MARCH 2023
BxWP2023-03

BORDEAUX ECONOMICS WORKING PAPERS
CAHIERS D'ECONOMIE DE BORDEAUX

Digital divides among micro-sized firms: Evidence from Sub-Saharan Africa

Damien GIROLLET
Univ. Bordeaux, CNRS, BSE, UMR 6060, F-33600 Pessac, France

Bordeaux Sciences Economiques
Bordeaux School of Economics

BSE UMR CNRS 6060
Université de Bordeaux
Avenue Léon Duguit, Bât. H
33608 Pessac – France
Tel : +33 (0)5.56.84.25.75
http://bse.u-bordeaux.fr/

Abstract

This paper explores digital inequalities in access and usage among 3,300 firms and entrepreneurs from eight sub-Saharan African countries. To account for informal firms' heterogeneity, we identify three segments: an upper tier of top performers, a lower tier of survivalists, and an intermediate segment composed of constrained gazelles. Although digital technologies are already used by most of the informal entrepreneurs in Sub-Saharan Africa, our findings suggest that the diffusion of these new technologies is uneven across informal firms, digital inequalities being rooted in the already existing socioeconomic inequalities. Indeed, digital inequalities align with the hierarchy of informal sectors in each country and are associated with entrepreneurs' and firms' characteristics. Using multivariate analysis, we find that gender and rural/urban digital divides persist in the productive sphere. At the same time, firms with a high level of informality, low profits, precarious operating conditions, and no access to financial services are less likely to use digital technologies.

Keywords: Digital technology, ICT, digital divide, informality, Africa.

JEL: D22, O17, O33, O55.

To cite this paper: **GIROLLET, Damien** (2023), **Digital divides among micro-sized firms: Evidence from Sub-Saharan Africa**, Bordeaux Economics Working Papers, BxWP2023-03

https://ideas.repec.org/p/grt/bdxewp/2023-03.html

Digital divides among micro-sized firms: Evidence from Sub-Saharan Africa

Damien Girollet

Abstract

This paper explores digital inequalities in access and usage among 3,300 firms and entrepreneurs from eight sub-Saharan African countries. To account for informal firms' heterogeneity, we identify three segments: an upper tier of top performers, a lower tier of survivalists, and an intermediate segment composed of constrained gazelles. Although digital technologies are already used by most of the informal entrepreneurs in Sub-Saharan Africa, our findings suggest that the diffusion of these new technologies is uneven across informal firms, digital inequalities being rooted in the already existing socioeconomic inequalities. Indeed, digital inequalities align with the hierarchy of informal sectors in each country and are associated with entrepreneurs' and firms' characteristics. Using multivariate analysis, we find that gender and rural/urban digital divides persist in the productive sphere. At the same time, firms with a high level of informality, low profits, precarious operating conditions, and no access to financial services are less likely to use digital technologies.

JEL classification: D22, O17, O33, O55

Keywords: Digital technology, ICT, digital divide, informality, Africa

1 Introduction

With an average informal employment rate of 89.2% in 2016 (Bonnet et al., 2019), the main form of economic activity in Sub-Saharan Africa (SSA) remains informality. Most workers and production units that produce legal goods and services thus remain at least partially organized outside or in conflict with business and labor market regulations. However, the disadvantages associated with informality have been highlighted in the economic literature over the past decade. Informal firms are, on average, less productive and smaller than their formal counterparts and face higher vulnerability due to severe economic, institutional, and social constraints (La Porta and Shleifer, 2014; Grimm et al., 2011). A large informal sector also considerably reduces the economy's tax base, implying lower fiscal resources and a lower capacity to finance public expenditure (Besley and Persson, 2014; Okunogbe and Santoro, 2022). Most implemented policies addressing informality challenges in SSA have therefore focused on the transition to formality, aiming to overcome such negative effects of informality and seeking to generate productivity gains, increase fiscal resources, and achieve greater economic and social inclusion. This push toward formalization has usually resulted in the implementation of a wide range of fiscal and business environment reforms aimed at either lowering the costs or increasing the benefits associated with formal sector participation.

However, existing evidence suggests that these formalization reforms have had limited success. The reduction in entry costs has little or no impact on the formalization of existing informal activities, with the vast majority of firms in developing countries remaining informal (Jessen and Kluve, 2021; Bruhn and McKenzie, 2014; Benhassine et al., 2018; De Giorgi and Rahman, 2013; De Giorgi et al., 2018; De Mel et al., 2013). Moreover, formalization does not seem to bring the expected benefits, as evidence of improved performance or access to financial services remains mixed (Campos et al., 2015; Floridi et al., 2020; Demenet et al., 2016; Berkel and Tarp, 2022), and the cost of these formalizing reforms exceeding the additional fiscal revenue generated (Benhassine et al., 2018).

Given the vulnerability of many informal firms and the limited performance of pro-formalization schemes, recent literature argues that a paradigm shift in policy priorities is needed to effectively address informality challenges in Sub-Saharan Africa (Choi et al., 2020; Nguimkeu and Okou, 2021; Kanbur, 2017). While pro-formalization programs consider the transition to formality as necessary to achieve greater productivity and inclusion, Roy and Khan (2021) suggest a shift towards policies that consider the reverse causal direction. Short to medium-term policies should focus on pro-productivity interventions to induce sustainable formalization and inclusion, as activities will not formalize unless they grow.

As a means of increasing the productivity of informal workers and businesses, digital technologies seem to hold great potential. A large and growing strand of empirical literature tends to confirm the existence of such a relationship at the firm level in the context of developed countries, with a significant positive impact of investments in digital technologies on labor productivity (Goldfarb and Tucker, 2019; Cardona et al., 2013). However, studies examining the role of digital technologies used in business performance in developing countries are sparse and often limited to formal firms for which data are more available and numerous (Commander et al., 2011; Motohashi, 2008; Cariolle et al., 2019; Paunov and Rollo, 2016; Cirera, 2016; Islam et al., 2018a). To date, very few studies have focused on informal firms in Sub-Saharan Africa, but all have found a significantly positive association between the use of digital technologies and their levels of sales, profits, or labor productivity (Esselaar et al., 2006; Atiyas and Dutz, 2021; Berrou et al., 2020; Eekhout et al., 2022; Danquah and Owusu, 2021; Gaglio et al., 2022; Islam et al., 2018b). Thus, digital technologies appear as a means to improve the efficiency of the production process of firms through different channels and mechanisms.

Indeed, digital technologies might address some market and state failures to which informal firms are particularly vulnerable. By reducing many specific economic costs such as search, replication, transportation, tracking, and verification costs (Goldfarb and Tucker, 2019), digital technologies improve firms' access to and use of information. Mobile phones

and the internet can provide faster, cheaper, and better access to information by offering immediate, on-demand, and reliable information (Aker and Mbiti, 2010; Aker and Blumenstock, 2015). Better information dissemination may improve coordination with the firm's partners by enabling frequent and instantaneous communication. Digital technologies may also provide access to financial services through mobile financial applications such as mobile money services. In addition to promoting financial inclusion, electronic transfers have significantly reduced the cost of sending and receiving money, and improved security and liquidity (Jack and Suri, 2014; Suri, 2017; Suri and Jack, 2016). Overall, informal enterprises that use digital technologies can benefit from reduced transaction costs, make their internal and external coordination more efficient, change their organizational structure, and extend their market reach. Similarly, higher digital inclusion provides new ways to enhance tax collection and compliance, and new opportunities to transition to formality through e-government processes (Apeti and Edoh, 2023; Mascagni and Nell, 2022; Chacaltana et al., 2018).

In this context, leveraging low-skilled-based digital technologies should be a good strategy to boost the productivity of informal firms in SSA (Nguimkeu and Okou, 2021). To promote the diffusion of digital technologies, most African countries have already implemented major telecommunications policy reforms since the early 2000s, which have coincided with growing investment in telecommunication infrastructures, mainly in the mobile segment (Moshi and Mwakatumbula, 2017). With 82% of the population of sub-Saharan Africa having mobile broadband coverage (ITU, 2021), the remaining challenge is to improve the quality of the internet connectivity in the region through the provision of a universal "meaningful connectivity", defined as a fast and unlimited daily broadband connection from a smartphone (A4AI, 2021). Therefore, substantial additional investment in infrastructure is required to improve the broadband value chain in African countries, including international connectivity capacity and local network infrastructures (The Broadband Commission for Sustainable Development, 2019; World Bank, 2016; Cariolle, 2020).

Nevertheless, investment in telecommunications infrastructure will not be enough to promote digital inclusion for all, as digital technology diffusion also brings new challenges and risks that can impede or reduce the benefits mentioned above. In particular, depending on supply-side and demand-side factors, the uneven diffusion of digital technologies generates digital inequalities - or digital divides - between and within countries. Initially defined as "the divide between those with access to new technologies and those without" (NTIA, 1999), the digital divide was firstly imbued with some technological determinism, in the belief that the universal dissemination of digital technologies is inevitable and would solve particular problems in the economy and society (Chandler, 1995; Gunkel, 2003). However, this prior consideration is no longer relevant as it has become clear that digital divides continue to expand even after physical access becomes universal. Notably, the literature shows that digital divides have emerged across populations having the same physical access to digital infrastructures, depending on individuals' socioeconomic and demographic characteristics (Hargittai, 2002). This second-level digital divide is often an extension of existing social inequalities, making digital inclusion a social rather than a technological challenge (Ragnedda and Muschert, 2013). As a result, the diffusion of digital technologies may not necessarily imply greater inclusion. It may induce the opposite effect by leading to the emergence of new inequalities and the reproduction of existing ones (Ragnedda et al., 2022; Van Dijk, 2017). Given the heterogeneous nature of informal activities in SSA, significant digital inequalities may emerge depending on existing social and economic vulnerabilities, and hinder most of them from benefiting from the expected productivity gains and greater inclusion.

We rely on the After Access business surveys conducted by Research ICT Africa (2020) in 2017-18 to explore digital inequalities in access and usage among 3,300 firms and entrepreneurs in eight sub-Saharan African countries[1]. Specifically, we focus on how the heterogeneity of informal firms is reflected in digital inequalities. First, using the methodology proposed by Grimm et al. (2012), we identify among each country three segments: an upper

[1]Senegal, Ghana, Nigeria, Rwanda, Kenya, Tanzania, Mozambique, and South Africa.

tier of top performers, a lower tier of survivalists, and an intermediate segment composed of constrained gazelles. We test the robustness of our partition by comparing its results with those of a clustering approach, and the construction of an index of firms' degree of informality. Then, we study the relationship between existing economic and social vulnerabilities of informal enterprises and their level of digital inclusion. After describing the level of ICT equipment of informal firms, their professional usage is explored through three major functions of ICT devices as described by Berrou et al. (2020): the external coordination function, the financial function, and the internal management function. We further perform a multivariate analysis to identify the main determinants of digital technologies usage for business purposes, considering both entrepreneur's and firm's characteristics, and controlling for the sector and regional effects. We conduct this analysis for the whole sample and then for each informal segment to observe the differences in the determinants of digital inequalities within the segments. Finally, we use a multivariate decomposition for nonlinear response models to decompose segments differences in the probabilities of digital technologies usage into differences in observable characteristics, on the one hand, and differences in the effect of these characteristics on the probability of using digital technologies, on the other.

We find evidence of significant heterogeneity among microenterprises in Sub-Saharan Africa, splited into three distinct segments: the top performers, the survivalists, and an intermediate segment of constrained gazelles. Apart from having specific entrepreneur and firm characteristics, these segments display strong disparities in their level of digital inclusion. The lower tier of survivalists is the most digitally excluded, both in terms of possession of ICT devices and professional usage. Conversely, the most successful firms, the so-called top performers, largely embrace digital technologies in their way of doing business. Although almost as well equipped as the top performers, firms in the intermediate segment exhibit a lower level of digital technologies usage for business purposes. Thus, digital inequalities align with the hierarchy of informal sectors in sub-Saharan Africa, and are associated with some characteristics of entrepreneurs and firms. The multivariate analysis finds a significant

gender gap and wide disparities between rural and urban areas in digital technologies usage for business reasons, with female-led businesses and rural enterprises having lower usage levels than others. Education level also appears to be an essential determinant of digital inequalities, confirming the importance of basic literacy in using computers and phones, and in accessing specific digital skills (Hargittai, 2002; van Deursen and Van Dijk, 2010). Concerning firm characteristics, we find that firms with a high level of informality, low profits, precarious operating conditions, and no access to financial services are less likely to use digital technologies. These results show that digital inequalities are congruent with social and economic inequalities traditionally observed among African informal firms. Nevertheless, the usage gap between top performers and the two others segments is not associated with similar factors according to the results of the multivariate decomposition. Regarding the comparison between the survivalists and the top performers, the second-level digital divide is mainly driven by differences in ownership of ICT devices, human capital, and level of economic inclusion. In contrast, the usage gap between constrained gazelles and top performers is mainly explained by differences in economic inclusion and induced differences in firm characteristics, as they tend to share similar entrepreneur characteristics.

Our paper contributes to the literature on the informal economy in two ways. First, we provide additional empirical evidence on informal firms' adoption and use of digital technologies in Sub-Saharan Africa. While such evidence remains scarce in the literature, this paper analyses the digital inclusion of informal firms from eight Sub-Saharan African countries, considering the diversity of devices and functions offered by digital technologies, and the heterogeneous nature of informal activities. Second, this paper explores the determinants of digital inequalities among informal firms in Sub-Saharan Africa. Our results confirm the existence of substantial digital inequalities among informal enterprises in sub-Saharan Africa. These inequalities are strongly associated with already existing social and economic vulnerabilities, reinforcing the exclusion of those who are already the most excluded. Addressing digital inequalities appears to have common and segment-specific levers, including access to

ICT devices, digital skills, electricity, and financial services. In the context of the policy shift described above, identifying the determinants of digital divides in access, usage, and benefit among informal firms are thus of prior importance, as they may orientate specific policies of support for informal firms and workers the more vulnerable to digital exclusion.

The remainder of this paper is organized as follows. Section 2 provides a literature review of digital divides and their determinants across individuals and firms in Sub-Saharan Africa. Section 3 describes the data and the conceptual framework for observing the uses of digital technologies. Section 4 presents the segmentation methods and the empirical strategy to identify the determinants of second-level digital divides. The results are discussed in Section 5, and Section 6 concludes.

2 Literature review

2.1 Digital divides definition

The digital divide concept initially referred to the gap between those who have and those who do not have access to certain forms of information and communication technologies (NTIA, 1999), mainly computers and the internet, although today, the concept also includes other digital technologies, hardware, and software, such as feature phones and smartphones (Van Deursen and Van Dijk, 2019). Initially, most studies considered this binary distinction between haves and have-nots, focusing on disparities in physical access (e.g., devices) and infrastructural access (e.g., coverage network) across demographic groups and countries (Riggins and Dewan, 2005). This access gap, referred to as the 'first-level digital divide', considered mainly digital inequalities through a narrow perspective, reducing access inequalities to mere technological and economic issues, which should disappear over time with market liberalization and deployment of telecommunication infrastructures (Thierer, 2000; Norris, 2003; Fuchs and Horak, 2008).

However, some scholars have argued that providing universal physical access to digital technologies will not be enough to bridge the digital divide, which in fact, arises when the use of digital technologies becomes ubiquitous in daily life (Van Dijk, 2017). This shift towards a broader perspective in the interpretation of digital inequalities refers to the 'second-level digital divide' (Hargittai, 2002; van Dijk, 2006). It aims to examine issues' beyond access', considering that digital inequalities mainly occur within the group of digital technology users itself, due to disparities in quality of access, digital skills, and actual usage of digital technologies (Hargittai, 2002; DiMaggio et al., 2004; van Deursen and van Dijk, 2014; Hilbert, 2011). Van Dijk and Hacker (2003) have also highlighted the complexity of access to digital technologies by theorizing it as a full process of technology appropriation going beyond mere physical access. Recognizing the digital divide as a complex and multidimensional phenomenon (van Dijk, 2006), the second-level digital divide has transcended the dichotomous "access approach", by considering digital inequalities as a continuum with real social implications (Ragnedda, 2019). Since digital inequalities are linked to existing social inequalities and can exacerbate them, the digital divide is now a social issue rather than just a technological one (Ragnedda and Muschert, 2013; Van Dijk, 2017).

The resource and appropriation theory developed by Van Dijk (2005) is in line with this theoretical shift, moving beyond methodological individualism and adopting a relational or network approach instead (Van Dijk, 2017). The core argument of this theory is that digital inequalities in access, skills, and usage are rooted in already existing social inequalities, as the latter lead to an unequal distribution of resources and, thus, to inequalities in the process of appropriation of digital technologies. Moreover, these digital inequalities may bring unequal benefits and tend to enhance already existing social inequalities. Indeed, even if access to and use of digital technologies do not automatically translate into tangible benefits, these first and second-level digital inequalities can also produce disparities in terms of social, economic, political, or cultural benefits, leading to a 'third-level digital divide' (Scheerder et al., 2017). Following this assumption, Heeks (2022) recently proposed a new conceptual framework to

understand the relationship between growing digital inclusion and digital inequalities in the global South. Through the concept of 'adverse digital incorporation', the author argues that inclusion in a digital system can lead to unequal benefits or even adverse outcomes for some groups and positive outcomes for others, resulting in greater digital inequalities despite growing digital inclusion.

2.2 Determinants of digital divides in sub-Saharan Africa at the individual-level

The determinants of the digital divide at the individual level have been widely studied in developed countries, but overlooked in Sub-Saharan Africa, and often limited to demographic and socioeconomic characteristics (Srinuan and Bohlin, 2011). Evidence suggests a rural-urban divide, consistent with findings in developed countries. Rural location is associated with lower access to and use of digital technologies, partly due to the lower mobile network and internet coverage, as remoteness and low population density make the deployment of terrestrial infrastructure less profitable than in urban areas (Buys et al., 2009). Similarly, access to a stable source of electricity appears to be a key factor in mobile phone ownership or internet adoption in West Africa (Forenbacher et al., 2019; Adeleke, 2021; Ochoa et al., 2022), highlighting the importance of complementary investment in electricity infrastructure to bridge the digital divide. This rural-urban divide can also be explained by the fact that income is an important determinant of digital inclusion in SSA (Birba and Diagne, 2012; Ochoa et al., 2022), as low-income households or individuals are less able to spend a significant portion of their income on digital devices or services. Affordability of digital services and devices remains a strong demand-side constraint in SSA, with very few countries currently meeting the accessibility target set by the UN Commission on Broadband for Sustainable Development of 2% of monthly GNI per capita for entry-level broadband service (ITU, 2021).

The lack of skills and capabilities to use digital technologies is another demand-side constraint. Thus, a lack of basic literacy is a barrier to digital inclusion, as most services and devices are designed for people who can read and write. Furthermore, knowledge of English or French is positively associated with greater digital inclusion, which can be explained by the greater availability of internet content in these languages (Ochoa et al., 2022; Penard et al., 2015; Pénard et al., 2012). More generally, because literacy and other digital skills are necessary to fully use and benefit from digital technologies, educational attainment appears to be a key determinant of digital inequalities (Deen-Swarray, 2016; Birba and Diagne, 2012; Ochoa et al., 2022; Hasbi and Dubus, 2020; Penard et al., 2015; Pénard et al., 2012; Forenbacher et al., 2019).

Basic demographic characteristics are also significant determinants of the digital divide in SSA. There is a gender gap in the access and use of digital technologies in most African countries (Ochoa et al., 2022; Birba and Diagne, 2012; Hasbi and Dubus, 2020; Penard et al., 2015; Pénard et al., 2012; Forenbacher et al., 2019; Gillwald et al., 2010). This gender gap may be due to differences in average digital skills or income levels, combined with social and cultural gender norms unfavorable to women (Hafkin and Taggart, 2001; Mumporeze and Prieler, 2017). Age of individuals is another common determinant of digital inequalities, with younger people being more likely to adopt digital technologies than older ones, probably due to higher education attainment and technology familiarity (Ochoa et al., 2022; Birba and Diagne, 2012; Hasbi and Dubus, 2020; Penard et al., 2015; Pénard et al., 2012; Forenbacher et al., 2019).

2.3 Digital divides among small firms in sub-Saharan Africa

There is little evidence on the level and determinants of the digital divide among firms in developing countries (Lythreatis et al., 2021; Srinuan and Bohlin, 2011). Yet the heterogeneous nature of these firms, which are for the most part informal microenterprises, has

been widely recognized as a key feature for several decades (La Porta and Shleifer, 2014; Cunningham and Maloney, 2001). The literature traditionally distinguishes between subsistence and growth-oriented entrepreneurs, which differ in their motivations, socioeconomic characteristics, growth potential, access to technology and capital, and level of financial inclusion (Benjamin et al., 2012; Schoar, 2010). Grimm et al. (2012) argues for the existence of another intermediate group, called "constrained gazelles", located between the lower tier of subsistence entrepreneurs (also defined as survivalists) and the higher tier of growth-oriented entrepreneurs (or top performers). Commonly considered as subsistence entrepreneurs, these constrained gazelles operate at low levels of capital and performance but share similar characteristics and skills with the top performers, revealing an untapped growth potential. Furthermore, addressing the firms' heterogeneity can also be done by going beyond the traditional dichotomy between formality and informality by recognizing the existence of a continuum of different levels of informality (Williams et al., 2016; Benjamin et al., 2012).

This heterogeneity in entrepreneurial attributes and firms' characteristics may lead to significant additional digital divides (Bhattacharya, 2019). First, the uneven distribution of resources among informal firms may induce inequalities in their opportunities to access and use digital technologies for business purposes. The high cost of ICTs is often the main barrier to ICT adoption and use cited by firms in Sub-Saharan Africa (Esselaar et al., 2006), highlighting that most firms face a lack of affordability of ICT equipment and associated services, potentially due to severe financial constraints such as low level of profits and no access to financial services. Thus, Atiyas and Dutz (2021) find that Senegalese micro-sized firms with access to bank loans are significantly more likely to own a smartphone. Moreover, the lack of knowledge and awareness of the potential benefits and the very existence of the new technology may contribute to slowing down adoption, especially for advanced technologies or uses (Esselaar et al., 2006). In addition to some industry-specific adoption patterns, firm size, level of physical capital, and access to electricity also appear to be positively cor-

related with access and use of digital technologies (Cirera et al., 2021; Berrou et al., 2020; Atiyas and Dutz, 2021). Second, entrepreneurial orientation may induce specific patterns in the appropriation process of digital technologies. According to the technological acceptance theory, the motivation, attitude, and intention to accept and use digital technologies may vary due to differences in perceived advantage, usefulness, and ease of use (Van Dijk, 2005; Chuttur, 2009). Hence, growth-oriented entrepreneurs, who have entrepreneurial traits that encourage the search for and pursuit of opportunities, as well as better management of risks and uncertainties, are expected to be more receptive to new digital technology innovations and more likely to adopt them (Tang and Konde, 2020; Fafchamps and Quinn, 2018). Similarly, competition in the domestic market and industry-level adoption of new technologies can affect technology adoption at the firm level through competitive concerns and spillover effects (Cirera et al., 2021). Finally, the level of formality has been pointed out as a key determinant of digital inclusion, higher level of formality being associated with greater access and use of digital technologies (Cirera et al., 2021; Esselaar et al., 2006; Deen-Swarray et al., 2013).

3 Data and descriptive statistics

3.1 The data

We use data from the After Access business surveys conducted by Research ICT Africa (2020) in 2017-18 across eight African countries[2]: Senegal, Ghana, Nigeria, Rwanda, Kenya, Tanzania, Mozambique, and South Africa. After Access surveys include a household survey and a business survey, conducted in 2018 for Senegal and Tanzania, and in 2017 otherwise.

[2]Due to important missing values, observations from Uganda were excluded from the analysis.

Household and business samples were generated randomly in each country, and the sampling procedure followed four steps[3]. First, based on a national census sampling frame, the enumeration areas (EA) in each country were separated into urban and rural ones. Second, the EAs were sampled for each stratum using probability proportional to size (PPS). Third, a census of all households and businesses in that area was conducted for each selected EA. Then, for each selected EA, a number of households and businesses were randomly sampled from this sampling frame. Thus the households and businesses surveyed are not directly linked, as the survey design is not similar to a 1-2-3 survey. As no other selection criteria were used, such as filter questions associated with a definition of informal enterprises, not all businesses surveyed are necessarily informal. We restrain the sample of surveyed businesses to non-farm activities, as agricultural activities are specific in their mode of production and are few in the sample. In addition, we delete some observations due to missing values on key variables. The final sample considered in the following analyses comprises 3,300 firms from eight sub-Saharan African countries[4].

Although businesses were randomly selected for the survey, they are not representative of the informal sectors in each country. Indeed, the sampling frame of the EAs is only representative at the household level, as each enumerator area has a household density of about 200. The repartition of businesses is not necessarily uniform within these EAs, particularly because of the potential concentration of informal businesses within activity areas such as markets. Furthermore, the listing of all businesses in each selected EAs omits the most hidden activities, those that take place in the homes of the entrepreneurs, as well as itinerant activities such as street vendors.

However, the data are particularly suitable for exploring in depth the level of digital inclusion of informal enterprises from diverse African countries with specific productive structures, business environments, or stages of digital technology diffusion. Indeed, harmonized surveys

[3]According to the technical documentation provided by RIA (2018).

[4]Appendix 1 presents the distribution of the surveyed firms across the eight African countries.

of informal enterprises in different countries asking them about their digital practices are valuable, and random sampling ensures the internal validity of the results.

The survey collects basic information on the firm's owner[5] and his/her business, such as socio-demographic characteristics of the entrepreneur, registration status of the business, sector of activity, labor force, relationship with suppliers and customers, economic performance, and level of physical assets. Businesses' sales, expenses, profit, and value of capital reported by entrepreneurs are expressed in local currency, so price differences between countries are adjusted via Purchasing Power Parity (PPP) conversion factors[6]. In addition, the survey investigates entrepreneurs' access to and use of digital technologies, exploring the multidimensionality of devices, interfaces, and functions that entrepreneurs may mobilize for business purposes.

3.2 Macroeconomic and digital environment across countries

The informal enterprises in the sample are located in eight different countries in sub-Saharan Africa with specific structural and macroeconomic environments, and varying levels of digital technology diffusion. Table 1 reports these disparities for each of the eight countries.

Of the eight countries considered, South Africa is a special case due to its upper-middle-income status. We also observe emerging economies with lower middle-income status, with Nigeria, Ghana, and Kenya. Senegal moved up to lower-middle-income status in 2018 but is still part of the least developed countries, with Tanzania, Mozambique, and Rwanda. By considering these eight countries, we have a good representation of what sub-Saharan African countries are in terms of income levels but not a good representation in terms of geographical distribution. Indeed, we do not consider any countries from Central Africa, as the countries in the sample are located in West Africa, East Africa, and South Africa. These different levels of development are correlated with significant heterogeneity in terms of urbanicity,

[5]Hereafter referred to as "entrepreneur".
[6]According to the year of the survey.

access to electricity, and literacy rates across countries. Except for South Africa, informality remains pervasive in all countries, with a high proportion of non-agricultural employment.

Infrastructural access is not a factor of inter-countries digital inequalities, as most of the population has access to at least a 3G mobile network coverage in 2017. While Nigeria and Senegal appeared to lag in 2017, the gap closed in 2018 as the proportion of the population with access to at least 3G coverage increased to 75.5% and 92%, respectively. More recent data shows that access to 3G or 4G networks is almost universal in all countries except Mozambique, Tanzania, and Nigeria, where between 25% and 15% of the population still lack access to such networks. Only two countries, South Africa and Nigeria meet the accessibility target of 2% of monthly GNI per capita for entry-level broadband service, confirming that affordability remains an important barrier to digital technologies adoption and use in Sub-Saharan Africa. Along with these disparities in infrastructural access and affordability, we observe digital inequalities in terms of usage. South Africa and Mozambique, the most and least economically developed countries in the sample, are also the most and least advanced in terms of digitalization. Indeed, while the spread of mobile telephony has been largely achieved in all countries, Mozambique is lagging behind. Conversely, half South Africa's population uses the internet, significantly outpacing other countries. We also note the well-known high adoption of mobile money services in Kenya.

Table 1: Structural characteristics and digital technology diffusion across countries in 2016-2017

	South Africa	Nigeria	Ghana	Kenya	Senegal	Tanzania	Rwanda	Mozambique
Stuctural characteristics								
Region	South	West	West	East	West	East	East	South
GNI per capita, PPP (constant 2017 Intl.$)[a]	13564	4963	4687	4231	3168	2402	1874	1250
Income level category[a]	Upper middle	Lower middle	Lower middle	Lower middle	Low	Low	Low	Low
Informal employment (% of non-agri. employment)[a]	35	-	83	70	90	72	68	87
Urban population (%)[a]	65,9	49,5	55,4	26,6	46,7	33,1	17,1	35,5
Access to electricity (%)[a]	84,4	54,4	79,0	56,0	61,7	32,3	34,1	24,3
Adult literacy rate (%)[a]	87	62	79	81,5	51,9	77,9	73,2	60,7
ICT diffusion								
Population covered by at least a 3G mobile network (%)[b]	98,6	54	80	85	66,3	85	93,4	80
Mobile-cellular sub-basket (as % of GNI p.c)[b]	1,5	1,8	2,4	2,3	17,3	6,4	8,4	6,0
Mobile-broadband, prepaid handset-based, 500 MB (as % of GNI p.c)[b]	1,3	1,7	4,5	4,4	8,3	3,0	4,4	5,2
Mobile-cellular subscriptions per 100 inhabitants[b]	156	75	122	87	104	71	72	42
Active mobile-broadband subscriptions per 100 inhabitants[b]	70	20	78	35	28	9	35	27
Internet users (%)[b]	56	33	38	18	30	16	17	8
Have mobile money account (%)[d]	19	5,6	39	72,9	31,8	38,5	31,1	21,9
ICT index								
Network Readiness Index (Rank out of 143)[c]	65 (1)	119 (6)	102 (4)	86 (3)	107 (5)	126 (8)	80 (2)	123 (7)
ICT Development Index (Rank out of 176)[b]	92 (1)	143 (5)	116 (2)	138 (3)	142 (4)	165 (8)	153 (7)	150 (6)

Notes: For the ICT indices, the ranking of the countries out of the eight observed is put in brackets.
Source: [a] World Development Indicators, World Bank (2017) ; [b] International Telecommunication Union, ICT Indicators Database, ITU (2017a), and ITU (2017b) ; [c] World Economic Forum, Baller et al. (2016) ; [d] Global Findex, Demirguc-Kunt et al. (2018).

These inter-country disparities in terms of digital inclusion are summarized by two composite indices: the Network Readiness Index (Baller et al., 2016) and the ICT Development Index (ITU, 2017b). Both indices are based on a different conceptual framework that aggregates indicators of different dimensions for each economy related to the regulatory environment of the telecommunications sector and/or the level of access, use, and impact of digital technologies[7]. These internationally comparable indicators make it possible to compare the level of digital transformation between economies. The overall ranking confirms that Sub-Saharan Africa remains the lowest-scoring region while displaying some heterogeneity among the countries in our sample. It confirms that South Africa is at an advanced stage of digital technology diffusion compared to other countries such as Tanzania and Mozambique. Thus, depending on the country, the informal enterprises in our sample do not benefit from the same macroeconomic environment, and the same level of digital inclusion. We intend to account for these differences between countries by including controls for regions within countries in our empirical strategy.

3.3 Firms and entrepreneurs characteristics description

Table 2 reports some firms' and entrepreneurs' characteristics for the total sample and each country. Although we observe significant differences across countries, it is difficult to draw any conclusions about the composition of countries' informal sectors, given that the samples for each country are not nationally representative[8].

In the whole sample, the entrepreneurs are mostly sole proprietors of their businesses and are 38 years old on average. The proportion of firms led by women is slightly lower than that of men, while 8% of businesses are joint ventures between men and women. The data also show that entrepreneurs' level of education remains low, with 36% of them having

[7]The NRI is a composite indicator made up of 53 individual indicators, while the IDI considers 11 indicators.

[8]See Mothobi et al. (2020) for a cross-country description of the sampled enterprises and entrepreneurs.

no education, and only 20% and 12% having reached secondary of tertiary education, respectively[9]. Half of the entrepreneurs state that the main reason for starting their business was the lack of other job opportunities and that they would have been unemployed and resourceless otherwise.

Most businesses are located in urban areas (63%). They have on average been set up for seven years and operate predominantly in the trade sector (68%), while firms in the manufacturing and service sectors represent only 12% and 20%, respectively, of sampled firms. The vast majority of businesses are not considered by any type of registration. However, one-third of enterprises are registered with some local or municipal authorities, and just over one-tenth are registered at the national level. This is consistent with the fact that most of the firms in the sample are not subject to any form of taxation. Indeed, while 39% pay taxes at the local or municipal level, only 17% are registered for national VAT or sales tax. The data also provide information on one of the other main formality criteria considered in Sub-Saharan Africa, namely keeping accounts in conformity with national or regional standards. It can be considered that only 5% keep advanced accounts that are probably compliant with standards, while 43% carry simple bookkeeping and the remainder none. This low level of formality is combined with low financial inclusion, as only 25% of businesses have access to a bank account.

The firms we consider are mainly micro-sized, as 97% have less than five full-time paid workers. Moreover, the absence of full-time paid workers is widespread, with 63% of entrepreneurs being self-employed without permanent employees. If the absence of workers is often observed, this is also the case for fixed assets, as 33% of the entrepreneurs declare that they do not own any fixed assets of value (machinery, vehicles, furniture, etc.).

[9]If there are several owners, we consider the one with the highest level of education. Similarly, for the owner's age, we consider the youngest owner.

Table 2: Characteristics of firms and entrepreneurs

	South Africa	Nigeria	Ghana	Kenya	Senegal	Tanzania	Rwanda	Mozambique	All
Entrepreneurs' characteristics									
Man	0.56	0.38	0.42	0.40	0.68	0.56	0.64	0.39	0.50
Woman	0.30	0.57	0.51	0.48	0.28	0.38	0.25	0.56	0.42
Joint venture (Woman and man)	0.13	0.04	0.07	0.12	0.04	0.05	0.11	0.05	0.07
Age	42.6	38.0	39.6	34.9	36.6	37.7	.	38.2	38.1
No education	0.09	0.28	0.18	0.02	0.51	0.68	0.55	0.47	0.36
Primary education	0.71	0.17	0.39	0.20	0.22	0.24	0.36	0.35	0.32
Secondary education	0.16	0.40	0.34	0.49	0.06	0.02	0.08	0.04	0.20
Tertiary education	0.04	0.15	0.08	0.29	0.21	0.05	0.02	0.13	0.12
No other opportunity	0.49	0.48	0.53	0.28	0.43	0.67	0.48	0.61	0.50
Firms' economic performance[a]									
Sales	1245	345	454	498	837	398	461	131	448
Value-added	527	129	235	280	586	187	169	28	205
Net profit	622	86	144	174	314	14	131	44	113
Firms' characteristics									
# of full-time paid workers	2.9	1.5	2.2	2.1	2.5	1.4	2.2	1.2	2.0
Any full-time paid worker	0.36	0.79	0.54	0.59	0.49	0.74	0.59	0.86	0.63
One or two full-time paid workers	0.45	0.14	0.34	0.34	0.34	0.24	0.32	0.12	0.28
Three or more full-time paid workers	0.09	0.04	0.06	0.04	0.11	0.02	0.05	0.02	0.05
Five or more full-time paid worker	0.10	0.03	0.06	0.04	0.07	0.01	0.05	0.01	0.04
Median physical capital[b]	1556	246	312	249	418	225	236	53	249
No capital	0.39	0.24	0.36	0.36	0.50	0.13	0.21	0.48	0.33
Urban	0.73	0.44	0.64	0.67	0.68	0.59	0.65	0.69	0.63
Manufacture	0.07	0.15	0.15	0.06	0.14	0.10	0.10	0.18	0.12
Service	0.27	0.18	0.20	0.26	0.21	0.15	0.24	0.07	0.20
Trade	0.65	0.67	0.65	0.69	0.64	0.75	0.66	0.76	0.68
Age of firm	5.7	7.6	6.8	3.6	6.1	4.8	17.6	6.0	7.1
Registration with any local or municipal authority	0.34	0.21	0.35	0.40	0.33	0.18	0.78	0.11	0.32
Registration at the national level	0.29	0.03	0.14	0.15	0.13	0.04	0.32	0.02	0.13
Pay local or municipal taxes	0.27	0.44	0.47	0.41	0.44	0.15	0.78	0.22	0.39
Registration for national VAT or sales tax	0.26	0.10	0.22	0.16	0.23	0.12	0.20	0.04	0.16
Formal accounts	0.06	0.02	0.01	0.01	0.21	0.02	0.02	0.02	0.05
Access to electricity	0.85	0.56	0.60	0.63	0.68	0.50	0.81	0.27	0.60
Bank account	0.55	0.26	0.32	0.28	0.14	0.09	0.38	0.10	0.25
Observations	364	492	387	388	470	457	349	393	3300

Notes: [a] Median values. [b] Median values among firms with non-zero capital.
Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

3.4 Global overview of digital technologies usage

Table 3 proposes an overview of access to and usage of digital technologies for business purposes for the total sample and each country.

We consider the first-level digital divide through the ownership of mobile phones, landlines, and computers. These devices offer different possibilities of use through which one or more professional functions of ICT can be mobilized. As these different tools are not necessarily substitutable, both in their specific uses and in the network of contacts that can be reached, the cumulative possession of these devices may offer entrepreneurs a greater capacity for use, and perhaps different benefits. The entrepreneurs in the sample widely own mobile phones, with a penetration rate of 79%, which is consistent with the fact that the access gap to mobile phones is narrowing in the region, making them increasingly ubiquitous (GSMA, 2021)[10]. Access to other ICT devices is extremely limited, with only 6% and 4% having access to a computer and a landline phone at the workplace, respectively, confirming the hegemony of the mobile phone and the limited penetration of these devices in Sub-Saharan Africa. Very few businesses have a computer or a landline without having a mobile phone. Overall, about 21% of the businesses in our sample do not have direct access to an ICT device, while the proportion of businesses with more than one ICT device reaches only 7%. Among firms that do not own a mobile phone, more than half claim they do not need one, and almost a quarter report that the main reason is that it is too expensive.

[10]However, the database does not allow for differences in the quality of devices owned by entrepreneurs, particularly between cell phones and smartphones. The ability of entrepreneurs to access a mobile broadband internet connection, a key determinant in internet access, is unknown. This figure, therefore, only guarantees access to basic telecommunication services for entrepreneurs.

Table 3: Digital technology ownership and usage among surveyed firms

	South Africa	Nigeria	Ghana	Kenya	Senegal	Tanzania	Rwanda	Mozambique	All
Ownership of ICT devices									
Mobile phone	0.92	0.66	0.78	0.87	0.89	0.81	0.67	0.7	0.79
Landline phone	0.14	0.02	0.04	0.01	0.09	0	0.01	0	0.04
Computer	0.19	0.01	0.05	0.03	0.1	0.03	0.04	0.02	0.06
Number of ICT devices owned									
Any	0.07	0.34	0.21	0.12	0.11	0.19	0.31	0.3	0.21
One	0.7	0.63	0.72	0.84	0.77	0.77	0.66	0.68	0.72
Two	0.13	0.02	0.05	0.03	0.07	0.04	0.02	0.02	0.05
Three	0.1	0	0.01	0	0.05	0	0	0	0.02
ICT usage for external coordination									
Bilateral coordination with suppliers	0.43	0.41	0.49	0.63	0.66	0.2	0.78	0.23	0.47
Bilateral coordination with customers	0.38	0.41	0.56	0.66	0.61	0.19	0.7	0.24	0.46
Multilateral coordination	0.28	0.09	0.13	0.11	0.23	0.06	0.05	0.03	0.12
ICT usage for financial transactions									
Online or mobile banking	0.22	0.04	0.07	0.12	0.08	0.06	0.09	0.11	0.09
Mobile money to send money	0.05	0.02	0.25	0.56	0.31	0.11	0.42	0.12	0.22
Mobile money to receive money	0.03	0.02	0.28	0.41	0.24	0.13	0.4	0.2	0.21
ICT usage for internal coordination									
Use of management software	0.09	0	0	0.01	0.07	0.01	0.02	0.01	0.02
Number of ICT functions used									
Any	0.37	0.44	0.27	0.12	0.19	0.66	0.15	0.56	0.35
One	0.21	0.23	0.23	0.15	0.2	0.14	0.08	0.17	0.18
Two	0.18	0.26	0.21	0.26	0.25	0.08	0.34	0.14	0.21
Three	0.11	0.04	0.12	0.19	0.11	0.07	0.11	0.07	0.1
Four	0.05	0.02	0.11	0.18	0.1	0.03	0.23	0.04	0.09
Five	0.06	0	0.06	0.08	0.09	0.01	0.07	0.02	0.05
Six	0.01	0.01	0.01	0.02	0.04	0.01	0.01	0.01	0.01
Seven	0	0	0	0.01	0.01	0	0.01	0	0
Observations	364	492	387	388	470	457	349	393	3300

Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

The second-level of the digital divide is observed through the actual usage of ICT and the diversity of functions mobilized for professional reasons. No information was collected on digital skills, and incomplete information was collected on the different interfaces used and the intensity of use. While it limits our consideration of the multidimensionality of digital technologies usage, the diversity of use is already relevant in characterizing second-level digital inequalities among African microenterprises. Subsequent analyses then focus on the purposes for which firms use ICTs, not the manner or frequency of their use.

The usage gap between firms appears to be even wider than the access gap, with 65% of businesses using ICT for business purposes. Among enterprises that do not have direct access to an ICT device, just over a third still report using ICT for business purposes. It highlights the complexity of assessing access to ICT devices in a context where shared devices and the interweaving of business and social spheres can provide opportunities for indirect access to ICT. Conversely, ownership of a mobile phone or other ICT devices does not necessarily imply professional use, as the use of digital technologies may remain confined to the social sphere. More than a quarter of enterprises with access to an ICT device do not use it for business reasons, illustrating the importance of going beyond access to ICT and observing actual usage. Given that most businesses have undergone some degree of digitalization by integrating digital technologies into their business operations, it is essential to explore for what purposes businesses use ICT devices.

Following Berrou et al. (2020), we consider three major functions of ICT devices in the business context: the external coordination function, the financial function, and the internal management function. The external coordination function considers ICT devices as tools to improve market coordination with the firm's external partners. Businesses can use digital technologies to communicate in a bilateral, or "one-to-one", way in the context of interpersonal relationships. The available data allow us to consider whether businesses use digital technologies for bilateral coordination upstream of production with suppliers or downstream of production with customers. These usages are the most widespread among the sampled

businesses, with 48% and 47% using digital technologies for bilateral coordination with their suppliers and customers. These two specific usages of digital technologies are significantly correlated, as three-quarters of the firms that practice bilateral coordination with their suppliers also do it with their customers, and conversely. However, face-to-face communication remains the most common mode of communication for most entrepreneurs, as only 24% and 18% report preferring ICT devices to communicate with their suppliers and customers, respectively. It is confirmed by the fact that ICT is most often used as a complementary communication tool to face-to-face communication rather than as a substitute, as few businesses (11%) exclusively use digital technologies to communicate with their suppliers.

Digital technologies also allow for more extensive market coordination through multilateral, or "one-to-many", coordination. This concerns the use of digital technologies, and mainly the use of the internet, for information retrieval, online sales or advertising, and group or mass communication through social media. Multilateral coordination is much less common among the businesses in our sample, with only 12% of them reporting that they use the internet or social media for business purposes. Although low, this is in line with the proportion of people using the internet in Africa in 2017, which stood at only 21.8% (ITU, 2017b). Then social media presence and internet use are advanced usages. In addition to requiring an internet device and an affordable and quality connection, such usage requires higher digital skills than those needed to communicate bilaterally with suppliers or customers.

Then, the financial function considers digital technologies as tools for financial inclusion, allowing people to benefit from financial services other than through banking and to carry out cashless financial transactions. Mobile money services are used for business purposes by 27% of the businesses in the sample. Among them, 77% receive transfers from their customers, and 63% pay their suppliers with these services. Some of these businesses also use mobile money to pay bills (48%), taxes (13%), or wages (14%). The use of mobile money for business purposes does not necessarily translate into sending and receiving transfers since 41% of users

of these services are only concerned with one of the types of transfers. In addition to the low adoption rate, the intensity of use of mobile money services is low among the businesses that use them, with only 13% of them receiving payments daily. Also considered is the use of online or mobile banking, which provides digital access to financial services based on the possession of a bank account (9%). It highlights the low financial inclusion of the businesses in our sample, with 57% of them neither having a bank account nor using mobile money services.

At last, the internal coordination function considers digital technologies as a tool for managing the firm's internal operations, such as employee management, account management, and administrative procedures. We consider internal coordination within the firm through the use of software for inventory management, accounting, or performance monitoring. Important resources in terms of digital skills and equipment are required to implement the use of such software since only computer software is considered here. Given the few enterprises with computers, using such management software concerns only 2% of the enterprises surveyed.

Finally, this overview hides significant disparities between countries, as the level and pattern of digital technologies diffusion among surveyed firms in each country vary. Access to ICT equipment remains a major issue in the digital inclusion process in some countries, while ICT equipment is now almost universal in others. Mobile phone ownership ranges from 66% in Nigeria to 92% in South Africa. Fixed phone and computer ownership concerns the South African and Senegalese samples essentially, with these devices being almost absent in the other countries. However, access divides remain relatively narrow across countries compared to usage divides.

The proportion of businesses using at least one of the functions of ICT for business reasons varies significantly between countries. Less than half of the businesses in Tanzania (34%) and Mozambique (44%) samples are involved in business usages of ICT. In comparison, three-

quarters or more of the businesses are users in Kenya (89%), Ghana (74%), Rwanda (85%), and Senegal (81%) samples. These regional inequalities could be explained by differences between countries in terms of structural and macroeconomic characteristics, level of ICT infrastructure and accessibility, as well as firms' and entrepreneurs' characteristics. However, the non-representativeness of the data at the national level does not allow us to judge the determinants of these cross-country inequalities, as they may be due to the structure of the samples in each country.

4 Empirical strategy

4.1 Heterogeneity of micro-sized firms: a deductive approach

To examine whether informal firms' heterogeneity is related to their level of digital inclusion, we rely on the method proposed by Grimm et al. (2012). This method, originally used to classify informal enterprises in the context of some West African countries (Grimm et al., 2012; Lavallée and Roubaud, 2019), has also been applied in other contexts such as the Democratic Republic of Congo by Adoho and Doumbia (2022), Ethiopia by Abebe et al. (2018), Mexico by Negrete-García (2018), and Morocco by Moosa (2019). The methodology follows a deductive approach as it assumes the existence of three distinct and homogeneous groups within the informal sectors: (1) top performers, a fixed proportion of entrepreneurs with the highest economic performance according to the selected criteria; (2) constrained gazelles, which are entrepreneurs with similar characteristics to the top performers, which gives them some economic potential, but are far from their economic performance levels; and (3) survivalists, a group of subsistence entrepreneurs with fundamentally different characteristics and limited economic potential.

4.1.1 Defining top performers

We define top performers based on a combination of two criteria as in Grimm et al. (2012). First, in each country, we select the top 40% of firms with the highest value of fixed physical capital. This size criterion allows us to identify businesses that have had the capacity and the motivation to grow in the past. This size criterion, therefore, reflects past performance and the propensity to invest in the business. In a way, it ensures that we retain growth-oriented entrepreneurs and firms with access to capital, two significant factors of success among micro and small enterprises (La Porta and Shleifer, 2014). The second step consists of retaining from these firms the 50% with the highest net profit. This performance criterion aims to capture the current performance of the firm. This deductive method automatically assigns about 20% of the sampled firms to the group of top performers.

Our approach differs somewhat from that initially proposed by Grimm et al. (2012), which identifies the group of top performers by first selecting the firms in the top 25% of physical capital value distribution, from which they select the 40% of firms with the highest capital profitability. First, while we consider the same size criterion, we consider the net profit instead of capital profitability to assess the current economic performance of firms. We argue that in our context, where firms are mainly engaged in commercial activities, capital productivity is not the most appropriate indicator to assess the success of firms. We also do not choose labor productivity because we do not have a measure of the volume of hours worked, and most firms are self-employed entrepreneurs with no other workforce. We also prefer the firm's net profit over the turnover, as it measures the entrepreneur's net income and the firm's capacity to reinvest. Second, we adapt the thresholds to identify a group of top performers representing 20% of the sample and not only 10%. Compared to Grimm et al. (2012), which use 1-2-3 surveys data, our sample of firms does not include the most vulnerable activities, such as household businesses or street vendors. Then we argue

that a 10% group of top performers is quite restrictive and risks assigning successful firms to the groups of constrained gazelles.

4.1.2 Identifying constrained gazelles and survivalists

Then, our empirical method aims to capture the heterogeneity of entrepreneurs who have not been defined as top performers. While traditionally this group of remaining entrepreneurs is defined as subsistence entrepreneurs, here we want to identify among them the entrepreneurs with the potential to become successful. These entrepreneurs, called constrained gazelles, have similar characteristics to top performers. Therefore, constrained gazelles are entrepreneurs with a high empirical probability of being top performers, given their observable characteristics. Following Grimm et al. (2012), the probability of being a top performer is estimated by the following probit model:

$$Pr(Y_i^{TP} = 1) = \Phi(\beta_0 + X_i'\beta_1 + \omega) \tag{1}$$

Where Y_i^{TP} is a dummy variable that takes the value 1 if the entrepreneur is a top performer, and 0 otherwise. X_i' is a vector of entrepreneur and firm characteristics. β_0 and β_1 are the vector of coefficients indicating how these characteristics influence the probability of being a top performer, and ω the error term. Finally, Φ is the cumulative density function of the standard normal distribution. Standard errors are clustered at the enumeration level.

To minimize endogeneity issues, the set of explanatory variables is limited to predetermined factors, i.e., characteristics that are observable prior to, or at the time of, the creation or starting management of the business by the entrepreneur. Although only predetermined factors are considered, we expect the model to identify a group of entrepreneurs, called constrained gazelles, with similar entrepreneurial behavior and characteristics to top performers but clearly different from survival entrepreneurs.

The predetermined factors considered as regressors in the model are the entrepreneur's gender and level of education[11], urban location, whether the main source of initial capital used to start the business is microfinance or bank loan, and the motivation of the entrepreneur to set up the business. This last variable is measured by a dummy variable which takes the value 1 if the entrepreneur started the business because he or she had no other job opportunity, i.e., he or she would be unemployed otherwise. As in Grimm et al. (2012), in addition to these variables, we add controls for sectors and countries, as well as for the age of the firm.

Based on this binary response model, we predict for all the entrepreneurs in our sample the probability of being a top performer, using the estimated parameters $\widehat{\beta_0}$ and $\widehat{\beta_1}$, and the observed vector of entrepreneur and business characteristics:

$$Pr(\widehat{Y_i}^{TP} = 1) = \Phi(\widehat{\beta_0} + X_i^{'}\widehat{\beta_1} + \omega) \tag{2}$$

Where "hats" indicate estimated parameters. In order to segment the sample of entrepreneurs in each country into these three groups - top performers, constrained gazelles, and survivalists -we use these predicted probabilities in the following way.

For each country, we calculate the average predicted probability of the top performer group identified in the previous section based on a double criterion of size and performance. The group of constrained gazelles must be defined in such a way that the average of the predicted probabilities of this group is identical to the average among the top performers. To do this, the second step consists of ranking the entrepreneurs not defined as top performers in descending order of their predicted probability. Then, starting with the non-top performer with the highest predicted probability, we add the following individuals until the average predicted probability of the constituted group is equal to the average predicted

[11]We do not include the entrepreneur's age due to missing data for all the sample of Rwanda. For the other countries, more than 90% of observations are assigned to the same group whether we consider or not the age of the owner in the model.

probability of the top performer group. Thus, the distribution of variables used as regressors in the model must be similar between the constrained gazelles and the top performers, and therefore the two groups share similar observable predetermined factors. The group of survival entrepreneurs includes the remaining entrepreneurs who are not top performers or constrained gazelles.

4.2 Alternative measures of heterogeneity

In the methodology described above, the choice of the criteria, the order in which they are applied, and the associated thresholds, partly determine the partition results. Although Moosa (2019) demonstrates that the modification in composition and characteristics of groups is limited when applying diverse alternative empirical specification[12], we test the robustness of our results by comparing this segmentation with that resulting from the application of cluster analysis and the construction of an index of firms' level of informality.

4.2.1 Cluster analysis: an inductive approach

We rely on exploratory data analysis methods to let the segmentation emerge from the data with a minimum of prior structure imposed, by exploring the similarities (or dissimilarities) between individuals in a multidimensional perspective. To implement this inductive classification, we rely on a Hierarchical Classification by Principal Components (HCPC) as detailed by Husson et al. (2010).

In line with the literature on heterogeneity among informal firms in developing countries, we consider four broad classes of variables (12 variables) in this clustering approach (Cunningham and Maloney, 2001). First, we include some characteristics of the entrepreneurs with two variables indicating their education level and motivation to start the business. Second, we include five variables displaying the size of the firm and its financial performance:

[12]Such as including only one criterion, inverse double-criteria, or choosing different thresholds.

the value of the stock of physical capital used, the value of monthly net profit, the value of monthly labor productivity, the number of full-time workers, and access to electricity in premises. The three continuous variables are grouped into terciles in each country, to ensure that the best-performing businesses in each country are assigned to the top tercile[13]. Third, as entrepreneurial traits and behaviors are expected to lead to heterogeneity in African informal sectors, four variables are added, including bookkeeping, separating business and personal finances, and having a bank account. Finally, the clustering approach considers the firms' participation in formal institutions through two variables: business registration and payment of taxes.

In the first stage, we use factor analysis as a pre-processing step for two reasons. First, factor analysis is used as a dimensionality reduction technique, which is relevant in our case, considering the large number of variables retained in the analysis and the complex relationships existing between them. Moreover, factor analysis allows the transformation of categorical variables into continuous ones. Due to the qualitative nature of the variables, the factor analysis was carried out using Multiple Correspondences Analysis (MCA). As we applied factor analysis as a pre-processing step with the main objective of obtaining the coordinates of the individuals in each component, factor analysis results are not interpreted. Based on the Kaiser criterion (Kaiser, 1960)[14], we retain the first eight dimensions, which reflect 58% of the total inertia.

We use the individuals' coordinates on these eight factors to carry out an Agglomerative Hierarchical Classification (AHC) to let a segmentation emerge from the data, categorizing entrepreneurs and their businesses into homogeneous groups. From the factor scores, the algorithm determines the Euclidean distances between individuals. Then, through the Ward's criterion method, individuals are aggregated into clusters in order to minimize the within inertia.

[13]For the value of the stock of physical capital, we define terciles by considering only enterprises with a nonzero value. Firms without capital are grouped in a different category.

[14]Which suggests retaining all dimensions whose inertia is greater than the average inertia.

The number of clusters is determined by analyzing the overall appearance of the hierarchical tree and the bar plot of within inertia gain. Partition into three groups is the most appropriate solution because the groups are homogeneous within themselves and distinct from each other. Solutions with five clusters, while providing inertia gains, constitute clusters of small size that could be less intelligible. Specifically, it divides the two extreme classes into two subgroups each, providing little additional information.

Despite AHC is an effective clustering algorithm and permits to determine the number of clusters without major contributions of the user, its results are not optimal and can be consolidated by k-means to maximize between-inertia. The partition obtained by the HCPC is introduced as the initial partition of the k-means algorithm, which, after a few iterations, consolidates the partition.

4.2.2 Degree of informality

Microenterprises heterogeneity can also be explored by looking beyond the traditional dichotomy between formality and informality. Indeed, there are many types of formalization processes, such as registration, tax payment, and compliance with accounting standards or labor regulations. These different processes are as many ways to maintain formal relationships with institutional actors operating at the local, regional, or national level. Therefore, the formal/informal dichotomy does not seem to capture the complexity and diversity of the relationships that microenterprises have with formal institutions. This implies the need to consider a continuum of degrees of informality instead through the combination of criteria to avoid an "all or nothing" criterion.

In line with Williams et al. (2016), we construct an index of the informality level of businesses by considering the following variables: (1) being registered with a local authority or municipality, (2) being registered at general registrar, (3) paying local or municipal taxes (tax stamps), (4) being registered for national VAT or sales tax, and (5) keeping accounts

according to national or regional standards. The cross-country nature of the survey implies that each of these criteria reflects different country-specific realities. This measure provides a six-point informality scale based on which we classify firms into four categories: totally informal, high informality, low informality, and totally formal[15].

4.3 Determinants of second-level digital divide

4.3.1 Multivariate analysis

To identify the main determinants of the second-level digital divide among microenterprises in Sub-Saharan Africa, we rely on a standard logit regression that estimates the effect of entrepreneur and firm-level characteristics on the probability to use digital technologies for business purposes. The dependent variable is a binary indicator taking the value 1 if the enterprise uses at least one of the following functions: bilateral coordination with suppliers, bilateral coordination with customers, multilateral coordination, sending money through mobile money services, receiving money through mobile money services, use of mobile or online banking, internal coordination through management software, and taking the value 0 otherwise. The following model is estimated :

$$DTuse_i = \beta_0 + \beta_1 X_i + \beta_2 B_i + \beta_3 DTownership_i + \beta_3 Region_i + \epsilon_i \tag{3}$$

Where $DTuse_i$ is a binary dependent variable that indicates whether a firm i uses digital technologies for business purposes. X_i is a vector of entrepreneur-level characteristics, including gender, educational attainment, and motivation to start the business[16]. B_i is a vector of firms' basic characteristics, including firm age, the logarithm of monthly profit, access to electricity, number of full-time paid workers, level of informality, sector of activity, and urban

[15]Totally informal firms meet zero criteria, high informal firms meet one or two criteria, low informal firms meet three or four criteria, and formal firms meet all criteria.

[16]We do not include the age of the entrepreneur due to missing data for all the sample of Rwanda

location. We add to this vector whether the entrepreneur has a bank account, separates business finances from personal finances, and keeps records, as well as some characteristics of the trading network, such as having formal partners, having businesses as customers, and having partners located further away than the surrounding villages and towns. As the possession of ICT devices influences the usage of digital technologies for business purposes, $DTownership_i$ is a set of dummies that indicate the number of ICT devices the firm possess among mobile phone, computer, and landline phone. $Region_i$ is a set of 82 region dummies controlling for intra- and inter-countries disparities in terms of levels of economic activity and infrastructural access. ϵ_i is an error term clustered at the Enumerator Area (EA) level. Only informal firms are considered in this regression, as being totally formal perfectly predicts the success of the binary outcome[17].

In order to observe whether there are differences in the determinants of digital technology use between the segments, similar regressions are run for each segment. As being located in some regions perfectly predict the success or failure of the binary outcome, we replace the region dummies with country dummies.

4.3.2 Multivariate decomposition

To further explain the sources of the differences in usage rates of digital technologies between the three segments, we perform a decomposition technique. We use a multivariate decomposition for nonlinear response models (Yun, 2004; Powers et al., 2011) , which extends the decomposition technique initially proposed by Oaxaca (1973) and Blinder (1973) to nonlinear models. This strategy allows attributing differences in binary outcome between groups to endowment and coefficient effects. The goal is to decompose segment differences in the probabilities of digital technologies usage into differences in observable characteristics (endowment effect), on the one hand, and differences in the effect that these characteristics have

[17]The 53 firms identified as being totally formal according to the informality index are then excluded from the analysis.

on the outcome of interest (coefficient effect), on the other. We restrain the interpretation of the decomposition results to the endowment effect, as results of the coefficient effect may be biased due to unobserved variables. Similar analyses have been carried out by Galperin and Arcidiacono (2021) to decompose the gender digital gap in Latin America.

We use the same predictors as in the multivariate analysis model above, except that we replace region effects with country effects, the ICT equipment ownership indicator by its individual components, and the level of informality by two dummies indicating whether the firms are registered and pay taxes, to avoid that these indicators predict success perfectly. Hence, we want to explain the observed differences in digital technologies usage prevalence between segments. Regarding the comparison between the survivalists and the top performers, the assumption is that the second-level digital divide is mainly driven by differences in ownership of ICT devices, human capital and entrepreneurial behavior, and level of economic inclusion (financial inclusion, level of informality, and access to public utilities). In contrast, differences between constrained gazelles and top performers are expected to be mainly explained by differences in the level of economic inclusion and induced differences in firms' characteristics, as they tend to share similar entrepreneur characteristics.

5 Results

5.1 Heterogeneity of micro-sized firms

5.1.1 Segmentation of the sampled firms: the deductive approach

Table 4 reports the estimated coefficients and the corresponding average marginal effects for equation 1, which estimates the probability of being a top performer on a set of predetermined factors.

Female-led firms are less likely to be top performers than men-led firms and joint ventures between men and women. The marginal effect indicates that the probability of being a top performer decreases by 9.5% points for firms managed by a female entrepreneur. This is a similar result that in Grimm et al. (2012) and other studies based on the same classification methodology, that relates to the gender gap in the economic performance of informal firms observed in the literature (Islam and Amin, 2022). The entrepreneur's education level is positively and significantly correlated with being a top performer. Compared to entrepreneurs with no education, primary education increases the probability of being a top performer by 11% points, while secondary and tertiary education increases it by 19% points. The entrepreneurial orientation of the entrepreneur, proxied by his motivation to start the business, is also correlated with top performance. Entrepreneurs who start their activities because they would be unemployed otherwise have a 6% points lower probability of being top performers, the latter being more often opportunity entrepreneurs. The firm's age has a significant and positive effect, as at the margin an increase in the firm's age by one year increases its probability of being a top performer by 0.7% points. However, this effect has a decreasing trend as the quadratic term is negative. This high magnitude is consistent with the fact that we identify top performers based on their level of physical capital stock value. This size criterion relates to the capital accumulation process of the firm, illustrating past performance and probably the firm's age. Similarly, having start capital financed by a formal loan from a microfinance institute or bank increases the probability of being a top performer by 12% points. We do not observe a significant effect for sectors, potentially due to our inability to disaggregate the sector to a lower level than manufacture, service, and trade. Finally, urban location increases the probability of being a top performer by 5% points, highlighting significant differences between rural firms and urban ones in terms of levels of performance.

Based on this regression from which we derive the predicted probabilities of being a top performer for each firm, we identify the group of constrained gazelles in sort that its

Table 4: Probability of being a top performer (probit model)

	(1) Coefficients	(2) Marginal effects
Female	-0.371***	-0.095***
	(0.055)	(0.014)
No education	(Ref.)	
Primary education	0.486***	0.110***
	(0.076)	(0.018)
Secondary education	0.738***	0.186***
	(0.092)	(0.025)
Tertiary education	0.739***	0.187***
	(0.103)	(0.029)
No other opportunity	-0.217***	-0.056***
	(0.060)	(0.015)
Age of firm	0.028***	0.007***
	(0.008)	(0.002)
Age of firm (squared)	-0.0004**	-0.0001**
	(0.0002)	(0.0001)
Initial capital (Formal loan = 1)	0.482***	0.124***
	(0.125)	(0.031)
Manufacturing	(Ref.)	
Service	0.051	0.014
	(0.095)	(0.026)
Trading	-0.101	-0.026
	(0.079)	(0.021)
Urban location	0.196**	0.050**
	(0.077)	(0.020)
Country effects	Yes	Yes
Pseudo-R2	0.082	
Observations	3300	3300

Notes: *p < 0.05,**p < 0.01, ***p < 0.001.
Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

average predicted probabilities are equal to the average in the group of top performers in each country (cf. Section 4.1.2). Table A1 reports the partition into three segments across the eight countries. Overall, along the group of top performers, representing about 20% of the sample, constrained gazelles regroup 40% of the sampled firms and survivalists the same proportion.

We check whether the partition verifies two criteria inherent to the methodology proposed by Grimm et al. (2012). First, by construction, the distribution of predicted probabilities of constrained should overlap the top performers' distribution, while the survivalists' dis-

tribution should be clearly distinct. It assures that, regarding the predetermined factors considered in equation 1, entrepreneurs that we identify as constrained gazelles are on average equally likely to be top performers than the actual top performers. In Figure 1, we report the kernel densities of predicted probabilities of being a top performer for each segment. We observe that kernel densities of constrained gazelles and top performers overlap, while the predicted probabilities of survivalists are concentrated at the low values, as expected. However, even among top performers, some firms have low predicted probabilities, confirming the limited predictive power of the probit model while only considering predetermined regressors.

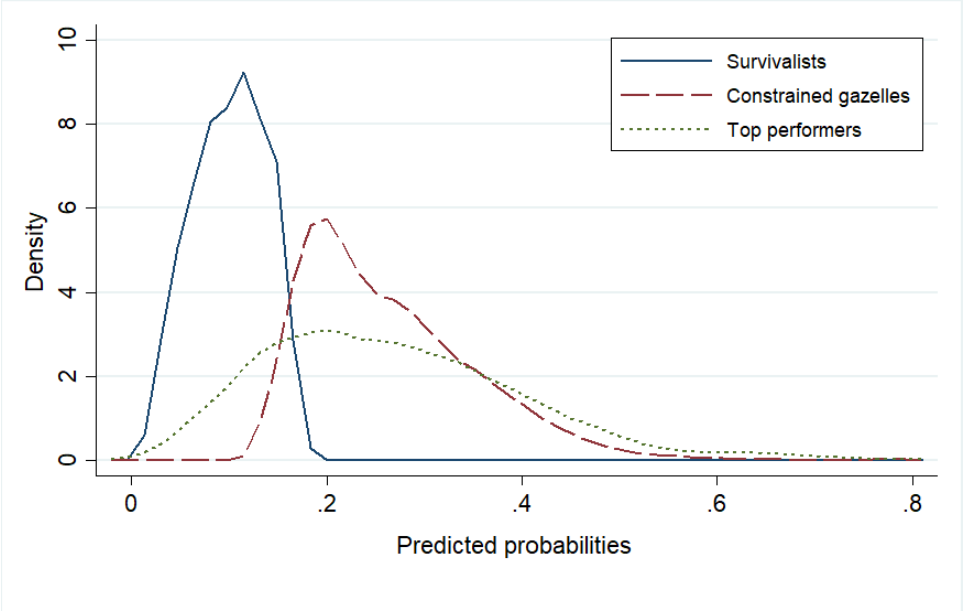

Figure 1: Predicted probabilities of being a top performer by segment

Second, constrained gazelles should have a similar distribution to survivalists in terms of physical capital stock and net profit, the two selected criteria used to identify top performers. The kernel densities of the two distributions[18] for each segment verify this condition, although the distributions of constrained gazelles are slightly more on the right than the ones of survivalists (Figures 2 and 3). It confirms that although constrained gazelles are on average equally likely to be top performers than the actual ones, they share similar levels of capital accumulation and economic performance to the lower tier of survivalists.

[18]Null values are not represented in these graphical distributions.

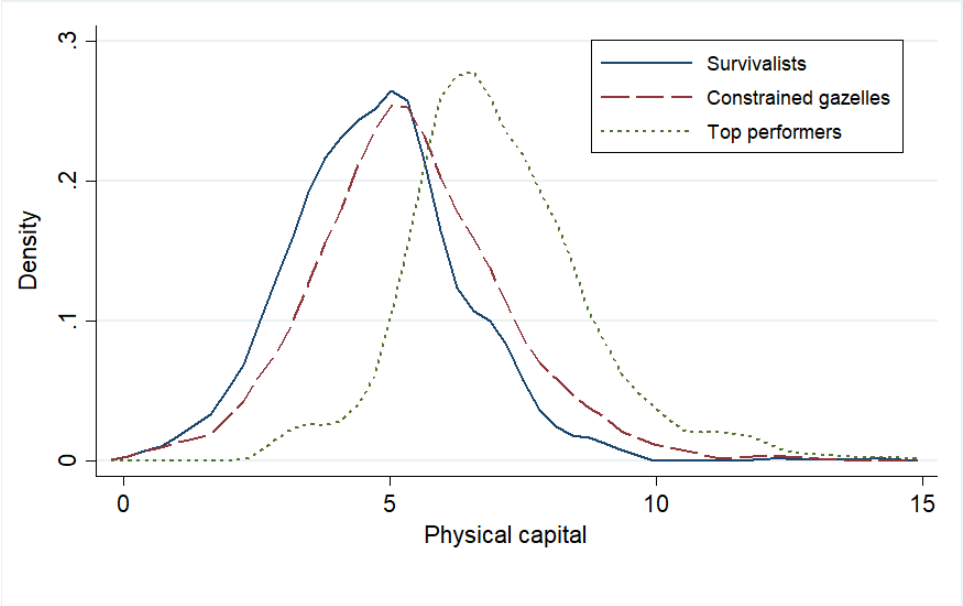

Figure 2: Distribution of physical capital by segment

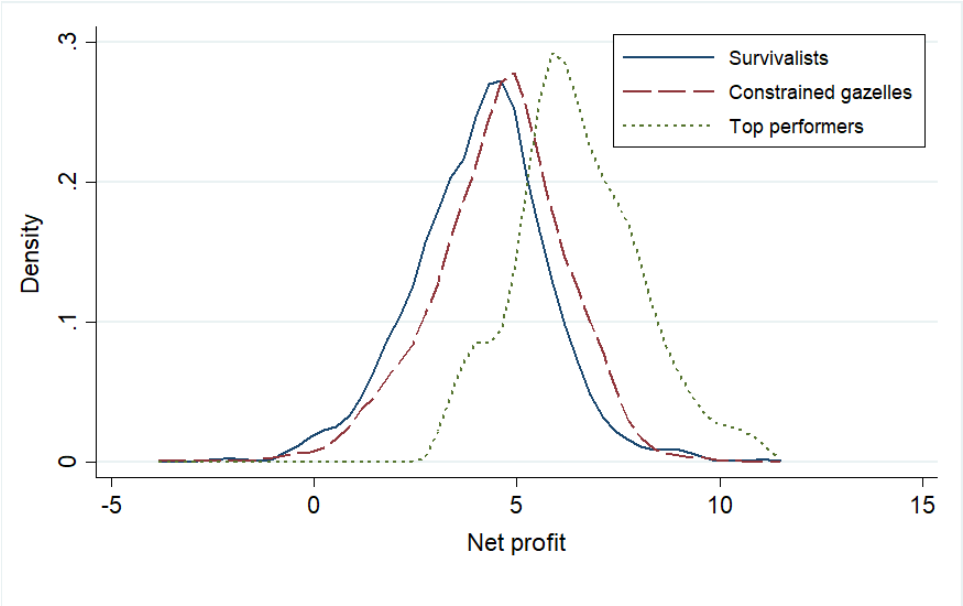

Figure 3: Distribution of net profit by segment

5.1.2 Segmentation and degree of informality

In Table 5, we present the distribution between the partition into the three segments and the levels of informality identified based on a five criteria index[19][20]. Overall, the index shows that the majority of the surveyed firms (54%) operate completely informally, i.e., without any registration with the authorities, payment of taxes, or formal accounts. Conversely, only 2% meet all the formality criteria, making them formal firms. Between them, 29% of the firms surveyed operate at a high level of informality (complying with only one or two of the criteria) and 15% at a low level of informality (complying with 3 or 4 of the criteria). A strong correlation stands out between the segments and the level of informality. Most survivalists (68%) are totally informal, confirming their high level of vulnerability. Conversely, top performers are only 30% in this case, highlighting that higher economic performances are associated with a lower level of informality. 5% of top performers are even formal firms, regrouping the two tiers of formal firms in the sample. Constrained gazelles are in an intermediate position, with half of them operating completely informally.

Table 5: Distribution of firms according to the level of informality

	(1)	(2)	(3)	(4)
	Survivalists	Constrained gazelles	Top performers	All
Totally formal	0.1	1.2	5.5	1.6
Low level of informality	8.3	16.2	28.8	15.6
High level of informality	23.2	31.5	31.4	28.9
Totally informal	68.5	51.2	30.4	53.9

Notes: Totally informal firms meet zero criteria, high informal firms meet one or two criteria, low informal firms meet three or four criteria, and formal firms meet all criteria.
Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

[19]The five criteria are: (1) being registered with a local authority or municipality, (2) being registered at general registrar, (3) paying local or municipal taxes (tax stamps), (4) being registered for national VAT or sales tax, and (5) keeping accounts according to national or regional standards.

[20]Distribution of firms by score is presented in Table A2.

5.1.3 Robustness check: deductive versus inductive approaches

Finally, we test the robustness of our segmentation by comparing it with that resulting from a Hierarchical Clustering on Principal Components (HCPC) computed on a large set of entrepreneurs' and firms' characteristics. The partition in three clusters reveals an upper tier of firms (25%) corresponding to our segment of top performers, another lower tier of firms (45%) corresponding to our segment of survivalists, and an intermediate segment (30%) corresponding to our segment of constrained gazelles.

Globally, it appears that the correlation between our main segmentation and the alternative measure is relatively high, as the classifications, although based on significantly different methodology and variables, assign the same segment to 63% of the sampled firms. Table 6 reports the transition matrix for firms between the two segmentation computing by the deductive and inductive approaches. It shows that 84% of top performers are also assigned to the upper tier cluster, demonstrating that our identification of the most successful firms is quite robust to the methodology used[21]. Our identification of survivalists is also robust, as 70% of them are identified in the lower tier of firms by the cluster analysis.

For the constrained gazelle, we observe less correspondence between the two partitions (45%). It is probably due to the specific definition of constrained gazelles in the methodology proposed by Grimm et al. (2012), which aims to identify a group of entrepreneurs with similar entrepreneur characteristics with top performers but displaying low capital accumulation and economic performance as survivalists. In this sense, our main segmentation identifies constrained gazelles according to their potential to become top performers based on a set of predetermined factors. Conversely, cluster analysis identifies an intermediate segment with specific observable characteristics clearly distinct from other clusters. Then, the fact that 39% of our constrained gazelles are assigned to the lower tier of firms in the cluster analysis

[21]The proportion of top performers identified by the cluster analysis (25%) confirms the pertinence of our choice to identify 20% of top performers in the deductive approach instead of only 10%.

is logical, as the latter observes the actual performance of firms to classy them and not their predicted economic potential.

Table 6: Transition matrix between deductive and inductive segmentations (%)

	Main segmentation - deductive approach			
Inductive approach	Survivalists	Constrained gazelles	Top performers	Total
Lower tier	70	39	7	45
Intermediate	25	45	9	30
Upper tier	5	16	84	25
Total	40	40	20	100

Notes: Column percentages are reported.
Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

5.2 Firm's and entrepreneur's characteristics differences between segments

Table 7 presents entrepreneur and business characteristics for the three segments identified. Above all, the three groups differ in their economic performance, with top performers being effectively the most successful businesses. Their average monthly sales in US$ PPP are more than six times higher than those of the survivalists and constrained gazelles. This gap widens further with value-added and net profit, with the average values of the top performers being eight times higher than those of the other segments. These observations are consistent with significant differences in the firm's size. While a significant proportion of survivalists (42%) and constrained gazelles (40%) do not own any physical assets of value, top performers are by far the most capital-intensive segment. These differences between the segments are a mechanical consequence of the method used to identify the top performers, which considers the stock of physical capital as a size criterion and net profit as a performance criterion.

As we identify constrained gazelles based on a set of predetermined factors, they share similar sector occupation and socio-demographic characteristics with top performers. Indeed,

Table 7: Comparison of groups by firm and entrepreneur characteristics

	(1)	(2)	(3)	(4)	(5)
	Survivalists	Constrained gazelles	Top performers	t-Test (1) vs. (2)	t-Test (2) vs. (3)
Entrepreneurs' characteristics					
Man	0,35	0,61	0,59	0***	0.429
Woman	0,62	0,29	0,29	0***	0.923
Joint venture	0,03	0,10	0,12	0***	0.157
Age	38,1	38.1	38,2	0.852	0.939
No education	0,66	0,13	0,20	0***	0***
Primary education	0,25	0,38	0,34	0***	0.147
Secondary education	0,07	0,29	0,27	0***	0.365
Tertiary education	0,02	0,20	0,18	0***	0.315
No other opportunity	0,67	0,38	0,39	0***	0.596
Firms' economic performance					
Sales	1066	1205	8187	0.591	0***
Value-added	782	558	4727	0.519	0***
Net profit	343	316	2558	0.736	0***
Firms' characteristics					
# of full-time paid workers	1.5	2.1	2.9	0***	0***
Any full-time paid worker	0.73	0.59	0.49	0***	0***
Average physical capital	1354	1254	12471	0.922	0***
No capital	0.42	0.40	0	0.304	0***
Urban	0.48	0.74	0.72	0***	0.372
Manufacture	0.10	0.14	0.14	0.004***	0.944
Service	0.10	0.25	0.26	0***	0.565
Trade	0.80	0.61	0.60	0***	0.642
Age of firms	5.8	7.8	8.1	0***	0.464
Access to electricity	0.46	0.64	0.83	0***	0***
Registration (local level)	0.21	0.33	0.55	0***	0***
Registration (national level)	0.06	0.14	0.26	0***	0***
Taxes payment (local level)	0.27	0.43	0.63	0***	0***
Taxes payment (national level)	0.25	0.42	0.60	0***	0***
Keep accounts	0.36	0.51	0.66	0***	0***
Separate finance	0.24	0.42	0.54	0***	0***
Bank account	0.14	0.27	0.46	0***	0***
Trade network's characteristics					
Any supplier	0.21	0.21	0.12	0.717	0***
One supplier	0.31	0.27	0.18	0.007***	0***
Two or three suppliers	0.32	0.33	0.40	0.309	0.002***
More than three suppliers	0.16	0.19	0.29	0.018**	0***
Formal partners	0.42	0.53	0.69	0***	0***
Businesses as customers	0.13	0.20	0.35	0***	0***
Not locally located trading partners	0.06	0.11	0.22	0***	0***
Very reliable suppliers	0.15	0.20	0.27	0.001***	0.001***
Relational contracting	0.48	0.50	0.56	0.177	0.018**
Long relationship	0.29	0.44	0.49	0***	0.015**
Weekly communication with suppliers	0.47	0.50	0.57	0.088*	0.004***
N	1312	1329	659		

Notes: *p < 0.05,**p < 0.01, ***p < 0.001.

Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

they are more represented in the manufacturing and service sectors than the survivalists. Findings also show that only 29% of top performers and constrained gazelles are exclusively run by women, while this figure rises to 62% among survivalists. This is in line with previous literature that found that women are underrepresented among top performers (Grimm et al., 2012). Two tiers of survivalists have no education compared to a minority of top performers and constrained gazelles, which is consistent with the exclusion from the formal labor market traditionally attributed to survivalists. Indeed, findings confirm that the survivalists are entrepreneurs by necessity, as 67% declare that they started their activity because they would have been unemployed otherwise. However, a significant proportion of top performers and constrained gazelles also make this claim, highlighting the importance of entrepreneurship in earning a living in Sub-Saharan Africa.

Table 7 shows other characteristics of each segment that we have not used to identify the segments. Structural business characteristics are consistent with these performance and capital accumulation disparities. The majority of the survivalist (73%) and constrained gazelles (59%) are in reality self-employed workers as they do not employ any full-time paid workers other than the owner. In contrast, the top performers are more likely to be employers with at least one additional full-time paid worker, but still exhibit a high proportion of self-employed (49%). Moreover, the most successful firms are also the oldest and those with the highest access to public utilities. Indeed, almost all top performers (83%) have access to electricity, while only 46% of survivalists and 64% of the constrained gazelles benefit from it. This difference in access to utilities probably illustrates significant disparities in operating conditions, with survivalists mainly conducting petty trading activities in markets or in precarious premises, while the top performers, and to a lesser extent the constrained gazelles, are better established.

Therefore, the three segments contain entrepreneurs with different socio-demographic characteristics and structurally different businesses. It is reflected in entrepreneurial behavior that differs, particularly in how they conduct their business and comply with regulations.

Top performers often keep financial records more than others, even if this is just simple bookkeeping. This management behavior is often accompanied by the strict separation of business accounts from personal finances. Moreover, as described above, the level of informality at which entrepreneurs in each segment conduct their business is also a meaningful distinction. Over half of the top performers are registered with local or national authorities, and almost two-thirds are subject to taxation. These differences in the level of informality between segments are also correlated with the level of access to banking services. Almost half of the top performers have a bank account, while only 14% and 27% of the survivalists and constrained gazelles have one, respectively.

Finally, the firms' trade network characteristics differ in each segment. Top performers integrate more complex value chains than the survivalists, who are limited to retailing locally supplied goods to nearby inhabitants. In addition to having locally based suppliers, the top performers also buy their goods and raw materials more often from suppliers further away geographically. The legal status of the suppliers with whom businesses deal also varies by segment. More than two-thirds of the top performers deal with formal businesses as suppliers, compared to 42% for survivalists and 53% for constrained gazelles. Almost all the firms in the sample have households as customers. However, the top performers and constrained gazelles are more involved in business-to-business relationships than survivalists. Although the main customers of the businesses are located in the neighborhood, whatever the segment, the top performers are more concerned with exports.

5.3 Informal heterogeneity and digital technologies

Table 8 reports ICT equipment and usage by segments. Descriptive statistics suggest significant access inequalities. Top performers have the greatest access to ICT, with only 8% of them having no ICT equipment compared to 18% of constrained gazelles and 31% of survivalists. We observe these inequalities of access for all the devices considered, with top

Table 8: Comparison of groups by ICT equipment and usage

	(1)	(2)	(3)	(4)	(5)
	Survivalists	Constrained gazelles	Top performers	t-Test (1) vs. (2)	t-Test (2) vs. (3)
Ownership of ICT devices					
Mobile phone	0.69	0.82	0.91	0***	0***
Landline	0.01	0.04	0.09	0***	0***
Computer	0.01	0.06	0.16	0***	0***
Number of ICT devices owned					
Any	0.31	0.18	0.08	0***	0***
One	0.68	0.76	0.75	0***	0.721
Two	0.02	0.05	0.11	0***	0***
Three	0	0.02	0.06	0***	0***
ICT usage for external coordination					
Bilateral coordination with suppliers	0.36	0.49	0.66	0***	0***
Bilateral coordination with customers	0.33	0.50	0.65	0***	0***
Multilateral coordination	0.06	0.13	0.25	0***	0***
ICT usage for financial transactions					
Online or mobile banking	0.04	0.09	0.020	0***	0***
Mobile money to send money	0.13	0.24	0.36	0***	0***
Mobile money to receive money	0.13	0.21	0.36	0***	0***
ICT usage for internal coordination					
Use of software	0	0.02	0.09	0***	0***
Number of ICT functions used					
Any	0.51	0.31	0.15	0***	0***
One	0.17	0.19	0.17	0.117	0.370
Two	0.19	0.24	0.20	0.002***	0.020**
Three	0.06	0.11	0.16	0***	0.002***
Four	0.05	0.10	0.15	0***	0***
Five	0.02	0.04	0.12	0.008***	0***
Six	0	0.01	0.06	0.001***	0***
Seven	0	0	0.01	0.160	0.001***
N	1312	1329	659		

Notes: *p < 0.05,**p < 0.01, ***p < 0.001.
Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

performers being the segment with the highest rate of mobile phone, landline, and computer equipment. Constrained gazelles are almost as well equipped as the top performers regarding mobile phones, the most widespread equipment. The possession of other ICTs mainly concerns these two segments, owning a landline or a computer, concerning very few survivalists. Thus, inequalities in access mainly affect survivalists, although constrained gazelles demonstrate also lower access to ICT than top performers.

Looking at the second-level of digital inequalities, survivalists remain the most digitally excluded firms since only 49% of them use ICT for at least one of the functions defined

above. The use of digital technologies for business purposes is much more widespread among constrained gazelles (69%) and top performers (85%). These differences in usage rates are likely to be explained by differences in equipment rates. However, the transition from access to actual usage appears not equally important for each segment. Indeed, while 86% of top performers with access to ICT equipment use it for business purposes, this figure drops to 76% among constrained gazelles and falls to 58% for survivalists.

The gaps in usage between the different segments are wide for each function of the digital technologies considered. Using ICTs for bilateral coordination with suppliers or customers is almost ubiquitous among top performers, with 79% exhibiting such behavior. While constrained gazelles also broadly use this type of function (64%), this is not the case for survivalists, whose use of bilateral coordination is less widely adopted (44%). The use of social media, or the internet in general, for multilateral coordination purposes concerns few firms, with only one-quarter of top performers and 13% of constrained gazelles. These advanced practices, such as online advertising, selling products and services online, or searching for information, are used by only 6% of survivalists.

The financial uses of digital technologies through mobile money or online banking services are also less widespread among survivalists. Only 17% of them receive or send money via mobile money for business purposes, compared to 29% and 43% among constrained gazelles and top performers, respectively. While they are already the most likely to have a bank account, top performers are also those who benefit the most from financial technology innovations. Not surprisingly, the few software users for inventory management, accounting, or performance monitoring are mostly top performers (68%).

5.4 Digital divides among informal firms

5.4.1 Determinants of digital technologies usage

Table 9 reports the results of the multivariate analysis, which estimates the probability of digital technologies usage for business purposes on observable entrepreneurs' and firms' characteristics in the eight African countries. In this table, we show the average marginal effects[22], which display the average change in predicted probabilities if the explaining variable is changed by one unit. The first column presents the average marginal effects obtained for the whole sample.

Table 9: Determinants of digital technologies usage for business purposes

	(1)	(2)	(3)	(4)
	All	Survivalists	Constrained gazelles	Top performers
Gender of the owner (ref. Man)				
Woman	-0.067***	-0.010***	-0.031	0.002
	(0.016)	(0.029)	(0.026)	(0.032)
Both	0.005	-0.068	0.035	0.024
	(0.029)	(0.073)	(0.039)	(0.058)
Education of owner (ref. None)				
Primary	0.035*	0.056	0.108***	0.015
	(0.019)	(0.035)	(0.037)	(0.031)
Secondary	0.060**	0.135**	0.089*	0.021
	(0.024)	(0.057)	(0.047)	(0.053)
Tertiary: Diploma /Certificate	-0.012	-0.076	0.030	0.024
	(0.028)	(0.081)	(0.047)	(0.058)
Age of the firm	0.002	-0.002	0.003	0.007*
	(0.001)	(0.003)	(0.003)	(0.004)
Age of the firm (squared)	-2.48e-05	2.75e-05	2.54e-05	-0.000136**
	(1.79e-05)	(4.01e-05)	(7.65e-05)	(6.13e-05)
Log monthly profits	0.019***	0.014	0.019**	0.001
	(0.005)	(0.010)	(0.008)	(0.017)
Number of full-time paid workers (Ref. None)				
1 or 2 full time workers	0.023	0.040	0.033	0.025
	(0.020)	(0.031)	(0.030)	(0.033)
3 or more full time workers	0.059*	0.004	0.091*	0.076
	(0.033)	(0.062)	(0.051)	(0.058)
Sector of activity (Ref. Manufacture)				
Service	-0.038	-0.004	-0.063*	-0.060

(continued on next page)

[22]Table A3 presents the underlying estimated coefficients.

	(1) All	(2) Survivalists	(3) Constrained gazelles	(4) Top performers
	(0.024)	(0.052)	(0.033)	(0.040)
Selling/Trading	-0.097***	-0.060	-0.119***	-0.143***
	(0.020)	(0.040)	(0.032)	(0.030)
Urban location	0.016	-0.007	-0.002	0.015
	(0.017)	(0.028)	(0.028)	(0.028)
Level of informality (Ref. Totally informal)				
High level of informality	0.099***	0.122***	0.092***	0.054*
	(0.019)	(0.034)	(0.026)	(0.031)
Low level of informality	0.0558**	-0.0139	0.155***	-0.0631
	(0.028)	(0.051)	(0.043)	(0.042)
Totally formal (omitted)	-	-	-	-
Access to electricity	0.054***	0.073**	0.059***	0.037
	(0.016)	(0.029)	(0.023)	(0.035)
Bank account	0.093***	0.070**	0.094***	0.070**
	(0.022)	(0.035)	(0.029)	(0.030)
Not keeping accounts	-0.068***	-0.078**	-0.039	-0.062**
	(0.020)	(0.034)	(0.024)	(0.029)
Separate finance	-0.024	0.008	-0.019	-0.015
	(0.016)	(0.031)	(0.023)	(0.030)
No other opportunity	-0.038***	-0.026	-0.062***	-0.018
	(0.015)	(0.029)	(0.023)	(0.025)
B2B relations	0.096***	0.141***	0.062*	0.070*
	(0.023)	(0.039)	(0.036)	(0.039)
Formal partners	0.062***	0.058**	0.063***	0.049*
	(0.015)	(0.024)	(0.024)	(0.028)
Not local trading partners	0.110***	0.280***	0.064	0.098**
	(0.033)	(0.072)	(0.048)	(0.045)
Number of ICT devices owned (Ref. Zero)				
1	0.168***	0.148***	0.228***	0.071
	(0.021)	(0.028)	(0.034)	(0.053)
2	0.290***	0.312***	0.321***	0.155**
	(0.042)	(0.114)	(0.058)	(0.062)
3	0.397***	-	0.392***	-
	(0.082)		(0.076)	
Region effects	Yes	No	No	No
Country effects	No	Yes	Yes	Yes
Observations	3191	1287	1297	601

Notes: *p < 0.05,**p < 0.01, ***p < 0.001. Robust standard errors are in parentheses.

Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

Results find some entrepreneurs' socio-demographic characteristics in the determinants of digital technologies usage for business purposes. In line with previous research, our findings show that firms led by women are 6.7% points less likely than firms led by men to use at least one function of digital technologies for business purposes. Education attainment is a significant driver too, as primary and secondary education increase the odds of using digital technologies for business purposes by 3.5% and 6% compared to entrepreneurs with no education. Although most entrepreneurs with tertiary education use digital technologies, no significant differences appear between them and entrepreneurs without education after controlling for other entrepreneurs' and firms' characteristics. No significant association is found between firms' digitization and urban location, which may be explained by the fact that the model includes region effects to control for cross-region differences in terms of telecommunication infrastructures and urbanization rate, for example. All these findings show that some socio-demographic determinants of the digital divide usually identified in Sub-Saharan Africa at the individual level, such as gender and education attainment, are diffused into the productive sphere and the sub-population of informal entrepreneurs[23]. Beyond the fact that this indicates that personal and professional usages of digital technologies are intrinsically linked, it invites us to deepen our understanding of the digital divide at the firm level by exploring the role of entrepreneurial attributes and firm characteristics.

Indeed, they appear to be related to the second-digital divide among informal firms as they have an impact on the odds of using at least one function of digital technologies. Congruent with findings about the role of household income in the literature, our results show that higher monthly profits are associated with digital technologies usage. Although the relation is endogenous, as the use of digital technologies may enhance firms' economic outcomes, it highlights that successful firms potentially have a greater ability to pay for

[23]We do not include the age of the entrepreneur in the model due to missing data for all observations in Rwanda. However, when estimating this model for the whole sample without Rwanda, a negative association is found between the age of the entrepreneur and the use of digital technologies, as entrepreneurs over 40 years old are significantly less likely to use them than younger ones (15-30 years old).

digital devices and services, and a greater need and willingness to use them. A significant association is found between digital technologies usage and firm size, as firms with three or more full-time paid workers are more likely to use digital technologies for business purposes than self-employed entrepreneurs.

Older businesses are not associated with a higher probability of usage, while the manufacturing sector's activities are more associated with digital technologies usage than trade activities (-9.7% points). These results highlight some sectoral specificities, as the nature of the economic activity probably shapes the professional usage of digital technologies due to specific organizational modes, business relationships, and sales channels. Moreover, types of business's partners appear to play a role in the firms' second-level digital divide. Having formal partners upstream or downstream of the production is associated with a rise of the odds of using digital technologies by 6.2% points, while having businesses as customers increases these odds by 9.6% points. It highlights some potential network effects in the diffusion of digital technologies, as having formal partners and expanding its customer base to businesses implies a greater need to adopt these technologies. Having business partners located further away than the surrounding villages or towns increases the likelihood of using digital technologies by 11% points. Finally, access to electricity appears to increase the probability of using digital technologies for business purposes by 5.4% points, confirming the importance of access to utilities and the quality of business premises.

Entrepreneurial behavior also plays a substantial role in the second-level digital divide among informal firms in Sub-Saharan Africa. The use of digital technologies for business purposes is more associated with opportunity entrepreneurs, as entrepreneurs who stated that they started their business due to the lack of other employment opportunities are 3.8% points less likely to be digital technology users. No significant association is found with the strict separation of business accounts from personal finances, but not keeping records decreases the firm's probability to use digital technologies by 6.8% points. Financial inclusion is a significant determinant of second-level digital divide among firms, as having a bank

account is significantly correlated with a higher probability of using these technologies, with a rise of 9.3% points. Finally, the level of informality remains a major determinant of digital inclusion. Compared to totally informal firms, all others are more likely to use digital technologies for business purposes.

All these findings are controlled by the number of ICT devices that firms possess. Not surprisingly, ICT device access is strongly correlated with using such technologies for business purposes, confirming the sequential process of digital divides. Moreover, it appears that the cumulative ownership of ICT devices is strongly associated with being a user. Compared to firms who do not own any ICT device, the probability of using digital technologies rises by 16.8% points for ownership of only one device, 29% points for ownership of two devices, and 39.7% for the combined ownership of a mobile phone, landline phone, and computer.

Within each segment, the second-level digital divide appears to follow specific patterns but also share similar factors (Table 10, columns 2, 3, and 4). Hence, the ownership of ICT devices, low level of informality, and financial inclusion through access to a bank account significantly increase the probability of using digital technologies for business purposes among the three segments. Similarly, characteristics of firms' trade networks are significantly associated with the use of digital technologies among the three segments, highlighting potential network effects on the adoption of digital technologies usage. However, access to electricity appears to be a significant determinant only among survivalists and constrained gazelles, which probably face higher external constraints than top performers. Similarly, entrepreneurs' characteristics are not associated with the second-level digital divide among top performers, these characteristics being only significant drivers among the most vulnerable informal segments. While the educational level is significant for both constrained gazelles and survivalists, the gender gap in the use of digital technologies for business purposes is only significant for survivalists. For constrained gazelles and top performers, some structural characteristics of the firm are involved in the second-level digital divide. Sector specificities appear for both, while the firm size and profit level are significant only for constrained

gazelles, and firm age is significant only for top performers. Finally, not keeping accounts decreases the probability of being a digital technology user only for survivalists and top performers.

5.4.2 Decomposition analysis of usage gaps between informal segments

Finally, to further analyze the digital technologies usage gap between the three segments, we perform a multivariate decomposition for nonlinear response models (Yun, 2004; Powers et al., 2014). Table 7 reports the detailed decomposition results of the comparison between survivalists and top performers (column 1) and between constrained gazelles and top performers (column 2). The decomposition aims to explain the sources of the observed usage rates across segments, that is 36% points between top performers and survivalists, and 16% points between top performers and constrained gazelles. These can be attributed to endowment and coefficient components, and within these components, to specific factors.

In the first column, we find that the observed difference in the use of digital technologies for business purposes between survivalists and top performers is mainly due to endowment effects (85%), i.e., differences in the characteristics of firms and entrepreneurs in the two segments. The coefficient effects explain only about 15% of the usage gap[24]. Clearly, the usage gap between top performers and survivalists is not mainly due to inequalities in material access, as bringing the mobile phone ownership of survivalists up to the level of top performers would only reduce the gap by 8%. Similarly, gender and education attainment differences explain a part of the endowment effect, but entrepreneurs' socio-demographic characteristics are not the main drivers of digital inequalities between the top performers and survivalists. Hence, other important factors are involved in the endowment effect, highlighting that digital inequalities between top performers and survivalists are mainly driven by disparities in levels

[24]Results of the coefficient effects decomposition are available upon request.

of economic inclusion, and trade network's characteristics. Indeed, shifting the survivalist's distributions on bank account ownership, level of informality, and electricity access to top performers levels would decrease the second-level digital divide between the two segments by more than 19% to the total. Above higher economic inclusion, the nature of the firm's value chain, through having formal partners, businesses as customers, or trading partners farther away than the surrounding villages and towns, appears to explain a large part of this gap (21% in total). It highlights the importance of potential network effects in the adoption of the use of digital technologies for business purposes, but also the varying usefulness of digital technologies across firms according to the level of complexity of their value chain. Hence, external constraints and environment play an important role in the uneven diffusion of digital technologies among the most successful informal firms and subsistence entrepreneurs. However, the endowment effects associated with secondary education and bookkeeping emphasize that internal constraints associated with digital and entrepreneurial skills may also partly explain the second-level digital divide between top performers and survivalists.

The second column shows that the usage gap between constrained gazelles and top performers is also largely explained by endowment effects (87%). Disparities in economic inclusion explain a significant part of the observed difference in digital technologies usage between the two segments, with the level of informality, access to electricity, and access to a bank account remaining important factors. Indeed, bringing the distribution of these variables for survivalists up to the level of the top performers' distributions, would reduce the usage gap between the two segments by about 34%. Moreover, equalizing the profit level of constrained gazelles to that of the top performers is expected to reduce the usage gap between the two segments by about 21.4%. The first-level digital divide also contributes to the usage gap, with ownership of mobile phones (11%) and computers (5.3%) contributing significantly to the endowment effect. It highlights that, more than for survivalists, material access and affordability of digital technology services are key factors in the usage inequalities between the constrained gazelles and the top performers. This is further supported by the fact that

Table 10: Decomposition of the usage gap between the top performers and the other segments

	(1) Survivalists vs. Top performers (%)	(2) Constrained gazelles vs. Top performers (%)
Prevalence of digital technologies usage		
Top performers	85	85
Survivalists / Constrained gazelles	49	69
Difference	36	16
Difference in characteristics	85.16	86.6
Women	7.9***	-0.1
No education	2.9	-2.5**
Primary education	0.8	-0.8***
Secondary education	4.9***	-0.7
Tertiary education	-4.4**	-0.3
Age of firm	-1.4	0.3
Monthly profits (log)	8.2	21.4***
Any paid workers	0.5	2.6
1 or 2 paid workers	0.5	-0.04
3 or more paid workers	-0.36	2.9**
Manufacturing sector	0.2	-0.2***
Service sector	0.7	-0.02
Trade/retail sector	2.1**	-0.01***
Urban	-0.6	0.002
Registration	2.5	7.0*
Payment of taxes	7.1**	9.6***
Electricity	6.8***	6.9***
Bank account	5.4**	10.1***
No opportunity	1.8	-0.9***
Separate finance	0.8	-1.2
Bookkeeping	5.7***	3.0
Businesses as customers	7.3***	5.1*
Formal partners	3.8**	5.7***
Not local partners	9.7***	3.5
Mobile phone	8.1***	10.8***
Landline phone	-	-
Computer	2.6	5.3*
Number of observations	1934	1936

Notes: *p < 0.05,**p < 0.01, ***p < 0.001.
Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

entrepreneurial attributes do not appear to play a role in these usage inequalities between the two segments, as constrained gazelles and top performers share similar socio-demographic characteristics and entrepreneurial behavior. Finally, the trade network's characteristics remain significant drivers in the endowment component of the usage gap with top performers, but to a less extent than for survivalists.

6 Conclusion

Major and ongoing investments in telecommunication infrastructure have greatly contributed to the diffusion of digital technologies in developing countries, particularly in Sub-Saharan Africa. However, despite similar physical access and network coverage, intra-country disparities persist in terms of material access and usage patterns. While these digital inequalities, or digital divides, between households or individuals have been largely studied, there is little evidence at the firm-level in Sub-Saharan Africa. Given that digital technologies have great potential to increase the productivity of informal workers and businesses by addressing market and state failures, identifying the determinants of digital inequalities among informal firms seems to be of prior importance.

In this paper, we explore digital inequalities in access and usage among 3,300 firms and entrepreneurs from eight sub-Saharan African countries. Acknowledging that the dichotomy between formal and informal firms is insufficient to assess appropriately economic activities in sub-Saharan Africa, we rely on a deductive approach proposed by Grimm et al. (2012) to classify the sampled businesses into three segments: a lower tier of survivalists, an intermediate segment of constrained gazelles, and an upper tier of top performers. Digital inequalities in access and usage between the identified segments are studied in a multidimensional way. Access is observed through the possession of the most common ICT devices in SSA (i.e., mobile phone, landline phone, and computer). We consider seven different usage functions to analyze the use of digital technologies for business purposes, divided between external coordination, financial transaction and internal coordination functions. We use multivariate analysis to identify the determinants of the use of digital technologies for business purposes among entrepreneur and firm characteristics, first for all informal firms and then for each segment. Next, in order to identify the main drivers of digital inequalities between top performers and the other segments, a multivariate decomposition for non-linear response models decomposes the digital technologies usage gap between segments.

Our findings suggest that digital technologies are already widely used by informal entrepreneurs in Sub-Saharan Africa, probably reshaping their business practices. However, the diffusion of these new technologies appears to be uneven within the heterogeneous set of informal firms. Digital inequalities seem to be rooted in the already existing socioeconomic vulnerabilities and inequalities. Despite the benefits attributed to the adoption of digital technologies by businesses, their uneven spread risks widening existing inequalities rather than creating new opportunities for inclusion. Indeed, this new technological dimension adds to the sources of exclusion already experienced by subsistence entrepreneurs, and is likely to dig the gap between the most successful entrepreneurs and entrepreneurs in the intermediate segment which seem to have an untapped economic potential. If material access is a prior condition for usage, policies intervention should focus on other important drivers of digital inequalities as access do not imply effective usage. Indeed, access to mobile phones is almost ubiquitous but second-digital divide remains strong in terms of usage and diversity of usage. Policy interventions should prevent the digital exclusion of the already most excluded by including digital skills in business training support, and find ways to make digital services affordable for all. Moreover, digital-based policy interventions for informal enterprises must also acknowledge that selection bias may occur due to digital inequalities, most vulnerable people being inaccessible or unable to benefit interventions from their own digital devices.

Finally, the adoption of the different functions of digital technologies is far from reaching saturation, with the use of the Internet and advanced business software still in its early stages. Future research on the evolution of usage over time using panel data would allow to observe the dynamics of adoption in different segments of the informal sectors in sub-Saharan Africa. Furthermore, while this paper explores inequalities in access and use, further empirical evidences on tangibles benefit of digital technologies usage on informal firms' performance are needed.

References

A4AI (2021). Alliance for Affordable Internet. A POLICY GUIDE: TOWARDS MEANINGFUL CONNECTIVITY. Technical report.

Abebe, G., Assefa, B., Gebre-eyesus, M., and Degu, T. (2018). Identifying dynamic and constrained entrepreneurs in low income countries: evidence from Ethiopia. Technical report.

Adeleke, R. (2021). Digital divide in Nigeria: The role of regional differentials. *African Journal of Science, Technology, Innovation and Development*, 13(3):333–346.

Adoho, F. M. and Doumbia, D. (2022). Informal Sector Heterogeneity and Income Inequality: Evidence from The Democratic Republic of Congo. *Journal of Economic Development*, 47(4):55–77.

Aker, J. C. and Blumenstock, J. E. (2015). The economic impacts of new technologies in Africa. *The Oxford handbook of Africa and economics*, 2:354–371.

Aker, J. C. and Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of economic Perspectives*, 24(3):207–232.

Apeti, A. E. and Edoh, E. D. (2023). Tax revenue and mobile money in developing countries. *Journal of Development Economics*, 161:103014.

Atiyas, and Dutz, M. A. (2021). Digital Technology Uses among Informal Micro-Sized Firms.

Baller, S., Dutta, S., and Lanvin, B. (2016). The global information technology report 2016. Technical report, World Economic Forum.

Benhassine, N., McKenzie, D., Pouliquen, V., and Santini, M. (2018). Does inducing informal firms to formalize make sense? Experimental evidence from Benin. *Journal of Public Economics*, 157(November 2017):1–14.

Benjamin, N., Mbaye, A. A., and Diop, I. T. (2012). *The informal sector in Francophone Africa: firm size, productivity, and institutions*. World Bank Publications.

Berkel, H. and Tarp, F. (2022). Informality and Firm Performance in Myanmar. *Journal of Development Studies*, 58(7):1363–1382.

Berrou, J.-P., Combarnous, F., Eekhout, T., and Mellet, K. (2020). My mobile, my market. *Reseaux*, 219(1):105–142.

Besley, T. and Persson, T. (2014). Why do developing countries tax so little? *Journal of economic perspectives*, 28(4):99–120.

Bhattacharya, R. (2019). ICT solutions for the informal sector in developing economies: What can one expect? *Electronic Journal of Information Systems in Developing Countries*, 85(3):1–7.

Birba, O. and Diagne, A. (2012). Determinants of adoption of Internet in Africa: Case of 17 sub-Saharan countries. *Structural change and economic dynamics*, 23(4):463–472.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, pages 436–455.

Bonnet, F., Vanek, J., and Chen, M. (2019). Women and men in the informal economy: A statistical brief. *International Labour Office, Geneva*, 20.

Bruhn, M. and McKenzie, D. (2014). Entry Regulation and the Formalization of Microenterprises in Developing Countries. *The World Bank Research Observer*, 29(2):186–201.

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

Campos, F., Goldstein, M., Mckenzie, D. J., and Campos, F. M. L. (2015). Short-term impacts of formalization assistance and a bank information session on business registration and access to finance in Malawi. *World Bank policy research working paper*, (7183).

Cardona, M., Kretschmer, T., and Strobel, T. (2013). ICT and productivity: Conclusions from the empirical literature. *Information Economics and Policy*, 25(3):109–125.

Cariolle, J. (2020). Digital spillovers and SMEs' performance in Sub-Saharan Africa. *FERDI Notes brèves/Policy briefs*.

Cariolle, J., Le Goff, M., and Santoni, O. (2019). Digital vulnerability and performance of firms in developing countries.

Chacaltana, J., Leung, V., and Miso, L. (2018). New Technologies and the Transition to Formality: The Trend Towards e–Formality. (January):1–38.

Chandler, D. (1995). Technological or media determinism.

Choi, J., Dutz, M. A., and Usman, Z. (2020). *The future of work in Africa: Harnessing the potential of digital technologies for all*. World Bank Publications.

Chuttur, M. (2009). Overview of the technology acceptance model: Origins, developments and future directions.

Cirera, X. (2016). ICT Use, Innovation, and Productivity: Evidence from Sub-Saharan Africa. *SSRN Electronic Journal*, (October).

Cirera, X., Comin, D., Cruz, M., and Lee, K. M. (2021). Firm-Level Adoption of Technologies in Senegal.

Commander, S., Harrison, R., and Menezes-Filho, N. (2011). ICT and productivity in developing countries: New firm-level evidence from Brazil and India. *Review of Economics and Statistics*, 93(2):528–541.

Cunningham, W. V. and Maloney, W. F. (2001). Heterogeneity among Mexico's microenterprises: An application of factor and cluster analysis. *Economic Development and Cultural Change*, 50(1):131–156.

Danquah, M. and Owusu, S. (2021). Digital technology and productivity of informal enterprises: Empirical evidence from Nigeria. Technical report.

De Giorgi, G., Ploenzke, M., and Rahman, A. (2018). Small firms' formalisation: The stick treatment. *The Journal of Development Studies*, 54(6):983–1001.

De Giorgi, G. and Rahman, A. (2013). SME's Registration: Evidence from an RCT in Bangladesh. *Economics Letters*, 120(3):573–578.

De Mel, S., McKenzie, D., and Woodruff, C. (2013). The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5(2):122–150.

Deen-Swarray, M. (2016). Toward digital inclusion: understanding the literacy effect on adoption and use of mobile phones and the internet in Africa. *Information Technologies International Development*, 12(2):pp–29.

Deen-Swarray, M., Moyo, M., and Stork, C. (2013). ICT access and usage among informal businesses in Africa. *info*.

Demenet, A., Razafindrakoto, M., and Roubaud, F. (2016). Do informal businesses gain from registration and how? Panel data evidence from Vietnam. *World Development*, 84:326–341.

Demirguc-Kunt, A., Klapper, L., Singer, D., and Ansar, S. (2018). *The Global Findex Database 2017: Measuring financial inclusion and the fintech revolution*. World Bank Publications.

DiMaggio, P., Hargittai, E., Celeste, C., and Shafer, S. (2004). Digital inequality: From unequal access to differentiated use. *Social inequality*, pages 355–400.

Eekhout, T., Berrou, J., and Combarnous, F. (2022). Entrepreneurs' mobile phone appropriation and technical efficiency of informal firms in Dakar (Senegal). *Journal of International Development*.

Esselaar, S., Stork, C., Ndiwalana, A., and Deen-Swarray, M. (2006). ICT usage and its impact on profitability of SMEs in 13 African countries. In *2006 International Conference on Information and Communication Technologies and Development*, pages 40–47. IEEE.

Fafchamps, M. and Quinn, S. (2018). Networks and manufacturing firms in Africa: Results from a randomized field experiment. *The World Bank Economic Review*, 32(3):656–675.

Floridi, A., Demena, B. A., and Wagner, N. (2020). Shedding light on the shadows of informality: A meta-analysis of formalization interventions targeted at informal firms. *Labour Economics*, 67:101925.

Forenbacher, I., Husnjak, S., Cvitić, I., and Jovović, I. (2019). Determinants of mobile phone ownership in Nigeria. *Telecommunications Policy*, 43(7):101812.

Fuchs, C. and Horak, E. (2008). Africa and the digital divide. *Telematics and Informatics*, 25(2):99–116.

Gaglio, C., Kraemer-Mbula, E., and Lorenz, E. (2022). The effects of digital transformation on innovation and productivity: Firm-level evidence of South African manufacturing micro and small enterprises. *Technological Forecasting and Social Change*, 182(June):121785.

Galperin, H. and Arcidiacono, M. (2021). Employment and the gender digital divide in Latin America: A decomposition analysis. *Telecommunications Policy*, 45(7):102166.

Gillwald, A., Milek, A., and Stork, C. (2010). Gender assessment of ICT access and usage in Africa. *Towards Evidence-based ICT Policy and Regulation*, 1(5).

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

Grimm, M., Knorringa, P., and Lay, J. (2012). Constrained Gazelles: High Potentials in West Africa's Informal Economy. *World Development*, 40(7):1352–1368.

Grimm, M., van der Hoeven, R., and Lay, J. (2011). Unlocking Potential.

GSMA (2021). The Mobile Economy Sub-Saharan Africa 2021. Technical report.

Gunkel, D. J. (2003). Second thoughts: toward a critique of the digital divide. *New media society*, 5(4):499–522.

Hafkin, N. J. and Taggart, N. (2001). *Gender, information technology, and developing countries: An analytic study*. Office of Women in Development, Bureau for Global Programs, Field Support ....

Hargittai, E. (2002). Second-level digital divide.

Hasbi, M. and Dubus, A. (2020). Determinants of mobile broadband use in developing economies: Evidence from Sub-Saharan Africa. *Telecommunications Policy*, 44(5):101944.

Heeks, R. (2022). Digital inequality beyond the digital divide: conceptualizing adverse digital incorporation in the global South. *Information Technology for Development*, pages 1–17.

Hilbert, M. (2011). The end justifies the definition: The manifold outlooks on the digital divide and their practical usefulness for policy-making. *Telecommunications Policy*, 35(8):715–736.

Husson, F., Josse, J., and Pages, J. (2010). Principal component methods-hierarchical clustering-partitional clustering: why would we need to choose for visualizing data. *Applied Mathematics Department*, 17.

Islam, A., Muzi, S., and Rodriguez Meza, J. L. (2018a). Does mobile money use increase firms' investment? Evidence from Enterprise Surveys in Kenya, Uganda, and Tanzania. *Small Business Economics*, 51(3):687–708.

Islam, A. M. and Amin, M. (2022). The Gender Labor Productivity Gap across Informal Firms.

Islam, M. M., Habes, E. M., and Alam, M. M. (2018b). The usage and social capital of mobile phones and their effect on the performance of microenterprise: An empirical study. *Technological Forecasting and Social Change*, 132:156–164.

ITU (2017a). ICT Prices 2017. Technical report, International Telecommunication Union.

ITU (2017b). Measuring the Information Society Report 2017. Technical report, International Telecommunication Union.

ITU (2021). Digital trends in Africa 2021.

Jack, W. and Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review*, 104(1):183–223.

Jessen, J. and Kluve, J. (2021). The effectiveness of interventions to reduce informality in low-and middle-income countries. *World Development*, 138:105256.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement*, 20(1):141–151.

Kanbur, R. (2017). Informality: Causes, consequences and policy responses. *Review of Development Economics*, 21(4):939–961.

La Porta, R. and Shleifer, A. (2014). Informality and development. *Journal of economic perspectives*, 28(3):109–126.

Lavallée, E. and Roubaud, F. (2019). Corruption in the informal sector: evidence from West Africa. *The Journal of Development Studies*, 55(6):1067–1080.

Lythreatis, S., Singh, S. K., and El-Kassar, A.-N. (2021). The digital divide: A review and future research agenda. *Technological Forecasting and Social Change*, page 121359.

Mascagni, G. and Nell, C. (2022). Tax Compliance in Rwanda: Evidence from a Message Field Experiment. *Economic Development and Cultural Change*, 70(2):587–623.

Moosa, D. (2019). Exploring heterogeneity of micro and small enterprises in Morocco. *Dalal Moosa*, page 141.

Moshi, G. C. and Mwakatumbula, H. J. (2017). Effects of political stability and sector regulations on investments in African mobile markets. *Telecommunications Policy*, 41(7-8):651–661.

Mothobi, O., Gillwald, A., and Aguera, P. (2020). A demand side view of informality and financial inclusion. *Policy Paper*, (9).

Motohashi, K. (2008). IT, enterprise reform, and productivity in Chinese manufacturing firms. *Journal of Asian Economics*, 19(4):325–333.

Mumporeze, N. and Prieler, M. (2017). Gender digital divide in Rwanda: A qualitative analysis of socioeconomic factors. *Telematics and Informatics*, 34(7):1285–1293.

Negrete-García, A. K. (2018). Constrained potential: A characterisation of Mexican microenterprises.

Nguimkeu, P. and Okou, C. (2021). Leveraging digital technologies to boost productivity in the informal sector in Sub-Saharan Africa. *Review of Policy Research*, 38(6):707–731.

Norris, P. (2003). Digital divide: Civic engagement, information poverty, and the Internet worldwide.

NTIA (1999). National Telecommunication and Information Administration. Falling through the net: Defining the digital divide. Technical report.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, pages 693–709.

Ochoa, R. G., Lach, S., Masaki, T., and Rodríguez-Castelán, C. (2022). Mobile internet adoption in West Africa. *Technology in Society*, 68:101845.

Okunogbe, O. and Santoro, F. (2022). Increasing Tax Collection in African Countries.

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

Penard, T., Poussing, N., Mukoko, B., and Tamokwe Piaptie, G. B. (2015). Internet adoption and usage patterns in Africa: Evidence from Cameroon. *Technology in Society*, 42:71–80.

Pénard, T., Poussing, N., Zomo Yebe, G., and Ella, N. (2012). Comparing the determinants of internet and cell phone use in Africa: evidence from Gabon. *Communications Strategies*, (86):65–83.

Powers, D. A., Yoshioka, H., and Yun, M.-S. (2011). mvdcmp: Multivariate decomposition for nonlinear response models. *The Stata Journal*, 11(4):556–576.

Ragnedda, M. (2019). Conceptualising the digital divide. *Mapping digital divide in Africa: A mediated analysis*, pages 27–44.

Ragnedda, M. and Muschert, G. W. (2013). *The digital divide*. Routledge Florence, KY.

Ragnedda, M., Ruiu, M. L., and Addeo, F. (2022). The self-reinforcing effect of digital and social exclusion: The inequality loop. *Telematics and Informatics*, 72(June):101852.

Research ICT Africa (2020). RIA ICT Access Survey 2017-2018 [dataset].

RIA (2018). Research ICT Africa. ICT HH and Business Survey Field Manual. Technical report.

Riggins, F. and Dewan, S. (2005). The Digital Divide: Current and Future Research Directions. *Journal of the Association for Information Systems*, 6(12):298–337.

Roy, P. and Khan, M. H. (2021). Digitizing Taxation and Premature Formalization in Developing Countries. *Development and Change*, 52(4):855–877.

Scheerder, A., Van Deursen, A., and Van Dijk, J. (2017). Determinants of Internet skills, uses and outcomes. A systematic review of the second-and third-level digital divide. *Telematics and informatics*, 34(8):1607–1624.

Schoar, A. (2010). The divide between subsistence and transformational entrepreneurship. *Innovation policy and the economy*, 10(1):57–81.

Srinuan, C. and Bohlin, E. (2011). Understanding the digital divide: A literature survey and ways forward. *22nd European Regional Conference of the International Telecommunications Society (ITS2011)*, page 39.

Suri, T. (2017). Mobile money. *Annual Review of Economics*, 9:497–520.

Suri, T. and Jack, W. (2016). The long-run poverty and gender impacts of mobile money. *Science*, 354(6317):1288–1292.

Tang, Y. K. and Konde, V. (2020). Differences in ICT use by entrepreneurial micro-firms: evidence from Zambia. *Information Technology for Development*, 26(2):268–291.

The Broadband Commission for Sustainable Development (2019). Connecting Africa Through Broadband. A strategy for doubling connectivity by 2021 and reaching universal access by 2030. Technical report.

Thierer, A. (2000). How free computers are filling the digital divide. *Heritage Foundation Backgrounder*, 1361:1–21.

van Deursen, A. and Van Dijk, J. A. G. M. (2010). Measuring internet skills. *International journal of human-computer interaction*, 26(10):891–916.

van Deursen, A. J. and van Dijk, J. A. (2014). The digital divide shifts to differences in usage. *New Media and Society*, 16(3):507–526.

Van Deursen, A. J. A. M. and Van Dijk, J. A. G. M. (2019). The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New media society*, 21(2):354–375.

Van Dijk, J. and Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *The information society*, 19(4):315–326.

van Dijk, J. A. (2006). Digital divide research, achievements and shortcomings. *Poetics*, 34(4-5):221–235.

Van Dijk, J. A. G. M. (2005). *The deepening divide: Inequality in the information society*. Sage publications.

Van Dijk, J. A. G. M. (2017). Digital divide: Impact of access. *The international encyclopedia of media effects*, pages 1–11.

Williams, C. C., Shahid, M. S., and Martínez, A. (2016). Determinants of the level of informality of informal micro-enterprises: Some evidence from the city of Lahore, Pakistan. *World Development*, 84:312–325.

World Bank (2016). *World development report 2016: Digital dividends*. World Bank Publications.

World Bank (2017). World Development Indicators. Technical report, Washington, DC. World Bank.

Yun, M.-S. (2004). Decomposing differences in the first moment. *Economics letters*, 82(2):275–280.

Appendix

Table A1: Distribution across segments by country

	Kenya	Mozambique	Ghana	Nigeria	Rwanda	South Africa	Tanzania	Senegal
Survivalists	0.27	0.36	0.46	0.39	0.45	0.38	0.44	0.41
Intermediate	0.53	0.44	0.34	0.41	0.35	0.42	0.36	0.39
Top performers	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20

Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

Table A2: Distribution of firms according to their informality index score

	(1)	(2)	(3)	(4)
Score	Survivalists	Constrained gazelles	Top performers	All
5	0.001	0.01	0.05	0.02
4	0.03	0.08	0.14	0.07
3	0.05	0.08	0.14	0.08
2	0.11	0.16	0.21	0.15
1	0.12	0.16	0.15	0.14
0	0.68	0.51	0.30	0.54

Notes: The five criteria are: (1) being registered with a local authority or municipality, (2) being registered at general registrar, (3) paying local or municipal taxes (tax stamps), (4) being registered for national VAT or sales tax, and (5) keeping accounts according to national or regional standards.
Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

Table A3: Determinants of digital technologies usage for business purposes

	(1) All	(2) Survivalists	(3) Constrained gazelles	(4) Top performers
Gender of owner (ref. Man)				
Woman	-0.497***	-0.615***	-0.233	0.0217
	(0.122)	(0.178)	(0.195)	(0.340)
Both	0.0391	-0.419	0.275	0.269
	(0.228)	(0.447)	(0.314)	(0.671)
Education of owner (ref. None)				
Primary	0.259*	0.345	0.802***	0.153
	(0.140)	(0.218)	(0.260)	(0.314)
Secondary	0.451**	0.835**	0.652*	0.217
	(0.182)	(0.353)	(0.340)	(0.564)
Tertiary: Diploma /Certificate	-0.0891	-0.483	0.216	0.254
	(0.207)	(0.533)	(0.335)	(0.617)
No other opportunity	-0.292***	-0.163	-0.480***	-0.197
	(0.112)	(0.183)	(0.175)	(0.270)
Age of the firm	0.0111	-0.0141	0.0194	0.0734*
	(0.0104)	(0.0173)	(0.0243)	(0.0386)
Age of the firm (squared)	-0.000189	0.000173	0.000196	-0.00145**
	(0.000136)	(0.000253)	(0.000589)	(0.000659)
Log monthly profits	0.144***	0.0871	0.146**	0.0156
	(0.0392)	(0.0613)	(0.0599)	(0.181)
Number of full time paid workers (Ref. None)				
1 or 2 full time workers	0.172	0.246	0.247	0.260
	(0.148)	(0.191)	(0.226)	(0.348)
3 or more full time workers	0.455*	0.0227	0.725*	0.903
	(0.261)	(0.386)	(0.422)	(0.794)
Level of informality (Ref. Totally informal)				
High level of informality	0.731***	0.744***	0.660***	0.618*
	(0.139)	(0.196)	(0.189)	(0.350)
Low level of informality	0.401**	-0.0852	1.193***	-0.589
	(0.199)	(0.317)	(0.369)	(0.410)
Totally formal (omitted)	-	-	-	-
Access to electricity	0.409***	0.463**	0.454**	0.394
	(0.124)	(0.183)	(0.178)	(0.377)
Separate finance	-0.185	0.0481	-0.144	-0.162
	(0.124)	(0.193)	(0.174)	(0.321)
Bank account	0.710***	0.440**	0.722***	0.747**
	(0.166)	(0.222)	(0.224)	(0.334)
No bookkeeping	-0.518***	-0.489**	-0.297	-0.658**
	(0.152)	(0.217)	(0.186)	(0.317)
Sector of activity (Ref. Manufacture)				
Service	-0.304	-0.0245	-0.533*	-1.006
	(0.194)	(0.323)	(0.291)	(0.697)
Selling/Trading	-0.751***	-0.375	-0.963***	-1.889***
	(0.161)	(0.249)	(0.283)	(0.588)
Urban location	0.122	-0.0462	-0.0119	0.157
	(0.133)	(0.174)	(0.212)	(0.303)
B2B relations	0.734***	0.890***	0.474*	0.744*
	(0.172)	(0.247)	(0.276)	(0.418)
Formal partners	0.471***	0.364**	0.488***	0.527*
	(0.119)	(0.156)	(0.183)	(0.292)
Not local trading partners	0.836***	1.763***	0.495	1.051**
	(0.251)	(0.455)	(0.370)	(0.502)
Number of ICT devices owned (Ref. Zero)				
1	1.159***	0.902***	1.515***	0.662
	(0.142)	(0.173)	(0.220)	(0.463)
2	2.129***	1.896***	2.291***	1.770**
	(0.348)	(0.730)	(0.500)	(0.741)

(continued on next page)

	(1) All	(2) Survivalists	(3) Constrained gazelles	(4) Top performers
3	3.299***	-	3.070***	-
	(1.117)		(0.901)	
Constant	0.776**	0.0581	-0.169	2.596*
	(0.389)	(0.567)	(0.642)	(1.431)
Region effects	Yes	No	No	No
Country effects	No	Yes	Yes	Yes
Pseudo R2	0.376	0.303	0.347	0.324
Observations	3191	1287	1297	601

Notes: *p < 0.05,**p < 0.01, ***p < 0.001. Robust standard errors in parentheses.
Source: Author's computations based on After Access Surveys, RIA, 2017-2018.

BxWP2023-03

BSE UMR CNRS 6060

Université de
Bordeaux
Avenue Léon
Duguit, Bât.
H 33608
Pessac,
France

Tel : +33 (0)5.56.84.25.75
http://bse.u-bordeaux.fr/

Derniers numéros – Last issues

- 2023-02 **Retombées du très haut débit sur les entreprises: Quels effets sur les performances des entreprises en France ?** *by Florian LEON & Laurent BERGE & Chloé DUVIVIER*
- 2023-01 **Les enjeux maritimes de l'Afrique coloniale française** *by Hubert BONIN*
- 2022-24 **Trains of Thought: High-Speed Rail and Innovation in China** *by Georgios TSIACHTSIRAS & Deyun YIN& Ernest MIGUELEZ &Rosina MORENO*
- 2022-23 **Experienced versus Decision Utility: Large-Scale Comparison for Income-Leisure Preferences** *by Alpaslan AKAY & Olivier BARGAIN & H. Xavier JARA*
- 2022-22 **The Economic Gains of closing the Employment Gender Gap: Evidence from Morocco** *by Olivier BARGAIN & Maria LO BUE*
- 2022-21 **Dynastic Measures of Intergenerational Mobility** *by Olivier BARGAIN & Maria C. LO BUE & Flaviana PALMISANO*
- 2022-20 **Liquidity matters when measuring bank output** *by Raphaël CHIAPPINI & Bertrand GROSLAMBERT & Olivier BRUNO*
- 2022-19 **Distributional National Accounts for Australia, 1991-2018** *by Matthew FISHER-POST & Nicolas HÉRAULT & Roger WILKINS*
- 2022-18 **French utilities committed to globalization (19th-21st centuries)** *by Hubert BONIN*

Ernest MIGUELEZ is the scientific coordinators of the Bordeaux Economics Working Papers. The layout and distribution are provided by Cyril MESMER.