*** (0.0828)	0.217** (0.0897)
% non-production workers			0.212 (0.154)	0.667*** (0.232)	0.434* (0.234)	0.864*** (0.248)	0.595*** (0.230)	1.005*** (0.275)
% skilled workers			-0.509*** (0.0452)	-0.188*** (0.0535)	-0.961*** (0.268)	-0.586*** (0.218)	-0.574*** (0.0885)	-0.245*** (0.0780)
Manager exp (years, ln)			0.0321 (0.0548)	0.0127 (0.0346)	-0.00900 (0.0748)	-0.0235 (0.0533)	-0.116* (0.0617)	-0.118** (0.0537)
Firm's age (years, ln)			0.192*** (0.0485)	0.0300 (0.0237)	0.353*** (0.0612)	0.172** (0.0744)	0.307*** (0.0790)	0.131** (0.0536)
% of foreign ownership			0.00774*** (0.00095)	0.00399*** (0.00075)	0.00838*** (0.00224)	0.00455** (0.00194)	0.00582* (0.00300)	0.00230 (0.00249)
% of state ownership			0.00516*** (0.00086)	0.000717 (0.00046)	0.00264 (0.0060)	-0.00150 (0.0054)	-0.00241 (0.00224)	-0.00596*** (0.00211)
Bank loan (0 or 1)			0.548*** (0.0482)	0.254*** (0.0466)	0.495*** (0.0963)	0.207* (0.107)	0.501*** (0.117)	0.212** (0.0863)
Electricity obstacle			0.0410 (0.0279)	0.0187 (0.0229)	0.183** (0.0828)	0.144** (0.0703)	0.0719 (0.0540)	0.0460 (0.0439)
<b>First stage estimates (A)</b>								
IV1	0.0437** (0.0195)	0.0794*** (0.0324)	0.1902*** (0.0581)	0.1902*** (0.0581)	0.1902*** (0.0581)	0.1902*** (0.0581)	0.1902*** (0.0581)	0.1902*** (0.0581)
IV2	-0.0020 (0.0014)	-0.0040 (0.0029)	-0.0137*** (0.0050)	-0.0137*** (0.0050)	-0.0137*** (0.0050)	-0.0137*** (0.0050)	-0.0137*** (0.0050)	-0.0137*** (0.0050)
(A) F-Stat	27.13	383.91	125.62	125.62	125.62	125.62	125.62	125.62
(B)-(C) F-Stat					13.40	13.40	136.92	136.92
Country-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cragg-Donald Wald F	1586.76	2523.38	4011.64	4011.64	18.179	18.179	20.748	20.748
LM stat	2.041	5.121*	8.595**	8.595**	0.953	0.953	1.474	1.474
Hansen J test (P-val.)	0.20	0.13	0.20	0.21	-	-	-	-
N	56,839	54,873	40,154	40,154	40,154	40,154	40,154	40,154

Note: \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%. Control estimates of the first-stage equation not reported. Standard errors are presented in parentheses, are robust to heteroscedasticity and clustered by country-fiscal year.



## 4.2 Intra-industry *versus* inter-industry spillovers

To study local internet spillover effects within and across industries, we decompose the location-level email incidence variable into:

- a variable measuring the spatial diffusion of email use among firms operating in the same industry as the firm considered, calculated as follows,

$$\overline{Internet\_intra}_{l(i)k(i)} = \frac{1}{N_{k,l} - 1} \left( \sum_{f \in k,l} Email_f - Email_i \right) \forall i \neq f, f \in K_l, l \in L \quad (5)$$

- a variable measuring the spatial diffusion of email use among the universe of firms operating in the same location but in the remaining set of industries, calculated as follows:

$$\overline{Internet\_inter}_{l(i)k(i)} = \frac{1}{N_{-k,l}} \left( \sum_{f \in -k,l} Email_f \right), \forall f \in -K_l, l \in L \quad (6)$$

Where  $N_{k,l}$  refers to the number of firms operating in industry  $k$  from location  $l$ ,  $N_{-k,l}$  represents the number of firms operating in other industries established in location  $l$ ,  $K_l$  represents the set of industries  $k$  established in location  $l$ , and  $L$  is the set of locations  $l$ . The fiscal-year subscript  $t$  is removed for the sake of simplification. The set of industries  $k$  is the result of stratification sectors grouping, detailed in Appendix B.4.<sup>15</sup>

We therefore conduct 2SLS estimations of the following system of equations:

$$Y_i = \gamma_0 + \gamma_1 \cdot \overline{Internet\_intra}_{l(i)k(i)} + \gamma_2 \cdot \overline{Internet\_inter}_{l(i)k(i)} + \gamma_3 \cdot \mathbf{X}_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + \epsilon_i \quad (7a)$$

$$\overline{Internet\_intra}_{l(i)k(i)t(i)} = \delta_0 + \delta_1 \cdot Z_{l(i)t(i)} + \delta_2 \cdot \mathbf{X}_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + \epsilon_{2l(i)k(i)t(i)} \quad (7b)$$

$$\overline{Internet\_inter}_{l(i)k(i)t(i)} = \zeta_0 + \zeta_1 \cdot Z_{l(i)t(i)} + \zeta_2 \cdot \mathbf{X}_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + \epsilon_{3l(i)k(i)t(i)} \quad (7c)$$

Where  $(\epsilon_{1i}, \epsilon_{2l(i)k(i)t(i)}, \epsilon_{3l(i)k(i)t(i)})$  is the error terms structure. The results reported in Table 4 (columns (1) and (2)) show opposite intra and inter-industry local spillover effects, depending on whether digital technology diffusion has taken place within or outside the industry where the firm operates. On the one hand, we find that the incidence of email use in the same industry has a positive and significant effect on the firm's sales and sales per worker. On the other hand, a larger local diffusion of email across other industries significantly deteriorates manufacturing firms' performance. To ensure that our findings are not biased by small sample size in specific locations-industries, we re-run estimations excluding locations-industries with less than 20 observations. The results, reported in columns (3) and (4), highlight robust, 1%-significant, but softer spillover effects. Therefore, in light of previous estimations, our analysis so far supports that positive intra-industry spillovers supersede negative inter-industry ones, but also indicates that the overall or net spillover effects could be much larger if the negative inter-industry spillovers were lower. In the next sub-sections, we provide additional insights into the mechanisms underlying such spillover effects.

Table 4: Inter- and intra-industry spillovers.

	(1)	(2)	(3)	(4)
Dep. Var:	Total sales	Sales per worker	Total sales	Sales per worker
	Baseline sample		Location-industry with $N_{k,l} \geq 20$	
(A) $\overline{Internet\_inter}$	-16.417*** (3.323)	-14.416*** (2.508)	-7.681*** (2.012)	-6.852*** (2.716)
(B) $\overline{Internet\_intra}$	23.799*** (3.313)	21.185*** (1.933)	8.362*** (1.037)	7.472*** (1.075)
	1st-stage estimates			
<b>Endog. var (A):</b>				
IV1	0.181*** (0.064)	0.181*** (0.064)	0.216** (0.092)	0.216** (0.092)
IV2 – Calibration: (t; t-4)	-0.0136** (0.005)	-0.0136** (0.005)	-0.017** (0.008)	-0.017** (0.008)
Weak-id SW F stat	22.18	22.18	73.33	73.33
Under-id SW Chi-2 stat.	22.68	22.68	74.95	74.95
<b>Endog. var (B):</b>				
IV1	0.128*** (0.032)	0.128*** (0.032)	0.206*** (0.071)	0.206*** (0.071)
IV2 – Calibration: (t; t-4)	-0.0077*** (0.0026)	-0.0077*** (0.0026)	-0.0107* (0.006)	-0.0107* (0.006)
Weak-id SW F stat	55.64	55.64	649.37	649.37
Under-id SW Chi-2 stat.	56.88	56.88	663.64	663.64
Control variables ( $X_i$ )	Yes	Yes	Yes	Yes
Control variables ( $X_j$ )	Yes	Yes	Yes	Yes
Fixed effects ( $d_{jt}, d_l, d_s$ )	Yes	Yes	Yes	Yes
N	39,673	39,673	25,504	25,504

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

### 4.3 Threshold spillover effects and firm's outward orientation

A first and obvious explanation for the existence of negative digital spillovers lies in the size of the internet user network. A critical mass of internet users might be necessary for network effects to become palpable (Grace, Kenny, & Qiang, 2003; Wu et al., 2021). By contrast, below a certain rate of technology diffusion within industries, first movers on the new technology may capture the market share of less productive competitors or use their market power to impose barriers to new entrants. Within locations, the introduction of the technology may increase the profitability of ICT-intensive industries at the expense of less ICT-intensive industries, with possible adverse consequences for local productive capacity (Choi et al., 2020). Moreover, Marsh et al. (2017) argue that it may take time for a technology to become widespread and fully exploited, which points to the possibility of threshold spillover effects induced by the delayed diffusion of digital technologies within locations. This possible delay in the transmission of positive local internet spillovers could occur within and across industries.

To investigate these non-linear spillover effects on manufacturing firm's performance, we introduce into the baseline equation (1) the squared term of the internet spillover variable<sup>16</sup>, yielding the following system of equations:

$$Y_i = \eta_0 + \eta_1 \cdot \overline{Internet}_{l(i)t(i)} + \eta_2 \cdot \overline{Internet}_{l(i)t(i)}^2 + \eta_3 \cdot \mathbf{X}_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + v_{1i} \quad (8a)$$

$$\overline{Internet}_{l(i)t(i)} = \beta_0 + \beta_1 \cdot Z_{l(i)t(i)} + \beta_2 \cdot \mathbf{X}_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + v_{2l(i)t(i)} \quad (8b)$$

$$\overline{Internet}_{l(i)t(i)}^2 = \theta_0 + \theta_1 \cdot Z_{l(i)t(i)} + \theta_2 \cdot \mathbf{X}_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + v_{3l(i)t(i)} \quad (8c)$$

With  $(v_{1i}, v_{2l(i)t(i)}, v_{3l(i)t(i)})$  the error terms structure.

2SLS estimates of equation (8), reported in Table 5, reveal non-linear internet spillover effects on manufacturing firms' sales and sales per worker. After testing for the presence of U-shaped spillover effects on sales, we reject, within a 1% confidence-level, the hypothesis of a monotone or inverse U-shaped relationship, and identify an email incidence cut-off equal to 0.53.

Second, we take the analysis in columns (3) to (6) further by replacing  $\overline{Internet}$  and  $\overline{Internet}^2$  with their corresponding inter-(intra-)industry email incidence variables in the set of endogenous regressors, controlling in parallel for the intra-(inter-)industry incidence. These additional estimations support a similar U-shaped inter-industry spillover effect on manufacturing firms' sales and sales per worker, with a close incidence cut-off value (0.45). Hence, an insufficient email user network size seems to lie behind negative (inter-industry) spillovers, suggesting that the internet penetration gap may widen the industrialisation gap between highly and poorly digitalised cities.<sup>17</sup> By contrast, intra-industry spillovers do not follow such a U-shaped curve, suggesting that knowledge and information spillovers, rather than network effects, could lie behind within-industry externalities.

However, insufficient network size does not preclude the capacity of a minority of high-performing firms displaying a higher capacity to absorb technological change to take advantage of the diffusion of ICTs across industries in a poorly digitised environment (Paunov & Rollo, 2015, 2016; Marsh et al., 2017). This absorptive capacity<sup>18</sup> is expected to be markedly higher in outward-oriented firms, defined as exporting, foreign or multi-plant firms (Glaeser et al., 1992; Paunov & Rollo, 2016; Farole & Winkler, 2012). Therefore, we interact  $\overline{Internet\_inter}$  with a dummy documenting the firm's inward-orientation:

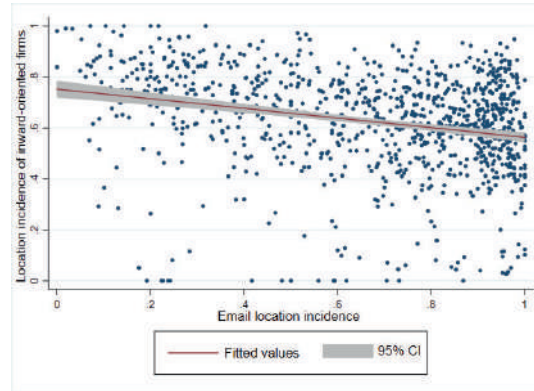
it is equal to one if, at the same time, the firm is single-plant, does not export, and has no foreign ownership.<sup>19</sup> Estimates presented in column (4) suggest that threshold negative (positive) inter-industry spillovers were actually driven by inward-(outward-)oriented firms. In fact, the diffusion of internet across industries is higher in locations where firms are more outward-oriented – the internet being indispensable to communicate with distant clients, suppliers, and other business partners –, with potentially a better chance of reaching the incidence threshold above which spillovers become positive. This interpretation is supported by Figure 6, plotting a negative (positive) correlation between the location-level incidence of inward- (outward-) oriented firms and the location-level email incidence, and additional estimations conducted in Online Appendix C.2.

Table 5: Threshold spillover effects.

Dep. Var:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total sales	Sales per worker	Total sales	Sales per worker	Total sales	Sales per worker	Total sales	Sales per worker
<b>Endogenous regressors:</b>								
(A) $\overline{Internet}$	-31.038*** (6.061)	-26.906*** (4.797)						
(B) $\overline{Internet}^2$	29.01*** (3.597)	25.581*** (2.604)						
(A) $\overline{Internet.intra}$			-26.050*** (3.479)	-23.018*** (2.712)			29.59*** (2.508)	26.75*** (2.701)
(B) $\overline{Internet.intra}^2$			28.770*** (2.565)	25.690*** (1.818)				
(A) $\overline{Internet.intra}$					188.30 (207.63)	168.30 (184.79)		
(B) $\overline{Internet.intra}^2$					-163.27 (196.45)	-146.01 (174.84)		
(C) $\overline{Internet} \times inward - or.$							-34.975*** (5.185)	-31.26*** (4.790)
<b>Additional controls:</b>								
$\overline{Internet.intra}$			0.821*** (0.310)	0.667** (0.272)				
$\overline{Internet.intra}$					-6.456 (6.807)	-5.508 (6.034)		
$Inward - orientation$							23.087*** (3.706)	20.915*** (3.337)
<b>1st-stage statistics</b>								
<b>Endog. var (A):</b>								
Weak-id SW F stat	58.01	58.01	81.91	81.91	3.75	3.75	86.75	86.75
Under-id SW Chi-2 stat.	59.30	59.30	83.75	83.75	3.83	3.83	88.70	88.70
<b>Endog. var (B):</b>								
Weak-id SW F stat	95.38	95.38	151.31	151.31	3.50	3.50	62.75	62.75
Under-id SW Chi-2 stat.	97.50	97.50	154.70	154.70	3.58	3.58	64.16	64.16
Lind & Melhum U shape T-test	5.12***	7.49***	7.49***	8.49***	0.75	0.75	-	-
Control variables ( $X_i$ )	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects ( $d_{jt}, d_t, d_s$ )	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	40,154	40,154	39,673	39,673	39,673	39,673	39,586	39,586

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-fiscal year. U-shape t-test tests the null that the relationship is monotone or inverse U-shaped.

Figure 6: Firms' inward orientation and email incidence in locations.



Note: 972 pooled observations (725 locations, 136 countries).

#### 4.4 Spatial internet spillovers on other manufacturing outcomes

In this subsection, we further the analysis by studying the net, inter-industry and intra-industry spillover effects on the firm's workforce size and output diversification. The workforce size is measured the number of full-time permanent and temporary workers (separately, expressed in logarithm), while output diversification is measured by the share of the main product in the firm's total annual sales (the lower the share, the more diversified the firm). In employment equations, we add the firm's total annual sales to the set of control variables  $X_{i,t}$ .

Results are reported in Table 6. Regarding net internet spillovers, we provide evidence of negative internet spillover effects on the number of full-time permanent workers (column (1)) and positive ones on firm's diversification (column (7)). According to our estimates, a 10 percentage point increase in location-level internet incidence leads to a 10% decrease in the number of permanent workers in manufactures, and to a 2.4-percentage point decrease in the share of their main output in total sales. These findings hide however conflicting inter-industry and intra-industry spillover effects. Indeed, estimates put in evidence positive inter-industry spillovers on both permanent and temporary workforces and on firm's output diversification, while negative intra-industry spillovers on these same three outcomes. Therefore, if previous evidence pointed to positive intra-industry spillovers on manufacturing firm's sales and sales per worker, this has been accompanied by a contraction in the workforce size and a greater output specialisation, thereby supporting Marshall's hypothesis of localisation externalities.<sup>20</sup> By contrast, inter-industry diffusion of internet, despite being detrimental to sales and labour productivity, is associated with greater output diversification and job creation, which comes in support of Jacob's hypothesis of diversification externalities.

Table 6: Spatial internet spillovers on other manufacturing outcomes.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep var:	FT permanent employees (ln)		FT temp. employee (ln)		% of main product in total sales	
(A) <i>Internet</i>	-1.019*** (0.293)		0.129 (0.476)		-24.07** (11.83)	
(B) <i>Internet_inter</i>		3.122*** (0.806)		11.32*** (3.406)		-291.2*** (50.62)
(C) <i>Internet_intra</i>		-4.708*** (0.754)		-12.55*** (2.862)		298.2*** (35.38)
First stage statistics						
Weak-id SW F-stat :						
(A)	124.58		125.72		125.38	
(B)		16.39		16.06		22.32
(C)		33.90		33.12		52.12
Under-id Chi-2 stat:						
(A)	254.70		257.12		256.35	
(B)		16.76		16.43		22.79
(C)		34.66		33.87		53.22
Controls + FEs	Yes	Yes	Yes	Yes	Yes	Yes
Total annual sales	Yes	Yes	Yes	Yes	No	No
Hansen test p-value	0.23	-	0.14	-	0.21	-
<i>N</i>	40,154	39,670	39,280	38,803	39,899	43,709

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

## 5 Robustness checks

In this section, we conduct a series of robustness checks which consist in ensuring that our findings are not undermined by omitted variable bias or location selection bias. All robustness checks' estimates are reported in Online Appendix B.

First, one possible caveat in our identification strategy is that unobserved characteristics of the firm's proximate environment could be directly conducive to firm's performance. Although including country-year and location fixed-effects, and controlling for the (time-varying) location's distance to connectivity infrastructure mitigates this concern, other time-varying local factors correlated with both local internet access and firm's performance, such as the business environment, could bias our result. To address this issue, we include in the set of control variables a proxy of local economic activity, orthogonal to local internet diffusion and the firm's own activity. This proxy is the residual of an OLS (or IV 2SLS) regression of the average location sales over firm's sales, location email diffusion, as well as location, country-year and sector fixed-effects:

$$\bar{Y}_{l(i)t(i)} = \alpha + \beta \cdot \overline{Internet}_{l(i)t(i)} + \gamma \cdot Y_i + d_{j(i)t(i)} + d_{l(i)} + d_{s(i)} + \varepsilon_{Y_i}^{OLS-IV} \quad (9)$$

Where  $\bar{Y}_{l(i)t(i)}$  is the average annual sales of other firms located in location  $l$ , that is, excluding firm  $i$ 's own sales, and  $\varepsilon_{Y_i}^{OLS-IV}$  the residual from an OLS or IV-2SLS estimation of Equation (9). Equations (1a) and (1b) are then re-estimated using the 2SLS-FE estimator and including the residual  $\varepsilon_{Y_i}^{OLS-IV}$  as



an additional control variable. Results, reported in Online Appendix B.1, suggest that the strength of internet spillovers was slightly overstated in baseline estimations due to indirect effects of local economic activity, but still, remain positive and significant at 1%.

Second, we follow Borusyak et al. and perform pre-trend tests by regressing the three-year lagged total sales (per worker) over  $\overline{Internet}_{l(i)t(i)}$ , instrumented by our baseline instrument set. Results are reported in column (3) of Online Appendix B.1 and do not provide evidence of an effect of internet spillovers on the lagged dependent variable. We also proceed to a reduced-form OLS estimation of the lagged dependent variable on instrumental variables IV1 and IV2, and do not find any significant relationship between IVs and the lagged dependent variable (column (4)). Moreover, to address the omitted variable bias that could arise from the non-random exposure to connectivity shocks (Borusyak & Hull, 2020), we perform falsification tests by regressing IV1 and IV2 on their “expected” value (including location and country-year FEs), and using the resulting residuals as “recentered” instruments.<sup>21</sup> Estimates are reported in column (4) and are robust and consistent with previous baseline estimations.

Third, we augment our baseline econometric specification by controlling for additional sources of omitted variable bias. In fact, there could be linkages between specific industries, channelling or mediating spillover effects, that our model does not control for. We therefore include in our specification industry-pair dummy variables, equal to one if the firm operates in a given pair of industries, zero otherwise. Moreover, we also replace stratification sector FEs by dummies indicating the four-digit ISIC code of the firm’s main product, to make sure that unobserved characteristics of the firm’s output do not alter estimated relationships. Estimates, reported in Online Appendix B.2, support the robustness of our results.

Fourth, we exclude landlocked countries from the sample because they cannot host SMCs, and can only be indirectly connected to them via the terrestrial cable network. Although the presence of fixed-effects controls for the consequences of this geographical feature on telecommunication outcomes, landlocked countries are particularly dependent on cross-border connectivity with neighbouring coastal countries hosting SMCs. The apparent absence of cable could therefore be associated with heterogeneous outcomes for these countries, because of missing information on terrestrial infrastructure deployment, cross-border connectivity, and cable-related internet disruptions. This restriction leads us to exclude 29 countries, corresponding to 5,444 manufacturing firms. These ultimate checks, reported in Online Appendix B.3, confirm the robustness of our findings.

Fifth, even though location FEs reduce the concern for location selection bias (Gaubert, 2018; Duranton & Puga, 2020), we rerun estimations excluding from the sample large, multi-plant, foreign and exporting firms, which are more likely to be mobile and to choose locations close to connectivity infrastructures (Dollar et al., 2005; Dollar, Hallward-Driemeier, & Mengistae, 2006). In addition, we make sure that the estimated relationships are not affected by a low number of firms in some locations, by imposing a minimum of 50 firms in locations. In fact, locations that are far from infrastructures may host a smaller number of firms, among which less performing firms may be over-represented. Results, reported in Online Appendixes B.4 and B.5, show that our findings remain robust to these ultimate sample restrictions.

## 6 Conclusion

Internet technologies play a critical role in environments where firms are strongly constrained in their daily interactions by large transactions costs, missing infrastructures and high levels of uncertainty. The adoption and diffusion of new technologies are likely to deliver net benefits, but these may not always be guaranteed due to limited digital absorptive capacity of some firms and industries, and to thresholds in network effects.

Focusing on manufacturing firms in a large sample of developing and transition economies, this paper shows that beyond the direct benefits of internet adoption by manufacturers, local diffusion of the email technology has yielded strong and positive spatial spillover effects on manufacturing firms' performance. In fact, IV estimations document a positive effect of an increased incidence of email use within locations on manufactures' revenues and labour productivity. We also find that these higher performance are accompanied by output diversification, driven by inter-industry spillovers, and workforce contraction, driven by intra-industry ones. However, these spillovers are subject to important threshold effects arising from the cross-industrial dimension: while the diffusion of the internet within an industry generates positive spillovers on firm performance whatever the incidence rate, the network of email users has to reach a critical size to produce positive externalities across industries. We indeed find evidence of negative spatial spillover effects from the diffusion of the internet across industries on manufacture sales and productivity below an email incidence threshold of approximately 50% of the local universe of firms. Last, these threshold effects seem related to the presence of outward-oriented firms, which tend to operate in places where internet is more diffused and demonstrate a greater absorptive capacity.

Overall, our empirical analysis, in line with recent findings on digitalisation in developing countries (Hjort & Poulsen, 2019; IMF, 2020), supports the idea that net spatial internet spillovers on developing countries' manufacturing sector are strong and positive. But evidence of U-shaped local spillover effects suggest that industrialisation paces may diverge between poorly and highly digitalised environments.

## **Acknowledgment**

The authors are grateful to Olivier Santoni for providing excellent research assistance. The authors also wish to thank Tanguy Bernard, Simone Bertoli, Vianney Dequiedt, Michaël Goujon, Michele Imbruno, Luc Jacolin, Rachid Laajaj, Alexis Le Nestour, Florian Léon, Mathilde Maurel, Jaime de Melo, Mark Plant, Eric Rougier, Petros Sekeris, Jules Tapsoba, and Martha Tesfaye for their useful suggestions, as well as participants at the CGDEV, Ferdi-Banque de France, and GRETHA research seminars, at the 2017 DIAL, the 2017 AFSE, the 2018 DEGIT XXIII, and the 2019 Royal Economic Society conferences, for their comments on a earlier but related paper.

## **Funding**

This work was supported by the Agence Nationale de la Recherche of the French government through the programme ‘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).

## References

- Aceto, G., Botta, A., Marchetta, P., Persico, V., & Pescapé, A. (2018). A comprehensive survey on internet outages. *Journal of Network and Computer Applications*, *113*, 36–63.
- Acharya, R. C. (2016). Ict use and total factor productivity growth: intangible capital or productive externalities? *Oxford Economic Papers*, *68*(1), 16–39.
- Acs, Z. J., Audretsch, D. B., & Feldman, M. P. (1994). R & d spillovers and recipient firm size. *The review of Economics and Statistics*, 336–340.
- Aker, J. (2017). Using digital technology for public service provision in developing countries potential and pitfalls. *Digital Revolutions in Public Finance*.
- Björkegren, D. (2019). The adoption of network goods: Evidence from the spread of mobile phones in rwanda. *The Review of Economic Studies*, *86*(3), 1033–1060.
- Borusyak, K., & Hull, P. (2020). *Non-random exposure to exogenous shocks: Theory and applications* (Tech. Rep.). National Bureau of Economic Research.
- Borusyak, K., Hull, P., & Jaravel, X. (2021). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, Forthcoming.
- Bresnahan, T., & Trajtenberg, M. (1995). General purpose technologies ‘engines of growth’? *Journal of Econometrics*, *65*(1), 83–108.
- Caragliu, A., de Dominicis, L., & de Groot, H. L. (2016). Both marshall and jacobs were right! *Economic Geography*, *92*(1), 87–111.
- Cardona, M., Kretschmer, T., & Strobel, T. (2013). Ict and productivity: conclusions from the empirical literature. *Information Economics and Policy*, *25*(3), 109–125.
- Cariolle, J. (2021). International connectivity and the digital divide in sub-saharan africa. *Information Economics and Policy*, *55*, 100901.
- Cariolle, J., Le Goff, M., & Santoni, O. (2019). Digital vulnerability and performance of firms in developing countries. *Banque de France Working Paper 709*.
- Carter, L. (2014). Submarine cables and natural hazards. In D. Burnett, R. Beckman, & T. Davenport (Eds.), *Submarine cables: The handbook of law and policy* (pp. 237–254). Martinus Nijhoff Publishers.
- Carter, L., Burnett, D., Drew, S., Marle, G., Hagadorn, L., Bartlett-McNeil, D., & Irvine, N. (2010). *Submarine cables and the oceans: connecting the world* (No. 31). UNEP/Earthprint.
- Chemin, M. (2020). Judicial efficiency and firm productivity: Evidence from a world database of judicial reforms. *Review of Economics and Statistics*, *102*(1), 49–64.
- Choi, J., Dutz, M. A., & Usman, Z. (2020). *The future of work in africa: Harnessing the potential of digital technologies for all*. World Bank Publications.
- Cingano, F., & Schivardi, F. (2004). Identifying the sources of local productivity growth. *Journal of the European Economic association*, *2*(4), 720–742.
- Crémer, J., Rey, P., & Tirole, J. (2000). Connectivity in the commercial internet. *The Journal of Industrial Economics*, *48*(4), 433–472.
- Diao, X., McMillan, M., & Rodrik, D. (2019). The recent growth boom in developing economies: A structural-change perspective. In *The palgrave handbook of development economics* (pp. 281–334).

Springer.

- Dollar, D., Hallward-Driemeier, M., & Mengistae, T. (2005). Investment climate and firm performance in developing economies. Economic Development and Cultural Change, *54*(1), 1–31.
- Dollar, D., Hallward-Driemeier, M., & Mengistae, T. (2006). Investment climate and international integration. World Development, *34*(9), 1498–1516.
- Duranton, G., & Puga, D. (2020). The economics of urban density. Journal of Economic Perspectives, *34*(3), 3–26.
- Eichengreen, B., Lafarguette, R., & Mehl, A. (2016). Cables, sharks and servers: Technology and the geography of the foreign exchange market (Tech. Rep.). National Bureau of Economic Research.
- Fabiani, S., Schivardi, F., & Trento, S. (2005). Ict adoption in italian manufacturing: firm-level evidence. Industrial and Corporate Change, *14*(2), 225–249.
- Farole, T., & Winkler, D. E. (2012). Foreign firm characteristics, absorptive capacity and the institutional framework: the role of mediating factors for fdi spillovers in low-and middle-income countries. World Bank policy research working paper(6265).
- Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. Regional Studies, *41*(5), 685–697.
- Freund, C., & Weinhold, D. (2002). The internet and international trade in services. American Economic Review, *92*(2), 236–240.
- Gaspar, J., & Glaeser, E. L. (1998). Information technology and the future of cities. Journal of Urban Economics, *43*(1), 136–156.
- Gaubert, C. (2018). Firm sorting and agglomeration. American Economic Review, *108*(11), 3117–53.
- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. Journal of Political Economy, *100*(6), 1126–1152.
- Goldfarb, A., & Tucker, C. (2019). Digital economics. Journal of Economic Literature, *57*(1), 3–43.
- Görg, H., & Greenaway, D. (2004). Much ado about nothing? do domestic firms really benefit from foreign direct investment? The World Bank Research Observer, *19*(2), 171–197.
- 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 Management, *12*(2), 48–63.
- Grace, J., Kenny, C., & Qiang, C. Z.-W. (2003). Information and communication technologies and broad-based development: A partial review of the evidence. The World Bank.
- Grubestic, T. H., O’Kelly, M. E., & Murray, A. T. (2003). A geographic perspective on commercial internet survivability. Telematics & Informatics, *20*(1), 51–69.
- Harrison, B., Kelley, M., & Gant, J. (1996). Innovative firm behavior and local milieu: exploring the intersection of agglomeration, firm effects, and technological change. Economic Geography, *72*(3), 233–258.
- Haskel, J. E., Pereira, S. C., & Slaughter, M. J. (2007). Does inward foreign direct investment boost the productivity of domestic firms? The review of economics and statistics, *89*(3), 482–496.
- Henderson, J. V. (2003). Marshall’s scale economies. Journal of Urban Economics, *53*(1), 1–28.

- Hjort, J., & Poulsen, J. (2019). The arrival of fast internet and employment in africa. American Economic Review, *109*(3), 1032–79.
- Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. Econometrica, *62*(2), 467–475.
- IMF. (2020). Regional economic outlook, october 2020, sub-saharan africa : A difficult road to recovery. USA: International Monetary Fund.
- Innovative business models for expanding fiber-optic networks and closing the access gaps (Tech. Rep.). (2018). World Bank Group.
- Jacobs, J. (1969). The economy of cities. Vintage.
- Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. American Economic Review, *75*(3), 424–440.
- Kneller, R., & Timmis, J. (2016). Ict and exporting: The effects of broadband on the extensive margin of business service exports. Review of International Economics, *24*(4), 757–796.
- Léon, F. (2020). The elusive quest for high-growth firms in africa: when other metrics of performance say nothing. Small Business Economics, 1–22.
- Lind, J. T., & Mehlum, H. (2010). With or without u? the appropriate test for a u-shaped relationship. Oxford Bulletin of Economics and Statistics, *72*(1), 109–118.
- Lucas Jr, R. E. (1988). On the mechanics of economic development. Journal of Monetary Economics, *22*(1), 3–42.
- Malecki, E. J. (2002). The economic geography of the internet's infrastructure. Economic Geography, *78*(4), 399–424.
- Marsh, I. W., Rincon-Aznar, A., Vecchi, M., & Venturini, F. (2017). We see ict spillovers everywhere but in the econometric evidence: a reassessment. Industrial and Corporate Change, *26*(6), 1067–1088.
- Marshall, A. (1890). Principles of economics. London: Macmillan.
- Martin, P., Mayer, T., & Mayneris, F. (2011). Spatial concentration and plant-level productivity in france. Journal of Urban Economics, *69*(2), 182–195.
- McMillan, M., Rodrik, D., & Sepulveda, C. P. (2016). The recent growth boom in developing economies: A structural-change perspective. In Structural change, fundamentals, and growth : a framework and case studies (pp. 281–334). Washington, DC : International Food Policy Research Institute,.
- OECD. (2014). International cables, gateways, backhaul and international exchange points. (Tech. Rep.). Organisation for Economic Cooperation and Development.
- Palmer-Felgate, A., Irvine, N., Ratcliffe, S., & Bah, S. S. (2013). Marine maintenance in the zones: A global comparison of repair commencement times. In Suboptic conference from ocean to cloud (pp. 22–25).
- 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.

- Rodrik, D. (2016a). An african growth miracle? Journal of African Economies, 27(1), 10–27.
- Rodrik, D. (2016b). Premature deindustrialization. Journal of Economic Growth, 21(1), 1–33.
- Rodrik, D. (2018). New technologies, global value chains, and developing economies (Tech. Rep.). National Bureau of Economic Research.
- Romer, P. M. (1986). Increasing returns and long-run growth. Journal of political economy, 94(5), 1002–1037.
- Sadowski, B. M., Maitland, C., & van Dongen, J. (2002). Strategic use of the internet by small-and medium-sized companies: an exploratory study. Information Economics and Policy, 14(1), 75–93.
- Simione, F., & Li, Y. (2021). The macroeconomic impacts of digitalization in sub-saharan africa: Evidence from submarine cables. IMF Working Papers, 2021(110).
- Stiroh, K. J. (2002). Are ict spillovers driving the new economy? Review of Income and Wealth, 48(1), 33–57.
- Storper, M., & Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. Journal of Economic Geography, 4(4), 351–370.
- Tybout, J. R. (2000). Manufacturing firms in developing countries: How well do they do, and why? Journal of Economic Literature, 38(1), 11–44.
- Weller, D., & Woodcock, B. (2013). Internet traffic exchange: Market developments and policy challenges.
- Wixe, S. (2015). The impact of spatial externalities: Skills, education and plant productivity. Regional Studies, 49(12), 2053–2069.
- Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.
- Wu, S., Wang, P., & Sun, B. (2021). Can the internet narrow regional economic disparities? Regional Studies, 1–14.



## Notes

<sup>1</sup>Goldfarb and Tucker (2019) define digital technologies, of which internet technologies are part, as “the representation of information in bits [...] rather than atoms”, which “reduces the cost of storage, computation and transmission of data” (p.3).

<sup>2</sup>In this paper, digitalisation refers to the growing use of digital technologies in socio-economic interactions.

<sup>3</sup>Paunov and Rollo (2015) address this issue, but in a very succinct way. They identify larger returns to email diffusion within industries for service firms than for those operating in the manufacturing sector.

<sup>4</sup>For the sake of our analysis clarity, we keep this dichotomy between intra and inter-industry spillovers in its original expression (Glaeser et al., 1992), having in mind the recent advances in the literature that put forward a diversification effect (induced by related local production variety) underlying intra-industry spillovers, and a portfolio effect (induced by unrelated local production variety) underlying inter-industry spillovers (Caragliu et al., 2016; Frenken et al., 2007).

<sup>5</sup>To our knowledge, empirical research on inter-industry spillovers from ICTs’ diffusion is rather scarce and mostly focused on industrialized economies (Harrisson et al., 1996; Marsh et al, 2017).

<sup>6</sup>In the WBES, the fiscal year corresponding to reported sales and firms’ characteristics diverges from the year of the survey round. See Appendix B.1

<sup>7</sup>This control is an ordered categorical variable: firms were asked whether access to electricity is i) not an obstacle, ii) a minor obstacle, iii) a moderate obstacle, iv) a major obstacle, or v) a very severe obstacle to their operations

<sup>8</sup>The nested nature of SMC deployment and exposure to shocks finds support in Appendix B.5, which shows a positive correlation between SMC number (V3 and V6) and outage days experience (V4 and V5) on the one hand, and a negative correlation between outage days experience (V4 and V5) and email adoption (V1), website adoption (V7), and email spillovers (V2) on the other hand. Moreover, as instruments are nested, they should respect the monotonicity condition necessary for 2SLS estimators with multiple IVs to be valid (Imbens & Angrist, 1994). In coastal countries, SMC laying is indeed a precondition to be exposed to SMC faults, in the same way as being affected by internet disruptions requires first having access to internet. This feature makes the hypothesis of homogeneous responses to IV1 and IV2 highly plausible. There are however in the baseline sample landlocked countries (Burundi, Zimbabwe, Ethiopia, Uzbekistan, and Armenia) with a record of internet disruptions caused by cable faults in their neighbouring coastal countries. IV estimations excluding landlocked countries are therefore conducted in the robustness section (Online Appendix B.3), and yield to first and second-stage estimates consistent with those obtained from the baseline sample.

<sup>9</sup>The International Cable Protection Committee Ltd (ICPC) is an organisation that illustrates the international concern for cable outages. It brings together more than 100 members representing telecommunication companies, government agencies and scientific organisations to protect cables and support the development, security, and continuity of international telecommunications.

<sup>10</sup>Note that regardless of whether firms have access to the internet via a computer or a mobile device, their connectivity will be similarly affected by SMCs-related internet disruptions.

<sup>11</sup>Primary raw data on outages’ time, duration and causes have been drawn and treated from the Subtel forum: <http://subtelforum.com/ca-faults-maintenance/>. They have been completed with information drawn from Akamai reports on the “State of Internet Connectivity” covering the 2008-2017 period, as well as manual internet searches.

<sup>12</sup>Estimations conducted without these two exclusion restrictions, reported in Online Appendix A.2, column(7), are consistent with baseline estimations.

<sup>13</sup>A distance raster map is defined from all coordinate points, which gives the distance of each firm’s location centroid to the nearest connectivity infrastructure. When the geographical stratification unit is the region or a group of cities, which sometimes happens in the data-set, we take the region’s or city group’s centroid as geographical coordinate. If a country does not host any SMC or IXP, the distance is calculated with respect to the closest infrastructure in neighbouring countries.

<sup>14</sup>Beyond the benefits of having multiple instruments for the analysis of inter- and intra-industry spillovers, estimations resulting from various IV calibrations, reported in Online Appendix A.2, confirm the nested nature of IV1 and IV2. First, using both instruments together reduces the risk of under-identification. Second, while first-stage estimations stress that the outage-based instrument (IV2) used as single instrument positively affects email location incidence, this effect has the expected negative sign after including the SMC-based instrument (IV1), or, after excluding from the sample Eastern and South-Eastern Asian countries which present the characteristics of having a large number of cables and being frequently

exposed to SMCs' outages. This indicates that a greater experience of SMCs' outages also results from a higher number of SMCs (as also shown by the correlation matrix in Appendix B.5 and the correlation graph in Online Appendix B.1). Third, reduced-form estimates reported in columns (8) to (10) of Online appendix A.2 confirm previous interpretation by showing a positive effect of IV1 or IV2 on firm's sales, supporting that IV2 carries information on both SMC capacity and vulnerability to shocks. Fourth, following Borusyak et al.(2021), Online Appendix A.3 reports IV estimates when IVs' components – i.e. aggregate connectivity shocks or distance to infrastructures – are separately fixed, that is, set to their value at the fiscal year of the first survey round. The resulting estimates and associated first-stage statistics support that identification results from variations in the aggregate component (i.e. SMC number and outage days) rather than variations in the weight component (i.e. distance to infrastructure) of IVs.

<sup>15</sup>This grouping of sectors into industries has been made to avoid location-level sample attrition that would have occurred if incidence variable had been computed using stratification sectors. Estimations based on an alternative sector grouping procedure (explained in Appendix B.4) are provided in Online Appendix C.1. Results thus obtained are consistent with those presented in this section.

<sup>16</sup>Following Wooldridge (2010, Section 9.5).

<sup>17</sup>As recently evidenced in a sample of Chinese cities by Wu et al. (2021).

<sup>18</sup>Defined as a firm's "capacity to assimilate the technological knowledge created outside the firm and to apply it within its production process" (Marsh et al., 2017, p.1068).

<sup>19</sup>Estimations using separately these characteristics of firms as interaction variables are reported in Online Appendix C.2. They show that inter-industry spillovers are positive when firms exports, are foreign owned, or are multi-plants.

<sup>20</sup>We also look at the spillover effects on workforce composition in Online Appendix C.3 and find evidence of negative internet spillovers on the number of permanent production and non production workers of similar size.

<sup>21</sup>Expected IV1 and IV2 are constructed by respectively drawing 2000 counter-factual connectivity shocks – i.e. shocks on SMCs' number or outage days – from a normal distribution with mean and standard deviation equal to the corresponding moments of these variables' historic distribution, and weighting them by the location distance to infrastructures. Expected instruments' value is set to zero in countries-year without SMCs. Following the same procedure, we also apply a Poisson distribution when constructing expected IV2, which applies to rare event such as plane accidents, and get robust results (not reported but can be provided upon request).

## Appendix

### Appendix A. Variables' sources and definitions.

Variables	Description	Sources
Total sales (USD, ln)	Total annual sales made during the last fiscal year in real USD	WBES
Real sales per worker (USD, ln)	Total sales weighted by the # of FT perm. workers adjusted for temp. workers	WBES
Email adoption (0,1)	Does the firm use emails to communicate with clients or customers (1=yes,0=no)	WBES
Website adoption (0,1)	Does the firm have its own website? (1=yes, 0=no)	WBES
% of state ownership	Percent of the firm owned by the government/State	WBES
Distance to connectivity infra (km, ln)	Distance to the closest submarine cable's landing stations or Internet Exchange Point	Telegeography, Packet Clearing House and PeeringDB
% of foreign ownership	Share of the firm owned by private foreign entities	WBES
% dir. indir. exports in sales	Share of sales that are exported directly or indirectly	WBES
Firm's age (years, ln)	Number of years since the creation of the firm	WBES
Initial # perm. FT employees(ln)	Number of full-time permanent employees when the started its operations	WBES
% Non-production workers	Share of non-production workers in total full-time permanent workers	WBES
% Skilled production workers	Share of skilled workers among total production workers	WBES
Manager experience (years, ln)	Number of years of experience the top manager has in the firm's sector	WBES
Bank loan (0,1)	Equal 1 when the firm has subscribed a credit or a loan from a financial institution, 0 otherwise	WBES
Electricity obstacle (ordered, 0 to 4)	To what extent is electricity an obstacle for the activity of the firm?	WBES

## Appendix B. Sample statistics

### B.1. Sample statistics, by fiscal year.

Fiscal year	Freq.	Percent	Fiscal year	Freq.	Percent
2005	5,447	13.57	2012	9,823	24.46
2006	2,871	7.15	2013	3,261	8.12
2007	3,208	7.99	2014	508	1.27
2008	1,749	4.36	2015	949	2.36
2009	4,480	11.16	2016	1,892	4.71
2010	1,609	4.01	2017	56	0.14
2011	4,301	10.71	<b>Total</b>	<b>40,154</b>	<b>100</b>

### B.2. Sample statistics, by stratification sectors.

Sector	Freq.	Percent	Sector	Freq.	Percent
Agriculture	3	0.01	Manuf. other non-metallic min..	2,670	6.65
Chemicals Chemical Prod.	3,411	8.49	Manuf. other transport equipm.	107	0.27
Construction	3	0.01	Manuf. paper prod.	513	1.28
Hotels restaurants	1	0.00	Manuf. petroleum prod.	89	0.22
Manuf. basic metals	1,159	2.89	Manuf. rubber & plastics	2,494	6.21
Manuf. electrical machinery	1,323	3.29	Manuf. textiles	2,708	6.74
Manuf. electronic equipment	229	0.57	Manuf. tobacco prod.	161	0.40
Manuf. fabricated metal prod.	3,349	8.34	Manuf. wearing apparel	5,043	12.56
Manuf. food prod.	7,999	19.92	Manuf. wood prod.	1,120	2.79
Manuf. furniture	2,048	5.10	Mining	2	0.00
Manuf. leather prod.	824	2.05	Other Manuf.	65	0.16
Manuf. machinery & equipment	2,342	5.83	Other business activities	2	0.00
Manuf. medical instruments	211	0.53	Publishing and printing	1,320	3.29
Manuf. motor vehicles	852	2.12	Recycling	91	0.23
Manuf. office	15	0.04	<b>Total</b>	<b>40,154</b>	<b>100</b>

B.3. Sample statistics, by country.

iso	Email incidence	Std. Dev.	Freq.	iso	Email incidence	Std. Dev.	Freq.	iso	Email incidence	Std. Dev.	Freq.
AFG	0.44	0.50	144	GTM	0.81	0.39	619	NGA	0.22	0.42	539
AGO	0.21	0.41	305	HND	0.67	0.47	395	NIC	0.53	0.50	495
ALB	0.71	0.46	136	HRV	0.91	0.29	371	NPL	0.57	0.50	355
ARG	0.98	0.15	1509	IDN	0.39	0.49	1821	PAK	0.44	0.50	437
ARM	0.70	0.46	107	IND	0.87	0.34	6241	PER	0.94	0.24	1343
AZE	0.41	0.49	121	IRQ	0.19	0.39	464	PHL	0.84	0.37	1231
BDI	0.47	0.50	154	JAM	0.88	0.33	57	PNG	0.95	0.21	22
BEN	0.70	0.46	50	JOR	0.73	0.45	197	PRY	0.82	0.38	417
BFA	0.74	0.44	70	KAZ	0.82	0.39	249	RUS	0.96	0.21	1227
BGD	0.45	0.50	1062	KEN	0.87	0.34	251	RWA	0.44	0.50	55
BGR	0.88	0.32	678	KGZ	0.62	0.49	152	SEN	0.38	0.49	414
BIH	0.93	0.26	177	KHM	0.52	0.50	95	SLE	0.23	0.42	74
BLR	0.85	0.36	142	LAO	0.23	0.42	53	SRB	0.95	0.22	155
BOL	0.85	0.36	420	LBN	0.80	0.40	159	SSD	0.32	0.47	79
BRA	0.95	0.21	781	LBR	0.30	0.46	69	SWZ	0.73	0.45	114
BWA	0.48	0.51	33	LKA	0.35	0.48	295	TCD	0.23	0.43	56
CHN	0.90	0.30	1454	LSO	0.67	0.47	64	THA	0.76	0.43	486
CIV	0.39	0.49	198	MAR	0.95	0.22	80	TJK	0.48	0.50	153
CMR	0.61	0.49	142	MDA	0.71	0.45	119	TTO	0.90	0.30	101
COD	0.27	0.45	443	MEX	0.77	0.42	1791	TUN	0.97	0.18	264
COL	0.94	0.24	1552	MKD	0.85	0.36	187	TUR	0.91	0.28	106
CRI	0.96	0.19	223	MLI	0.30	0.46	427	TZA	0.40	0.49	257
DJI	0.77	0.43	22	MMR	0.30	0.46	620	UGA	0.34	0.47	494
DOM	0.91	0.29	90	MNE	0.75	0.44	48	UKR	0.77	0.42	697
ECU	0.96	0.21	451	MNG	0.55	0.50	100	URY	0.84	0.37	464
ETH	0.64	0.48	417	MOZ	0.22	0.42	340	VEN	0.93	0.25	60
GEO	0.61	0.49	147	MRT	0.60	0.49	110	YEM	0.29	0.45	237
GHA	0.41	0.49	540	MWI	0.25	0.46	8	ZAF	0.79	0.41	524
GIN	0.22	0.41	139	MYS	0.77	0.42	452	ZMB	0.58	0.49	553
GMB	0.30	0.47	33	NAM	0.76	0.43	152				
GNB	0.47	0.50	45	NER	0.45	0.51	20	Total	0.72	0.45	40154

#### B.4. Correspondence between industries and sector stratification.

Industry	Sample		Sector strat.	Sample	
	Total	Estimation		Total	Estimation
Construction	7,580 (16)	3	Construction	7,580 (16)	3
Manuf of machinery, equip., & motor vehicles	21,407	8,072	Manuf. of electrical machinery	2,360	1,323
			Manuf of electronic equipment	444	229
			Manuf of medical instruments	540	211
			<i>Manuf of motor vehicles</i> †	1,392	852
			Manuf of other transport equipment	365	107
			Manuf of paper prod.	1,162	513
			Manuf of machinery & equip.	4,951	2,342
			<i>Manufacture of rubber &amp; plastics prod.*</i>	5,016	2,494
			<i>Sale-maintenance-repair of motor vehicles</i> †	5,177 (19)	1
			Other Manufacturing	1,696	331
Other services	4,455 (14)	0	Recycling	260	91
			Manuf. of tobacco prod.	267	161
			Manuf. of office	57	15
			Other services	4,374	0
			Real estate activities	57	0
			Recreational, cultural & sporting activities	12	0
Textiles Garments	18,123	8,575	Electricity, gas, steam & hot water supply	12	0
			Manuf. of textiles	6,021	2,708
			Manuf. of leather products	1,946	824
Manufacture of metals	9,535	4,508	Manuf. of wearing apparel	10,156	5,043
			Manuf of basic metals	2,135	1,159
			Manuf of fabricated metal prod.	7,400	3,349
Manufacture of extracted resources	9,107	3,881	<i>Manuf of petroleum prod.</i> ‡	205	89
			Manuf of wood prod.	2,727	1,120
			<i>Mining</i> ‡	8	2
			Manuf of other non-metallic mineral prod.	6,167	2,670
			Post & telecommunications	747	
ICTs & media	6,710	1,320	Publishing & printing	3,110	1,320
			Computer & related activities	2,853	
			Supporting & auxiliary transport activities	2,292	0
Transportation & storage	5,171 (7)	0	Transport communications	2,513	0
			Water transport	228	0
			Air transport	138	0
			Wholesale trade commission trade	10,779	0
Wholesale	10,779 (20)	0	Retail trade	25,626	0
Retail	25,626 (76)	0	Manufacture of food products	18,223	7,999
Food	18,223	7,999	Manufacture of furniture	4,500	2,048
Furniture	4,500	2,048	Hotels restaurants	7,669	1
Hospitality Tourism	7,669 (12)	1	Chemicals Chemical Prod.	6,407	3,411
Chemicals Chemical Prod.	6,407	3,411	Other business activities	51	2
Rest of universe	112 (19)	5	Agriculture	13	3
			Health & social work	48	

Note: the sub-column "Total" reports the total number of both service and manufacturing firms, while the column "Estimation" reports the number of manufacturing firms in the estimation sample. The number of manufacturing firms in the total sample is sometimes specified in parenthesis to explain sample attrition in the estimation sample.\* The sector is considered as a separate industry in robustness analysis. † Sectors are grouped together as a separate industry in robustness analysis.‡ Sectors are dropped from "Manufacture of extracted resources" in robustness analysis.

B.5. Dependant, independent, and instrumental variables correlation matrix.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18
V1: Email adoption	1.00																	
V2: Email spillovers	0.53	1.00																
V3: SMC x Distance infra	0.10	0.19	1.00															
V4: E(Outage days x Distance) [T; t-4]	-0.06	-0.10	0.20	1.00														
V5: Sum outage days [t; t-4]	-0.05	-0.09	0.25	0.91	1.00													
V6: SMC number	0.08	0.16	0.77	0.38	0.35	1.00												
V7: Website adoption	0.51	0.37	0.18	-0.05	-0.04	0.11	1.00											
V8: Exports (% sales)	0.15	0.06	0.02	0.02	0.03	0.01	0.15	1.00										
V9: Initial # perm. FT employees (ln)	0.17	0.03	0.09	0.11	0.12	0.10	0.19	0.22	1.00									
V10: % non-production workers	0.12	0.15	0.00	-0.07	-0.07	-0.05	0.10	-0.09	-0.09	1.00								
V11: % skilled workers	-0.09	-0.07	0.01	0.01	0.02	-0.01	-0.03	-0.01	-0.03	-0.01	1.00							
V12: Manager exp (years, ln)	0.14	0.20	0.09	-0.06	-0.05	0.05	0.13	0.06	0.02	0.03	0.01	1.00						
V13: Firm's age (years, ln)	0.15	0.16	0.15	0.00	0.00	0.13	0.17	0.06	0.11	0.06	-0.02	0.45	1.00					
V14: % of foreign ownership	0.10	0.02	-0.03	-0.04	-0.03	-0.08	0.09	0.19	0.16	0.03	-0.02	-0.03	-0.02	1.00				
V15: % of state ownership	0.01	-0.02	-0.03	0.02	0.02	-0.03	0.03	0.02	0.12	0.00	0.01	-0.02	0.05	-0.01	1.00			
V16: Bank loan (0 or 1)	0.24	0.26	0.019	-0.06	-0.05	-0.023	0.20	0.09	0.04	0.09	-0.08	0.13	0.14	-0.01	0.005	1.00		
V17: Electricity obstacle	-0.08	-0.19	-0.06	-0.010	-0.02	-0.06	-0.07	-0.00	-0.02	-0.02	-0.02	-0.02	-0.02	0.01	-0.02	-0.00	1.00	
V18: Distance to connectivity infra (km, ln)	-0.07	-0.13	-0.42	0.12	0.005	-0.05	-0.07	-0.05	-0.01	-0.06	-0.01	-0.09	-0.10	-0.06	0.02	-0.09	0.00	1.00

Online Appendix : <https://ferdi.fr/dl/df-25sCWb9YGk7Fr2M9NqnWckPS/wp288-online-appendix.pdf>



*“Sur quoi la fondera-t-il l'économie du monde qu'il veut gouverner? Sera-ce sur le caprice de chaque particulier? Quelle confusion! Sera-ce sur la justice? Il l'ignore.”*

Pascal



Created in 2003 , the **Fondation pour les études et recherches sur le développement international** aims to promote a fuller understanding of international economic development and the factors that influence it.

 **Contact**

[www.ferdi.fr](http://www.ferdi.fr)

[contact@ferdi.fr](mailto:contact@ferdi.fr)

+33 (0)4 73 17 75 30