UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2021/114

Digital technology and productivity of informal enterprises

Empirical evidence from Nigeria

Michael Danquah[1] and Solomon Owusu[2]

July 2021

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: The lingering policy dilemma facing many governments in sub-Saharan Africa in recent years is what can be done in the short to medium term to boost the output and incomes of individuals and enterprises in the informal sector, given the size and persistence of the sector in the region. In this paper we examine the structural impact of access and usage of digital technology by informal enterprises on labour productivity. Using a sample of non-farm informal enterprises in Nigeria, we employ IV LASSO techniques to carry out our analysis. The structural parameters of our IV LASSO estimates show that labour productivity is significantly higher for enterprises that use digital technology than for non-users. Further analysis reveals that benefits arise more strongly in larger enterprises in the upper segment of the informal sector. Our findings have key implications for the ongoing discussion on the role of digital technology and government regulatory and policy frameworks for ICT in the region.

Key words: digital technologies, informal sector, productivity, IV LASSO, Nigeria

JEL classification: D22, O17, O33, O55

---

[1] UNU-WIDER, corresponding author: danquah@wider.unu.edu; [2] UNU-MERIT, Maastricht, Netherlands, and German Development Institute, Bonn, Germany

This study has been prepared within the UNU-WIDER project Transforming informal work and livelihoods.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-054-2

https://doi.org/10.35188/UNU-WIDER/2021/054-2

Typescript prepared by Joseph Laredo.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

The importance of the informal sector to Africa's economy has been widely emphasized. Over the years, the informal sector has remained large and has continued to function as a key part of the structure of the economies in sub-Saharan Africa (SSA). The informal sector in SSA accounts for more than 40 per cent of output and 85.8 per cent of employment in the region (African Development Bank 2013; ILO 2018).[1] However, the sector is largely characterized by small and inefficient firms, run by poorly educated entrepreneurs, and is often branded as unproductive (La Porta and Shleifer 2011, 2014; Mohammad and Islam 2015). In a region where employment opportunities in the formal sector are limited, the informal sector remains a significant source of income to support growth and development. Yet, African governments over the years have been facing a persistent policy dilemma of how to tackle the challenges of the sector, boost productivity, and create decent jobs and livelihoods for the poor.

In many African countries, the dominant policy over the years has centred on the rollout of a proliferation of formalization policies—formal sector private jobs for the poor—in a bid to formalize the informal sector. These prohibitively high-cost interventions are intended to stimulate or induce the transition of informal firms and workers to the formal sector. These pro-formalization reforms have involved using a range of incentives and enforcement tools such as reducing registration costs, providing tax and banking support, and investing in business advisory services and training. Yet, despite the expensive formalization programmes implemented, post-reform formalization rates have been minimal (Benhassine et al. 2018; Bruhn and McKenzie 2014; Campos et al. 2015; Choi et al. 2020; Grimm et al. 2012; Lince 2011).

The minimal impact of formalization reforms in developing economies, especially in Africa, is evident in the persistence and composition of the informal sector after decades of pro-formalization policy implementation. Ulyssea (2010) attributed the low formalization to the fact that the majority of informal businesses are low-productivity survivors, which are therefore likely to stay informal even when the formal sector's entry costs are eliminated. Evidence from Africa also shows that formalization has no material and significant effect on firm productivity (Campos et al. 2015; Dioa et al. 2018; Lince 2011), thus raising the possibility that even firms that have taken advantage of policy to formalize may revert to the informal sector. Following from the fact that many development institutions still advocate for formalization policies (Ohnsorge et al. 2021; World Bank 2019), albeit from a global and long-term viewpoint, the question that remains for countries in SSA is: What policy interventions can be put in place in the short to medium term, given the scale and persistence of the sector, to at least boost the output and incomes of these informal enterprises?

Increasing the productivity of enterprises in the informal sector through pro-productivity policy interventions such as leveraging ICT and associated low-skill-biased digital technologies, rather than wholesale formalization in the short to medium term, is beginning to feature prominently in the policy discussion about how to transform the sector (Campos and Gassier 2017; Choi et al. 2020; Cusolito and Maloney 2018). It is posited that a sensible pathway in the short to medium term in Africa would be to refocus policies and reforms on improving the productivity of small-scale, low-productivity firms and enhance the skills of informal workers using ICT. There is the likelihood that in the long run, more productive informal enterprises may seek to formalize as they grow and realize the gains of being formal. Context-specific and appropriate digital technologies

[1] This includes agricultural employment. We focus on non-farm informal enterprises in our study.

can present prospects to boost the productivity of enterprises operating informally. Digital technologies can enable owners of informal enterprises and their workers not only to learn but also to gain access to improved inputs and larger markets, obtain adequate capital and credit, and reduce transaction risk in the day-to-day running of their activities (Aker 2008; Aker and Mbithi 2010; Choi et al. 2020; Ongori and Migiro 2010). For instance, ICT can increase the proximity of businesses to markets and facilitate the creation of new and efficient ways of reaching customers and communicating with suppliers. This enhanced connection between informal enterprises and customers through digital technologies can bolster SSA economies.

The good news is that Africa is becoming more connected and digital development is growing at an increasing pace. Between 2013 and 2018, the share of the population in the region using the internet increased from 13 per cent to 30 per cent (Bayen 2018; Hjort and Poulsen 2019). The level of human capital in the region is not a binding constraint to the adoption of technology from the world technological frontier (Danquah and Ouattara 2014; Knoche and Huang 2012). Given that the predominant activities of informal enterprises are mostly in retail trade, manufacture of wearing apparel, and manufacture of food and beverages, the increasing digitization of the region can play a significant role in enhancing the value chain of informal businesses, unlock their potential, and increase their efficiency.

Some studies have examined the effect of innovation on firm growth and employment among informal enterprises in SSA (e.g. Avenyo et al. 2019; Fu et al. 2015). However, other than anecdotal evidence from case studies (Choi et al. 2020; Dean-Swarray et al. 2013; Nguimkeu and Okou 2020), there has until now been no study in Africa that has provided rigorous empirical evidence on how digital technologies could impact the productivity of informal enterprises. Specifically, we propose that access to and usage of digital technology (henceforth DT) by an informal enterprise will positively impact its productivity. In this paper, we use two waves (2010 and 2012) of Living Standards Measurement Study (LSMS) panel survey data in Nigeria to provide robust empirical evidence to support the anecdotal studies. Such evidence should induce governments and development partners to pay more attention to the role of DT as an appropriate policy option in dealing with the challenges associated with the informal sector.

The country context for the study, Nigeria, is important given the size of the informal sector and its pertinent role in livelihoods and the economy. The informal sector in Nigeria is made up of more than 41.5 million enterprises, the majority of these being located in Lagos (about 11.5 per cent). These enterprises contribute around 50 per cent to the nation's GDP, making the informal sector a significant source of livelihood and economic growth (SMEDAN 2019). Current estimates show that 93 per cent of all employment in Nigeria is informal, with about 95 per cent of women and 90 per cent of men working in this sector. The majority of informal workers are own-account workers who operate in vulnerable working conditions with little or no access to social protection and are reliant on a daily income to feed their families (ILO 2018).

The connection to DT is also very important. For instance, Rohman and Bohlin (2011) have shown that increases in broadband penetration by 10 percentage points will increase Nigeria's GDP by 1 per cent. Following the liberalization of the telecoms sector, the major providers—MTN, Globacom, Etisalat, and Airtel—have invested billions of dollars in terrestrial fibre and 2G and 3G base stations. As a result, access to mobile networks and the internet has increased steadily, such that Nigeria is now one of the largest telecoms markets in Africa and the biggest internet user in the region (Economic Intelligence Unit 2016). Although these improvements in DT have aided Nigeria's economic diversification, the effects are unequally distributed across the country. The majority of the populace are still grappling to access the full benefits of DT due, among other factors, to a costly regulatory and policy framework, geographical challenges, and a poor connective infrastructure (Forenbacher et al. 2019). The disaggregation of national data by income levels shows substantial inequalities. For instance, for the 53.5 per cent of Nigerians who live on or below the poverty line ($1.90), the cost of a month's DT usage represents roughly 10 per cent of their annual income (Economic Intelligence Unit 2016). This indicates that access to and usage of DT remain beyond the means of many Nigerians and therefore that a huge proportion of the population are still unconnected. The case of Nigeria therefore provides important insights of relevance to the development of the informal sector and DT usage in other SSA countries.

Based on a sample of non-farm enterprises that are not registered with any government agency in Nigeria, we estimate the structural effect of usage of DT on labour productivity using the IV LASSO technique. This method allows us to deal with the issue of unconfoundedness in the selection of observables and unobservables, which confounds many econometric estimations. The findings show that usage of DT significantly boosts the labour productivity of informal enterprises. The structural parameters of the IV LASSO regressions show that labour productivity is between 65 and 77 per cent higher for enterprises that use DT than for enterprises that do not. Further analysis reveals that these benefits arise more strongly in larger enterprises in the upper-tier segment of the informal sector. These findings are highly relevant, particularly in shaping discussions on ICT policy towards the informal sector. The excessive taxes imposed on DT (especially the internet) by African governments in recent years may have severe consequences on access to and usage of DT by informal enterprises (Choi et al. 2020; Stork and Esselaar 2019; Stork et al. 2020). The study sends clear signals to policy-makers on the need to formulate vigorous policies to facilitate access to and usage of DT by informal enterprises and foster low-skill-biased technology development tailored to the needs of the sector.

The rest of the paper proceeds as follows. Section 2 presents a review of the relationship between DT and productivity. In Section 3, we introduce and explain how the IV LASSO methodology is employed in the analysis. The description and source of data are also contained in this section. A detailed discussion of the empirical results is presented in Section 4. The conclusions and policy implications of the study are in Section 5.

2 How does digital technology affect the productivity of informal enterprises?

Studies examining the impact of DT on economic activity have been proliferating in recent years. At the macro level, the growth-accounting literature has proposed a strong relationship between DT investment and productivity growth at the country level (Goldfarb and Tucker 2019). Solow (1987: 1) claimed that 'you can see the computer age everywhere except in the productivity statistics'. The Solow paradox persisted for many years, but it was finally addressed in the growth-accounting literature, where the overall effect of ICT on economywide productivity was examined. Jorgenson et al. (2008) and Bloom et al. (2010) found that the dramatic improvement in labour productivity growth in the US in the mid-1990s was largely driven by the massive DT investment boom and usage of the late 1980s. The authors also found not only that comparable EU economies had not experienced similar productivity growth but that the productivity growth gap between them had widened. While the productivity gap between the US and the EU in 1995 stood at 1.8 per cent, the gap had grown to 9.8 per cent by 2004—an increase that was largely driven by stronger productivity growth in the US ICT production and market services sectors. This divergence in productivity growth between the US and the EU raised new questions about the increasing economic importance of ICT.

Moving away from the macro analysis, there is a large and growing literature examining this relationship at the firm or micro level. With few exceptions, this literature has found evidence of a productivity-enhancing effect of DT. Using cross-sectional data on firms in New Zealand, Grimes et al. (2012) find that firms using digital technologies have 7–10 per cent higher labour productivity. Maliranta and Rouvinen (2006) found similar results in Finland. Using Norwegian firm-level data, Akerman et al. (2015) find that increasing DT availability by 10 percentage points raises firm output by 0.4 per cent. In contrast to these positive results, in Germany, Bertschek et al. (2013) find no evidence of a positive and significant effect of DT on labour productivity. This is confirmed in Haller and Lyons (2015) for Irish firms. In Italy, Colombo et al. (2013) demonstrate that the quality of DT provision and its contribution to the operations of firms plays an important role in the DT–productivity relation.

Factors such as firm size and age, organizational change, regulation, skills, and geography are identified as having a significant impact on this relationship (Acemoglu and Autor 2011; Brynjolfsson and Hitt 2003; Brynjolfsson and Saunders 2010; Draca et al. 2009). Using a large multi-country micro-firm-level panel dataset of 19,000 firms in 13 EU countries and a small sample of US firms over 11 years, Bloom et al. (2012) find that ICT increases productivity. However, the effect differs across countries and types of firms. The findings of the study show that, unlike other multinational firms, UK-based US multinational firms experienced the same productivity improvement as US firms based in the US, suggesting that US firms are better organized than their counterparts in other countries and this allows them to be efficient users of ICT (Goldfarb and Tucker 2019).

While studies examining the productivity effect of DT in advanced economies are widespread, in many developing countries, particularly in Africa, studies examining this relationship are very limited at the micro level and almost non-existent in the informal sector. Leaving aside papers that have examined the effect of innovation on firm growth and employment among informal enterprises in the region (Avenyo et al. 2019; Fu et al. 2015) and anecdotal evidence from case studies (Aker and Mbiti 2010; Blimpo and Owusu 2020; Choi et al. 2020; Dean-Swarray et al. 2013; Nguimkeu and Okou 2020), there are no studies that have undertaken rigorous empirical evidence-based research on the influence of DT on the productivity of informal enterprises in Africa.

We fill this gap in the literature by examining the impact of DT on the productivity of informal enterprises in Nigeria. DT can offer opportunities to enhance the productivity of informal non-farm enterprises through a number of channels and mechanisms. DT use can improve access to and use of information, reduce search costs, and improve the coordination and management of supply chains, which translates into greater market and productive efficiency among economic agents. By so doing, DTs increase the proximity of informal enterprises to markets and facilitate the creation of new and efficient ways for these enterprises to reach customers and communicate with suppliers (Aker and Mbiti 2010; Jensen 2007). We call this the *access to information, transaction cost, and efficiency channel*.

For example, given the prevalence of imperfect and asymmetric information on markets in SSA, enterprises in the region have for years relied on conventional search mechanisms such as personal travel and radio to find information in a variety of areas, including input and output prices, and potential buyers and sellers. As detailed in Aker and Mbiti (2010: 10), in Niger, for example,

> [P]ersonal travel requires transport and opportunity costs, which can be relatively high with a combination of long distances and poor roads. . . . [A]n average trip to a market located 65 kilometers away can take 2–4 hours roundtrip, as compared to a two-minute call.

Informal enterprises also rely heavily on radio to access information. However, radio generally provides a limited range of information. While newspapers represent an alternative source of market information, their usage is largely in urban areas, is costly, and is not accessible to uneducated people, many of whom are in the informal sector. DT has introduced a new search facility that offers several advantages including (potentially) reduced costs. Although DT requires an initial sunk cost investment, it should be relatively cheaper after this due to lower variable costs associated with its use compared with the cost of equivalent travel and other opportunity costs. DT also allows informal enterprises to obtain real-time information routinely, rather than waiting for newspapers to be published or radio programmes to be broadcast (Aker 2008).

Informal enterprises may be able to leverage the opportunities afforded by DT through access to information and reduced transaction costs to broaden their market reach. Informal enterprises using DT are able to expand their market reach by searching across bigger consumer and producer markets, sell in more markets, both local and international, and generate more market contacts (e.g. through social media) as compared with their non-DT counterparts (Aker 2008). We call this channel the *market access channel*. These factors collectively translate into gains in job creation, revenue generation, competitiveness, and productivity (Ongori and Migiro 2010).

Informal enterprises could also leverage the opportunities of DT to access quality intermediate inputs that were previously unavailable to them and learn about best practices through access to the wider local and international markets. Through this medium, informal enterprises are afforded opportunities for (social) learning and upskilling as they become exposed to the latest technologies and efficient means of production. This can in turn affect the rate of technology adoption and increase productivity through the introduction of new product varieties and the extension of existing production into both local and international market (Bas and Strauss-Kahn 2015; Ndubuisi and Owusu 2021; Pahl and Timmer 2020). We call this channel the *quality intermediate inputs and learning channel*.

In addition to providing informal enterprises with important links to markets and information about fair prices for their products, DT enables them to engage in proper financial bookkeeping and thereby to access loans and finance to expand their operations, innovate, and invest in production facilities and a skilled workforce (Blimpo and Owusu 2020). DT also improves information flow between members of a network, both locally and across international boundaries. The increased speed of information flows within the network allows informal enterprises the opportunity to obtain credits from multiple sources including from networks outside the country and to respond better to shocks in the domestic financial market while obtaining information about alternative investment opportunities (Aker and Mbiti 2010). We call this the *access to credit and risk reduction channel*. For example, Kenya's M-Pesa now functions as a fully fledged financial service, providing money transfer services and offering loans and savings to businesses without access to the formal banking network, the majority of which are in the informal sector (Mbiti and Weil 2011). Other examples include the use of mobile money transfers in Ghana and other African countries, which has helped many informal firms overcome constraints and obstacles in accessing finance for their operations.

3 Methodology and data

3.1 Description of IV LASSO technique

The success of empirical research seeking to estimate structural or causal effects depends on using the right sets of instruments and controls in order to deal with unconfoundedness in the selection of unobservables and observables, respectively. In this paper we use a machine-learning method—the IV LASSO—to perform this model selection and inference that remain valid following model selection. The IV LASSO method used in this study offers an approach to estimating structural parameters in the presence of many potential instruments and controls centred on techniques for estimating sparse high-dimensional models (Belloni et al. 2012, 2014; Chernozhukov et al. 2015). In this case, the high-dimensional technique is used to choose which instruments and control variables to use. The IV LASSO method depends on an approximate sparsity assumption and the usage of high-quality variable selection combined with the use of appropriate moment functions.

Following, and also using the notations of, Chernozhukov et al. (2015) (henceforth CHS) in the ensuing presentation, we discuss the IV model, estimation approach, and algorithms using the post-regularization and post-double selection (PDS) methods of CHS (2015) and Belloni et al. (2014), respectively.

We consider a linear IV model:

$$y_{it} = \alpha_0 d_{it} + x'_{it}\beta_0 + \varepsilon_{it} \tag{1}$$

$$d_{it} = x'_{it}\tau_0 + z'_{it}\delta_0 + u_{it} \tag{2}$$

where $E[(z'_{it}, x'_{it})'\varepsilon_{it}] = E[(z'_{it}, x'_{it})'u_{it}] = 0$. y_{it} is the outcome variable (labour productivity), d_{it} is the endogenous variable (usage of DT), α is the coefficient of interest, x_{it} is a p_n^x vector of the exogenous control variable, z_{it} is a p_n^z vector of instruments, and *n* is the sample size. p_n^x, p_n^z are a large set of intuitively chosen potential control variables and instruments.

We might have it that z_{it} and x_{it} are correlated so that z_{it} are only valid instruments after accounting for x_{it}. Specifically, we let $z_{it} = \prod x_{it} + \zeta_{it}$, for $\prod$ is a $p_n^z \times p_n^x$ matrix and ζ_{it} is a p_n^z vector of unobservables with $E[x_{it}\,\zeta_{it}] = 0$. Replacing this expression for z_{it} as a function of x_{it} into (2) and then further replacing into (1) offers a system for y_i and d_i that is contingent only on x_{it}:

$$y_{it} = x'_{it}\theta_0 + \rho_{it}^y \tag{3}$$

$$d_{it} = x'_{it}\vartheta_0 + \rho_{it}^d \tag{4}$$

where $E\left[x_{it}\,\rho_{it}^y\right] = 0$ and $E\left[x_{it}\,\rho_{it}^d\right] = 0$. As indicated, our model allows for a large number of instruments and a large number of controls in our settings. Also, in order to estimate a baseline model without instruments—that is, accounting only for unconfoundedness in the selection of observables—we accommodate an exogenous case for d_{it} by placing $p_n^z = 0$ and imposing the additional condition $E[d_{it}\,\varepsilon_{it}] = 0$.

Given that the dimensions of $\eta_0 = (\theta'_0, \vartheta'_0, \tau'_0, \delta'_0)'$ may be large or even larger than n (high dimensional parameter), instructive estimation and inference about α_0 is not possible without imposing restrictions on η_0. An important approach and structure that has been used in the literature is approximate sparsity of the high-dimensional linear model. Approximate sparsity enforces a restriction that only *S* variables amongst all of variables p_n^z, p_n^x, where *S* is much smaller than n, have associated coefficients that are dissimilar from 0, while allowing a non-zero approximation error $r_{p,i}$. Therefore, estimators for this model attempt to learn the identities of the variables with large non-zero coefficients, while concurrently estimating these coefficients. This approach allows the researcher to consider many variables and to use the data to learn which of the many variables are the most relevant. The set-up also accommodates the case where the researcher does not know a priori exactly which suitable variables—be they instruments or controls—should be included in a model. Under the assumption of sparsity[2], we can assume that:

$$\|\eta_0\|_0 \leq S_{n,} \qquad S_n^2 \log(p_n^z + p_n^x)^3 / n \rightarrow 0,$$

where $\|\eta_0\|_0$ is the ℓ_0 'norm' of η_0 and denotes the number of non-zero components of η_0. In this case, sparsity requires that among the $p_n^z + p_n^x$ observed variables, the number of variables with non-zero coefficients is small in relation to the sample size. The sparsity assumption then reduces the problem of estimating α_0 to a problem of deciding which instruments and controls to use in equations (1) and (2). The bigger issue in doing this is the likelihood of making serious variable selection mistakes. For instance, a variable may be considered as important when in actual fact it has a zero coefficient and therefore has no true predictive power, or a variable may be omitted from the model in spite of having a non-zero coefficient. Both kinds of errors may adversely affect post-model-selection estimators and inference for α. Whilst the spurious inclusion of irrelevant variables after being considered predictive from examining the data results in overfitting, the omission of relevant x variables leads to standard omitted variables bias. Again, when relevant z variables are omitted, one loses identification power. Although the spurious inclusion of variables that are not relevant can be avoided by the use of contemporary, principled data-mining methods such as LASSO with appropriate tuning parameters (see Belloni et al. 2012, 2014, 2015; Tibshirani 1996), safeguarding against this type of mistake comes at the cost of needing to acknowledge that the omission of relevant explanatory variables is likely to happen. This is because, while methods such as LASSO will correctly find strong predictors, it has been shown that such techniques have non-negligible likelihood of losing predictors with small but non-zero coefficients (Chernozhukov et al. 2015). The omission of such explanatory variables can have substantial effects on drawing inferences for parameters of interest such as α in our model (Leeb and Pötscher 2008). The IV LASSO method used in this paper overcomes this difficulty by centring the estimation and inference on techniques that are robust to this type of model selection error. The methodology used in this study estimates equations that are locally insensitive to this type of mistakes, called orthogonal moment functions (Belloni et al. 2015; Chernozhukov et al. 2015).

The moment condition in our IV model is given by:

$$M(\alpha_0 \colon \eta_0) = 0, \quad M(\alpha, \eta) \coloneqq E\,[\psi_i(\alpha, \eta)], \tag{5}$$

where $\psi_i(\alpha,\eta) = (\tilde{\rho}_{it}^y - \tilde{\rho}_{it}^d \alpha)\tilde{v}_{it}$ for $\eta \coloneqq (\theta', \vartheta', \tau', \delta')$, $\tilde{\rho}_{it}^y \coloneqq y_{it} - x_{it}'\theta$, $\tilde{\rho}_{it}^d \coloneqq d_{it} - x_{it}'\vartheta$, and $\tilde{v}_{it} \coloneqq x_{it}'\tau + z_{it}'\delta - x_{it}'\vartheta$. When we set $\eta = \eta_0$, we have $\tilde{\rho}_{it}^y = \rho_{it}^y = y_{it} - x_{it}'\theta_0$, $\tilde{\rho}_{it}^d = \rho_{it}^d = d_{it} - x_{it}'\vartheta_0$ and $\tilde{v}_{it} = v_{it} \coloneqq x_{it}'\tau_0 + z_{it}'\delta_0 - x_{it}'\vartheta_0 = \zeta_{it}'\delta_0$.

In this case, a small selection error will have little impact on the estimation of α_0 by stating that the following orthogonality condition holds:

$$\frac{\partial}{\partial \eta} M(\alpha_0, \eta)|_{\eta=\eta_0} = 0. \tag{6}$$

Put differently, missing the true value η_0 by a small amount does not nullify the moments condition. Thus estimators $\hat{\alpha}$ of α_0 based on the empirical analog of equation (5),

[2] For a detailed presentation of the generalization to approximate sparsity, see Chernozhukov et al. (2015).

$$\widehat{M}(\hat{\alpha}, \hat{\eta}) = 0 \quad (7)$$

with $\widehat{M}(\alpha, \eta) \coloneqq n^{-1}\sum_{i=1}^{n}[\psi_i(\alpha, \eta)]$, can be shown to be 'immunized' or 'orthogonalized' against small selection errors. A comprehensive general formulation of orthogonal moments functions for use in sparse high-dimensional models and a number of estimation and inference results are presented in Belloni et al. (2014).

It can be seen that using the empirical version of equation (5) to estimate α_0 is the same as employing the usual IV regression of ρ^y on ρ^dusing v as instruments. Following from this, CHS posit the following algorithm for estimating α_0 centred on the 'double-selection' strategy of Belloni et al. (2014). Here, we first do a LASSO or post-LASSO regression of d_{it} on x_{it}, z_{it} to obtain $\hat{\tau}$ and $\hat{\delta}$. Then we do a LASSO or post-LASSO regression of y_{it} on x_{it} to get θ. This is followed by another LASSO or post-LASSO regression of $\widehat{d_{it}} = x'_{it}\hat{\tau} + z'_{it}\hat{\delta}$ on x_{it} to get ϑ. Letting $\hat{\rho}^y_{it} \coloneqq y_{it} - x'_{it}\theta$, $\hat{\rho}^d_{it} \coloneqq d_{it} - x'_{it}\vartheta$ and $\hat{v}_{it} \coloneqq x'_{it}\tau + z'_{it}\delta - x'_{it}\vartheta$, we retrieve estimator $\hat{\alpha}$ from equation (7) by using standard IV regression of $\hat{\rho}^y_{it}$ on $\hat{\rho}^d_{it}$ with $\hat{v}_{it}$ as the instruments. Inference is performed on α_0 using $\hat{\alpha}$ or the related score statistics and conventional heteroscedasticity robust standard errors.

Again, we use the alternative algorithms following from the PDS strategy of Belloni et al. (2014) that would yield similar asymptotic properties. Following the double-selection strategy, we run a LASSO regression of d_{it}on x_{it} and z_{it}, followed by a LASSO regression of d_{it} on x_{it} and another LASSO regression of y_{it} on x_{it}. Then we form a 2SLS estimator using instruments selected in step one and controlling for the union of controls selected in the three LASSO steps. The precise statement and proof of the properties of $\hat{\alpha}$ obtained from these algorithms can be found in Belloni et al. (2014) and Chernozhukov et al. (2015).

To ensure that relevant z variables are not excluded, we use the weak identification robust inference as in Belloni et al. (2012, 2013) and confidence sets based on Chernozhukov et al. (2013) super score weak identification robust tests. The super score test of statistical significance of the instruments is computed following the IV LASSO estimation. Here, the null hypothesis is that the coefficient on endogenous regressor d_{it} is H_O= b (d_{it}) = 0. The rejection of the null indicates that the instruments are valid, that is, orthogonal to the true disturbance (Ahrens et al. 2018).

3.2 Description and source of data

The dataset used in this paper comes from the Nigeria Living Standards Measurement Survey (LSMS-ISA) database sourced from the World Bank. The survey covers seven[3] countries: Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda. The survey consists of three parts, namely, community, agriculture, and household questionnaires. The last component, the household questionnaire, captures information on household demographics, employment, non-farm enterprises, and other factors. Our study is related to the last component and focuses on non-farm enterprises.

The database allows comparisons between the countries included in the LSMS-ISA project. However, some shortcomings remain. For instance, as documented in Nagler (2015), while all seven countries include a section covering non-farm-enterprises, items (specific questions)

[3] There are eight partner countries in SSA that are part of the LSMS-ISA projects, namely, Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda.

included in the survey are not completely harmonized across all the countries, and in a number of cases, not all questions are included in all questionnaires. Moreover, questions that are asked in all or at least the majority of countries can contain different answer possibilities, also reflecting the specific country context. For example, questions about access to DT are more comprehensive in the Nigeria dataset. The questionnaires in the other countries ask different questions regarding access to DT, so that these are not perfectly comparable with those contained in the Nigerian dataset; hence our decision to use the Nigerian dataset for this study. Even in the Nigerian dataset, out of the four waves available (2010, 2012, 2015, 2019), we find consistency only in the 2010 and 2012 questionnaires as far as our variable of interest (DT access) is concerned, hence our focus on the years 2010 and 2012.

In our analysis, we use these two waves (2010, 2012) of the LSMS panel survey in Nigeria to provide robust empirical evidence of the effect of access to and usage of DT on the productivity of informal non-farm enterprises. We follow the approach in the extant literature (Aker and Mbiti 2010; Jensen 2007) to measure our main explanatory variable—access to and usage of DT. Access to and usage of DT is a dummy variable equal to 1 if an enterprise uses a mobile phone, computer, or the internet and 0 if otherwise. Section five of the household questionnaire asks whether a survey respondent has access to: a mobile phone (S5q8), computer (S5q11), internet (S5q14). Survey participants have only two possible responses: 'yes' or 'no'. We measure productivity as total sales per worker in naira.

In our estimation model, we control for an array of potential variables that influence both labour productivity and usage of DT by firms. These include variables that capture demographics, firm and owner characteristics, access to public infrastructure, finance, and networks, as well as locational variables.

Since the estimates of the effect of DT on productivity might be susceptible to reverse causality, i.e. more productive informal enterprises may have more resources to invest in DT, we apply an instrumental variables approach. With respect to instruments, we choose a set of potential and well targeted good instruments from the literature to select from. Under the exclusion restriction, the ideal set of chosen instruments using IV LASSO provides a high-quality prediction of the endogenous variable—usage of DT. Using IV LASSO with proper penalty parameters theoretically guarantees that any instruments selected are not spuriously correlated to the endogenous variable but have true explanatory power. This indicates that IV LASSO could select no instruments at all as there may be no set of variables with satisfactory predictive power to achieve the required standard. The potential instruments based on economic intuition from the literature include the history of owning a DT device, DT intensity per local government area, coverage of DT infrastructure per local government area, and average incidence of lightning strikes per local government area.

The history of an individual within an enterprise owning a DT device such as a smartphone or tablet has been used in the literature as an instrument for usage and diffusion of DT (Bertschek and Niebel 2016). In this case, the history of the respondent owning a DT device might be a good predictor for the extent to which the individual is using DT services. In this paper, we use the variation in the history of a respondent owning a DT device in the two years of the panel (2010 and 2012): ownership in both years, ownership in either 2012 or 2010, ownership in 2010 but not 2012, or no access to DT in both years. We measure this using a dummy variable equal to 1 if the respondent used DT in both 2010 and 2012 and zero if otherwise. We follow a similar approach to construct the second DT history variable for the second, third, and fourth strategy. This is argued in the literature to be a valid instrument, as the entire enterprise's productivity will not depend on the number of years a single respondent working in the enterprise has had access to DT.

Again, the prevalence of DT infrastructure promotes coverage and allows the diffusion and adoption of DT in an area (Hodler and Raschky 2017; Manacorda and Tesei 2020). This means that access to or usage of DT by individuals depends on the extent of coverage and availability of service in a particular area. Due to the lack of fixed phone lines and high-speed internet cabling in Africa, mobile phones are the most commonly used way to access DT, and GSM technology accounts for almost 100 per cent of mobile technology in Africa. In this paper, we use data on two variables to capture this. The first variable is the average local mobile phone coverage per local government area in Nigeria as a proxy for coverage of DT (Piet et al. 2009). This is sourced from the Global System for Mobile Communications (GSM) Association (Manacorda and Tesei 2020). Second, we compute the average DT intensity per local government area using our dataset. In the Nigerian LSMS household questionnaire, respondents are asked to state the number of mobile phones, computers, and internet subscriptions they have (section 5 S5q10, S5q13, and S5q16). We sum and average this at the local government level. The average DT intensity per local government area is used as a valid instrument in other studies (e.g. Bertschek and Niebel 2016). The variable has been proven to be a good predictor for the share of workers having access to DT within an enterprise as it reflects the diffusion of DT use by workers across the local government area, hence affecting the labour productivity of all enterprises within the local government area.

Finally, we also use the incidence of lightning strike intensity as an instrument. Here, the argument is that frequent electrostatic discharges during storms are known to damage digital infrastructure such as that used for mobile phones and negatively affect connectivity, thereby disrupting the demand for and supply of DT (Andersen et al. 2012; Manacorda and Tesei 2020). In this case, the intensity of lightning strikes in a local government area may be correlated with the usage and availability of DT by individuals in the area. We show that in Africa areas with higher-than-average incidences of lightning strikes display slow DT adoption rates. Data for this variable are sourced from Manacorda and Tesei (2020). The appendix contains detailed information on how the variable is measured. Table A1 reports the summary statistics and measurement of all the variables used in our final estimation.

4 Empirical results

In this section, we discuss in detail the results of the structural effect of access to and usage of DT on labour productivity. First, we discuss the findings for the CHS and PDS methods where we do not account for unconfoundedness in the selection of unobservables—that is, no instruments used—followed by the IV LASSO approach, where we introduce the instruments. Starting with the CHS, where the LASSO-selected controls are used to construct orthogonalized versions of labour productivity and our variable of interest, usage of DT, we present the orthogonalized versions based on the LASSO and post-LASSO estimated coefficients. The post-LASSO is OLS applied to LASSO-selected variables; it is convenient to implement and works well in terms of convergence and bias compared with the LASSO (see Belloni and Chernozhukov 2013; Belloni et al. 2012). In the CHS method, the selected high-dimensional controls for labour productivity and usage of DT are partialled out using the LASSO or post-LASSO coefficient. In both LASSO and post-LASSO, the structural parameters show that the usage of DT significantly impacts the labour productivity of firms in the informal sector in Nigeria (see Tables 1A and 1B). Enterprises using DT have a 65 and 47 per cent higher labour productivity than non-users using the LASSO and post-LASSO estimates, respectively.

With respect to the PDS, LASSO is used to select a set of variables that are helpful for explaining labour productivity, and a set of variables that are helpful for explaining usage of DT. We estimate the structural parameter α by OLS regression of labour productivity on the usage of DT and the union of the variables selected for explaining labour productivity and usage of DT. In this case, we use variables that are important for either of the two predictive relationships to guard against omitted variables bias when estimating α. The PDS method is equivalent to Frisch–Waugh–Lovell partialling-out of all selected controls; therefore, we can draw inferences on the causal variable, but not on the selected high-dimensional controls. The estimated structural parameter of usage of DT using the PDS-selected variables and the full set of selected controls also shows that enterprises that use DT enhance their labour productivity by 47 per cent compared with non-users (Table 1C).

Table 1A: OLS using CHS LASSO-orthogonalized variables

Dep var: Log Lab. Productivity	Coeff.	Robust std. error
Usage of DT	0.6488***	0.2083

Table 1B: OLS using CHS post-LASSO-orthogonalized variables

Dep var: Log Lab. Productivity	Coeff.	Robust std. error
Usage of DT	0.4671**	0.1998

Table 1C: OLS with PDS-selected variables and full regressor set

Dep var: Log Lab. Productivity	Coeff.	Robust std. error
Usage of DT	0.4727**	0.2019
No. of workers	-0.0162	0.0029
Age owner	-0.2844	0.0349
State (Adamawa)	0.0995	0.5485
State (Kaduna)	-0.9023	0.2806
State (Katsina)	-0.1671	0.3030
State (Rivers)	1.0272	0.1774
Sector of activity (non-tradeable services)	-0.3808	0.1498
Sector of activity (tradeable services)	0.3302	0.1054
Main source of light fuel (candles)	0.2104	0.1686
Operation of activity (home)	-0.4243	0.1047
Effort to get credit (yes)	0.9312	0.1890
Cons	10.6371	0.2790
Observations	705	
No. of high-dimensional controls	114	
Selected controls	11	

Note: standard errors and test statistics valid only for usage of DT variable. ***p<0.01, **p<0.05, *p<0.1

Source: authors' calculations based on described data.

To help deal with the potential for simultaneity between usage of DT and labour productivity, we introduce our set of 10 instruments and run the IV LASSO estimations using CHS and PDS methods. We begin the discussion by looking at the first stage estimations and the battery of weak identification tests. The first stage estimation is a LASSO regression of usage of DT on the selected instruments and controls (Table 2A). The LASSO-selected instruments are ownership of a DT device in years 2010 and 2012, and not owning a DT device in years 2010 and 2012. The first-stage relationship shows that these two variables, after accounting for the LASSO-selected control regressors, are strongly correlated with the endogenous variable and therefore appear to be strong instruments. The robust weak identification F statistics are very high for the IVs of the optimal LASSO and post-LASSO as well as the full IV set. We also reject the null hypothesis that the instruments are not valid or weakly identified for the LASSO and post-LASSO orthogonalized based on the super score weak identification robust tests.

In the CHS estimations, the selected variables are again used to construct orthogonalized versions of our dependent variable (labour productivity), endogenous variable (usage of DT), and control variables, and to create optimal instruments from the LASSO-selected IVs. The orthogonalized versions based on LASSO and post-LASSO estimated coefficients using the optimal IVs created for the endogenous regressor are presented in Tables 2B and 2C. Compared with the CHS estimates without using instruments, the OLS estimates using the CHS LASSO and post-LASSO orthogonalized versions seem to be biased downwards. The structural effect of usage of DT in both LASSO and post-LASSO regressions is to boost labour productivity significantly: by 76 and 77 per cent, respectively.

In the PDS estimation, we use the LASSO-selected controls and instruments in a post-regularization IV estimation. As indicated, we form a 2SLS estimator using the selected instruments for the usage of DT with the union of selected high-dimensional controls for both labour productivity and usage of DT as control variables. The results for the structural parameter α using IV with PDS-selected variables and full regressor set are consistent with the CHS estimates. The average labour productivity difference between enterprises that use DT and non-users is 65 per cent, which is higher than the estimates of the OLS version (see Table 2D).

Table 2A: First-stage estimation(s)

Dep var: Usage of DT	**Coeff.**	**Robust std. error**
No. of workers	0.0030	0.0045
Age of owner	-0.0003	0.0003
State (Adamawa)	-0.4110***	0.0707
State (Kaduna)	-0.1445***	0.0365
State (Katsina)	-0.0872**	0.0395
State (Rivers)	-0.0097	0.0231
Sector of activity (non-tradeable services)	0.0020	0.0195
Sector of activity (tradeable services)	-0.0008	0.0137
Main source of light fuel (candles)	0.0087	0.0223
Operation of activity (home)	-0.0308**	0.0136
Effort to get credit (yes)	-0.0041	0.0247
DT history 1	0.2029***	0.0166
DT history 2	-0.7871***	0.0408
_cons	0.8211	0.0290
Weak identification F stats, robust (full set)	359.91	

Structural equation:

Table 2B: IV using CHS LASSO-orthogonalized variables

Dep var: Log Lab. Productivity	Coeff.	Robust std. error
Usage of DT	0.7636***	0.2874

Table 2C: IV using CHS post-LASSO-orthogonalized variables

Dep var: Log Lab. Productivity	Coeff.	Robust std. error
Usage of DT	0.7759***	0.2863

Table 2D: IV with PDS-selected variables and full regressor set

Dep var: Log Lab. Productivity	Coeff	Robust std. error
Usage of DT	0.6507**	0.2829
No. of workers	-0.2848	0.0349
Age of owner	-0.0160	0.0029
State (Adamawa)	0.1869	0.5574
State (Kaduna)	-0.8790	0.2819
State (Katsina)	-0.1143	0.3088
State (Rivers)	1.0272	0.1775
Sector of activity (non-tradeable services)	-0.3828	0.1499
Sector of activity (tradeable services)	0.3332	0.1055
Main source of light fuel (candles)	0.2300	0.1701
Operation of activity (home)	-0.4194	0.1049
Effort to get credit (yes)	0.9267	0.1892
Cons	10.45034	0.3480
Observations	705	
No. of high-dimensional controls	114	
Selected controls	11	
Number of instruments	10	
Selected instruments	2	

Note: standard errors and test statistics valid only for usage of DT variable. ***p<0.01, **p<0.05, *p<0.1

Source: authors' calculations based on described data.

Again, we attempt to look at the structural effect of access to and usage of DT in relation to the size of the enterprise. Here, we look at the DT effect when the enterprise employs 1 person against the effect when employment is more than 1. We are only able to produce results without accounting for unconfoundedness in the unobservables. This is due to the fact that there are only a few enterprises in our sample that have more than 1 employee and therefore we are not able to create optimal IVs using our IV LASSO method. However, our structural estimates show that in the tier where enterprises have more than 1 employee, the average labour productivity difference between enterprises that use DT and non-users is much more pronounced than the tier where enterprises have only 1 employee (see Tables 3A and 3B). The structural effect of DT is almost double among enterprises that have more than 1 employee. This clearly shows the substantial effect of DT on labour productivity in enterprises that are in the upper size segment of the informal sector. The increased labour productivity and expansion in output for these upper-tier enterprises may have the potential to propel them to formalize in the long run.

Table 3A: Structural estimates if enterprise size equal to 1

OLS using CHS LASSO-orthogonalized variables

Dep var: Log Lab. Productivity	**Coeff.**	**Robust std. error**
Usage of DT	0. 6182***	0.2137

OLS using CHS post-LASSO-orthogonalized variables

Dep var: Log Lab. Productivity	**Coeff.**	**Robust std. error**
Usage of DT	0.5475***	0.2138

OLS with PDS-selected variables and full regressor set

Dep var: Log Lab. Productivity	**Coeff.**	**Robust std. error**
Usage of DT	0.5530***	0.2151
Age owner	-0.0185	0.0031
State (Adamawa)	0.2119	0.5609
State (Katsina)	-0.2547	0.3997
Sector of activity (tradeable services)	0.5835	0.0993
Main source of light fuel (candles)	0.1143	0.1783
Operation of activity (home)	-0.3877	0.1137
Cons	10.2372	0.2882
Observations	606	
No. of high-dimensional controls	114	
Selected controls	6	

Table 3B: Structural estimates if enterprise size is greater than 1

OLS using CHS LASSO-orthogonalized variables

Dep var: Log Lab. Productivity	**Coeff.**	**Robust std. error**
Usage of DT	1.9834***	0.6359

OLS using CHS post-LASSO-orthogonalized variables

Dep var: Log Lab. Productivity	**Coeff.**	**Robust std. error**
Usage of DT	1.6867**	0.8095

OLS with PDS-selected variables and full regressor set

Dep var: Log Lab. Productivity	**Coeff.**	**Robust std. error**
Usage of DT	1.7634**	0.7779
No. of workers	-0.2428	0.0467
State (Bauchi)	-0.7896	0.3495
State (Kaduna)	-1.0813	0.3243
State (Rivers)	1.1246	0.3514
State (Sokoto)	-0.9637	0.1412
Sector of activity (non-tradeable services)	-0.1947	0.1500

Freq. of black outs (several times in a year)	0.3753	0.4256
Source of capital (proceeds from family business)	3.2224	0.8435
Cons	8.4459	0.7991
Observations	99	
No. of high-dimensional controls	114	
Selected controls	8	

Note: standard errors and test statistics valid only for usage of DT variable. ***p<0.01, **p<0.05, *p<0.1

Source: authors' calculations based on described data.

This sturdy impact of DT on labour productivity in the informal sector may stem from the role of DT in helping enterprises in Nigeria to 'leapfrog' poor roads and middlemen and connect consumers to markets. For example, through the use of social media, DT enables informal enterprises to broaden their market reach to search, compete in, and sell in more markets (Aker 2008; Ongori and Migiro 2010). By helping overcome the imperfect and asymmetric information problem, DT enables informal enterprises to access and compare input prices, output prices, and potential buyers and sellers of their output. DT is also able to help enterprises in the informal sector mitigate barriers to accessing quality intermediate inputs that were previously unavailable to them and learn about best practices. The rise of mobile money and simple applications in Nigeria and many other SSA countries has provided the opportunity for enterprises to save and contribute to different financial schemes that allow them to access small loans and insurance. DT in this regard is able to assist informal enterprises to expand their operations and invest in building their human and production capabilities. Mobile money in particular facilitates efficient and seamless transactions between enterprises, suppliers, and customers. Informal enterprises are also able to obtain credits from multiple sources—including members of their social network both within a country and across international boundaries—to finance their operations and respond better to shocks in the domestic financial market (Aker and Mbiti 2010).

5 Conclusion and policy implications

In this paper, we have examined the structural impact of access to and usage of digital technology (DT) on the labour productivity of informal non-farm enterprises in Nigeria using panel LSMS-ISA survey data. To the best of our knowledge, this study represents the first attempt to empirically examine this impact in SSA. Given the many issues (e.g., endogeneity, omitted variable bias) that bedevil econometric estimations, we use IV LASSO techniques to carry out the estimations in this study. The findings from both our initial econometric estimates (without instruments) and the IV LASSO estimations for the structural parameter—DT access and usage—show that labour productivity is significantly higher for informal enterprises that use DTy than for non-users. Further analysis reveals that these benefits arise more strongly in enterprises in the upper size segment of the informal sector.

These findings echo the anecdotal evidence and the broader literature on the productivity-enhancing effect of access to and usage of DT at the micro informal level. The context of the study—Nigeria—is also intuitive as empirical studies of this kind on SSA are particularly lacking. The increased labour productivity and the expansion in the output of these firms due to access to DT suggests that DT may have the potential to propel these enterprises to formalize in the long run.

The findings of the paper suggest that a reasonable pathway in the short to medium term in SSA would be to refocus policies and reforms on enhancing the productivity of workers and enterprises in the informal sector using DT. Given this, there is a need to formulate well designed policies that target the removal of barriers to the access to and use of DTs, such as costly regulation and high tariffs. Such policies would enable many enterprises in the informal sector in Nigeria and other SSA countries to access and use DT. Additional policy considerations should target the provision of training in digital literacy while addressing supply-side barriers to access to connectivity infrastructure, e.g. through investment in the electricity and DT network infrastructure.

References

Acemoglu, D., and D. Autor (2011). 'Skills, Tasks and Technologies: Implications for Employment and Earnings'. In D. Card and O. Ashenfelter (eds), *Handbook of Labor Economics (Volume 4B)*, pp. 1043–171. Amsterdam and Boston: Elsevier, North-Holland. https://doi.org/10.1016/S0169-7218(11)02410-5

African Development Bank (2013). *Recognizing Africa's Informal Sector*. Abidjan: African Development Bank.

Ahrens, A., C.B. Hansen, and M.E. Schaffer (2018). 'PDSLASSO: Stata Module for Post-selection and Post-regularization OLS or IV Estimation and Inference'. Available at: http://ideas.repec.org/c/boc/bocode/s458459.html (accessed 20 June 2021).

Aker, J.C., and I.M. Mbiti (2010). 'Mobile Phones and Economic Development in Africa'. *Journal of Economic Perspectives*, 24(3): 207–32. https://doi.org/10.1257/jep.24.3.207

Aker, J.C. (2008). 'Does Digital Divide or Provide? The Impact of Cell Phones on Grain Markets in Niger'. Center for Global Development Working Paper 154. Washington, DC: Center for Global Development. https://doi.org/10.2139/ssrn.1093374

Akerman, A., I. Gaarder, and M. Mogstad (2015). 'The Skill Complementarity of Broadband Internet'. *Quarterly Journal of Economics*, 130(4): 1781–824. https://doi.org/10.3386/w20826

Andersen, T.B., J. Bentzen, C.J. Dalgaard, and P. Selaya (2012). 'Lightning, IT Diffusion, and Economic Growth across US States'. *Review of Economics and Statistics*, 94(4): 903–24. https://doi.org/10.1162/REST_a_00316

Avenyo, E.K., M. Konte, and P. Mohnen (2019). 'The Employment Impact of Product Innovations in Sub-Saharan Africa: Firm-level Evidence'. *Research Policy*, 48: 103806. https://doi.org/10.1016/j.respol.2019.103806

Bas, M., and V. Strauss-Kahn (2015). 'Input–Trade Liberalization, Export Prices, and Quality Upgrading'. *Journal of International Economics*, 95(2): 250–62. https://doi.org/10.1016/j.jinteco.2014.12.005

Bayen, M. (2018). 'Africa: a Look at 442 Active Tech Hubs of the Continent'. Mobile for Development (blog), 22 March. Available at: https://www.gsma.com/mobilefordevelopment/programme/ecosystem-accelerator/africa-a-look-at-the-442-active-tech-hubs-of-the-continent/ (accessed 20 June 2021).

Belloni, A., and V. Chernozhukov (2013). 'Least Squares after Model Selection in High-dimensional Sparse Models'. *Bernoulli*, 19(2): 521–47. https://doi.org/10.3150/11-BEJ410

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012). 'Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain'. *Econometrica*, 80(6): 2369–429. https://doi.org/10.3982/ECTA9626

Belloni, A., V. Chernozhukov, and C. Hansen (2014). 'Inference on Treatment Effects after Selection among High-Dimensional Controls'. *The Review of Economic Studies*, 81(2): 608–50. https://doi.org/10.1093/restud/rdt044

Belloni, A., V. Chernozhukov, I. Fernández-Val, and C. Hansen (2015). 'Program Evaluation with High-dimensional Data'. CEMMAP Working Paper CWP55/15. London: Centre for Microdata Methods and Practice.

Benhassine, N., D. McKenzie, V. Pouliquen, and M. Santini (2018). 'Does Inducing Informal Firms to Formalize Make Sense? Experimental Evidence from Benin'. *Journal of Public Economics*, 157: 1–14. https://doi.org/10.1016/j.jpubeco.2017.11.004

Bertschek, I., and T. Niebel (2016). 'Mobile and More Productive? Firm-level Evidence on the Productivity Effects of Mobile Internet Use'. *Telecommunication Policy*, 40(9): 888–98. https://doi.org/10.1016/j.telpol.2016.05.007

Bertschek, I., D. Cerquera, and G.J. Klein (2013). 'More Bits—More Bucks? Measuring the Impact of Broadband Internet on Firm Performance'. *Information Economics and Policy*, 25(3): 190–203. https://doi.org/10.1016/j.infoecopol.2012.11.002

Blimpo, M.P., and S. Owusu (2020). 'Human Capital'. In J. Choi, M.A. Dutz, and Z. Usman (eds), *The Future of Work: Harnessing the Potential of Digital Technology for All*, pp. 77–120. Washington, DC: The World Bank.

Bloom, N., M. Draca, T. Kretschmer, and R. Sadun (2010). 'The Economic Impact of ICT'. Final Report. London: London School of Economics Centre for Economic Performance.

Bloom, N., R. Sadun, and J. Van Reenen (2012). 'Americans Do It Better: US Multinationals and the Productivity Miracle'. *American Economic Review*, 102(1): 167–201. https://doi.org/10.1257/aer.102.1.167

Bruhn, M., and D. McKenzie (2014). 'Entry Regulation and the Formalization of Microenterprises in Developing Countries'. *The World Bank Research Observer*, 29(2): 186–201. https://doi.org/10.1093/wbro/lku002

Brynjolfsson, E., and L.M. Hitt (2003). 'Computing Productivity: Firm-Level Evidence'. *Review of Economics and Statistics*, 85(4): 793–808. https://doi.org/10.1162/003465303772815736

Brynjolfsson, E., and A. Saunders (2010). *Wired for Innovation: How Information Technology Is Reshaping the Economy*. Cambridge and London: MIT Press. https://doi.org/10.7551/mitpress/8484.001.0001

Campos, F., and M. Gassier (2017). 'Gender and Enterprise Development in Sub-Saharan Africa: a Review of Constraints and Effective Interventions'. Policy Research Working Paper 8239. Washington, DC: The World Bank. https://doi.org/10.1596/1813-9450-8239

Campos, F., M. Goldstein, and D. McKenzie (2015). 'Short-term Impacts of Formalization Assistance and a Bank Information Session on Business Registration and Access to Finance in Malawi'. Washington, DC: The World Bank. https://doi.org/10.1596/1813-9450-7183

Chernozhukov, V., D. Chetverikov, and K. Kato (2013). 'Gaussian Approximations and Multiplier Bootstrap for Maxima of Sums of High-dimensional Random Vectors'. *Annals of Statistics*, 41(6): 2786–819. https://doi.org/10.1214/13-AOS1161

Chernozhukov, V., C. Hansen, and M. Spindler (2015). 'Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments'. *American Economic Review*, 105(5): 486–90. https://doi.org/10.1257/aer.p20151022

Choi, J., M.A. Dutz, and Z. Usman (eds) (2020). *The Future of Work in Africa: Harnessing the Potential of Digital Technologies for All*. Washington, DC: World Bank Publications. https://doi.org/10.1596/978-1-4648-1445-7

Colombo, M.G., A. Croce, and L. Grilli (2013). 'ICT Services and Small Businesses' Productivity Gains: an Analysis of the Adoption of Broadband Internet Technology'. *Information Economics and Policy*, 25(3): 171–89. https://doi.org/10.1016/j.infoecopol.2012.11.001

Cusolito, A.P., and W.F. Maloney (2018). *Productivity Revisited: Shifting Paradigms in Analysis and Policy*. Washington, DC: World Bank Publications. https://doi.org/10.1596/978-1-4648-1334-4

Danquah, M., and B. Ouattara (2014). 'Productivity Growth, Human Capital and Distance to Frontier in Sub-Saharan Africa'. *Journal of Economic Development*, 39(4): 27–48. https://doi.org/10.35866/caujed.2014.39.4.002

Deen-Swarray, M., M. Moyo, and C. Stork (2013). 'ICT Access and Usage among Informal Businesses in Africa'. *info*, 15(5): 52–68. https://doi.org/10.1108/info-05-2013-0025

Diao, X., J. Kweka, and M. McMillan (2018). 'Small Firms, Structural Change and Labor Productivity Growth in Africa: Evidence from Tanzania'. *World Development*, 105: 400–15. https://doi.org/10.1016/j.worlddev.2017.12.016

Draca, M., R. Sadun, and J. Van Reenen (2009). 'Productivity and ICTs: a Review of the Evidence'. In C. Avgerou, R. Mansell, F. Quah, and R. Siverstone (eds), *The Oxford Handbook of Information and Communication Technologies*, pp. 100–47. Oxford and New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199548798.003.0005

Economic Intelligence Unit (2016). 'Building a digital Nigeria'. Available at: https://eiuperspectives.economist.com/technology-innovation/building-digital-nigeria (accessed 12 May 2021).

Fu, X., P. Mohnen, and G. Zanello (2015). 'Innovation, Informality, and Firm's Growth in Low-income Countries'. TMCD Working Paper TMD-WP-72. Oxford: Technology and Management Centre for Development.

Goldfarb, A., and C. Tucker (2019). 'Digital Economics'. *Journal of Economic Literature*, 57(1): 3–43. https://doi.org/10.1257/jel.20171452

Grimes, A., C. Ren, and P. Stevens (2012). 'The Need for Speed: Impacts of Internet Connectivity on Firm Productivity'. *Journal of Productivity Analysis*, 37(2): 187–201. https://doi.org/10.1007/s11123-011-0237-z

Grimm, M., P. Knorringa, and J. Lay (2012). *Constrained Gazelles: High Potentials in West Africa's Informal Economy*. Washington, DC: The World Bank. https://doi.org/10.1596/26877

Haller, S.A., and S. Lyons (2015). 'Broadband Adoption and Firm Productivity: Evidence from Irish Manufacturing Firms. *Telecommunications Policy*, 39(1): 1–13. https://doi.org/10.1016/j.telpol.2014.10.003

Hjort, J., and Poulsen (2019). 'The Arrival of Fast Internet and Employment Creation in Africa'. *America Economic Review*, 109(3): 1032–79. https://doi.org/10.1257/aer.20161385

Hodler, R., and P.A. Raschky (2017). 'Ethnic Politics and the Diffusion of Mobile Technology in Africa'. *Economics Letters*, 159: 78–81. https://doi.org/10.1016/j.econlet.2017.07.011

ILO (2018). *Women and Men in the Informal Economy: A Statistical Picture (Third edition)*. Geneva: International Labour Office.

Jensen, R. (2007). 'The Digital Provide: Information (Technology), Market Performance and Welfare in the South Indian Fisheries Sector.' *Quarterly Journal of Economics*, 122(3): 879−924. https://doi.org/10.1162/qjec.122.3.879

Jorgenson, D.W., M.S. Ho, and K.J. Stiroh (2008). 'A Retrospective Look at the U.S. Productivity Growth Resurgence'. *Journal of Economic Perspectives*, 22(1): 3–24. https://doi.org/10.1257/jep.22.1.3

Knoche, H., and J. Huang (2012). 'Text Is Not the Enemy: How Illiterate People Use Their Mobile Phones'. NUI workshop at CHI'12, Austin, Texas, USA.

La Porta, R., and A. Shleifer (2011). 'The Unofficial Economy in Africa'. NBER Working Paper 16821. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w16821

La Porta, R., and A. Shleifer (2014). 'Informality and Development'. NBER Working Paper 20205. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w20205

Leeb, H., and B.M. Pötscher (2008). 'Can One Estimate the Unconditional Distribution of Post-model-selection Estimators?'. *Econometric Theory*, 24(2): 338–76. https://doi.org/10.1017/S0266466608080158

Lince, S. (2011). 'The Informal Sector in Jinja, Uganda: Implications of Formalization and Regulation'. *African Studies Review*, 54(2): 73–93. https://doi.org/10.1353/arw.2011.0029

Maliranta, M., and P. Rouvinen (2006). 'Informational Mobility and Productivity: Finnish Evidence'. *Economics of Innovation and New Technology*, 15(6): 605–16. https://doi.org/10.1080/10438590500418935

Manacorda, M., and A. Tesei (2020). 'Liberation Technology: Mobile Phones and Political Mobilization in Africa'. *Econometrica*, 88(2): 533–67. https://doi.org/10.3982/ECTA14392

Mbiti, I., N. David, and D.N. Weil (2011). 'Mobile Money: the Impact of M-Pesa in Kenya'. NBER Working Paper 17129. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w17129

Mohammad, A., and A. Islam (2015). 'Are Large Informal Firms More Productive than the Small Informal Firms? Evidence from Firm-level Surveys in Africa'. *World Development*, 74: 374–85. https://doi.org/10.1016/j.worlddev.2015.05.008

Nagler, P. (2015). *Occupational Choice in the Developing World*. Maastricht: Boekenplan.

Ndubuisi, G., and S. Owusu (2021). 'How Important Is GVC Participation to Export Upgrading?'. *The World Economy*, January (online). https://doi.org/10.1111/twec.13102

Nguimkeu, P., and C. Okou (2020). *Increasing Informal Sector Productivity*. Washington, DC: World Bank Publications.

Ohnsorge, F., and S. Yu (eds) (2021). *The Long Shadow of Informality: Challenges and Policies*. Washington, DC: World Bank Publications.

Ongori, H., and S.O. Migiro (2010). 'Information and Communication Technologies Adoption in SMEs: Literature Review'. *Journal of Chinese Entrepreneurship*, 2(1): 93–104. https://doi.org/10.1108/17561391011019041

Pahl, S., and M. Timmer (2020). 'Do Global Value Chains Enhance Economic Upgrading? A Long View'. *Journal of Development Studies*, 56(9): 1–23. https://doi.org/10.1080/00220388.2019.1702159

Piet, B., S. Dasgupta, S.T. Timothy, and D. Wheeler (2009). 'Determinants of a Digital Divide in Sub-Saharan Africa: a Spatial Econometric Analysis of Cell Phone Coverage'. *World Development*, 37(9): 1494–505. https://doi.org/10.1016/j.worlddev.2009.01.011

Rohman, I.K., and E. Bohlin (2011). 'Does Broadband Speed Really Matter for Driving Economic Growth? Investigating OECD Countries'. Study for LM Ericsson, Arthur D. Little, and Chalmers University of Technology, Gothenburg, Sweden. Available at: https://www.econstor.eu/obitstream/10419/60385/1/72027561X.pdf (accessed 20 June 2021).

SMEDAN (2019). 'MSME National Survey Report'. Abuja: Small and Medium Scale Enterprise Development Agency of Nigeria.

Solow, R.M. (1987). 'We'd Better Watch Out'. Book Review. *New York Times*, 12 July.

Stork, C., and S. Esselaar (2019). 'Regulatory and Tax Treatment of OTTs in Africa'. Available at: https://www.researchgate.net/profile/Christoph-Stork-4/publication/337648378_Regulatory_and_tax_treatment_of_OTTs_in_Africa/links/5de1effda6fdcc2837f6dadc/Regulatory-and-tax-treatment-of-OTTs-in-Africa.pdf (accessed 21 June 2021).

Stork, C., H.S. Nwana, S. Esselaar, and M. Koyabe (2020). *Over the Top (OTT) Applications and the Internet Value Chain*. London: Commonwealth Telecommunications Organisation.

Tibshirani, R. (1996). 'Regression Shrinkage and Selection via the Lasso'. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1): 267–88. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Ulyssea, G. (2010). 'Regulation of Entry, Labor Market Institutions and the Informal Sector'. *Journal of Development Economics*, 91(1): 87–99. https://doi.org/10.1016/j.jdeveco.2009.07.001

World Bank (2019). *World Development Report 2019: The Changing Nature of Work*. Washington, DC: The World Bank.

Appendix

Table A1: Description of all regression variables and summary statistics

Variable	Description	Mean	SD	Range
Labour productivity(log)	Continuous: measures sales per employee	9.61	1.35	3.68–17.32
Labour productivity(log)	Continuous: measures sales per employee among DT users	9.71	1.32	3.68–17.32
Labour productivity(log)	Continuous: measures sales per employee among non-DT users	8.97	1.36	5.01–13.81
DT access and usage	Binary: measures the enterprise's access to DT; it assumes a value of 1 if the enterprise has access to a mobile phone or computer or uses the internet and 0 if otherwise	0.86	0.33	0–1
Sales(log)	Continuous: captures the enterprise's total sales for each month (in naira)	9.73	1.33	3.68–18.42
Firm size	Continuous: captures the number of persons engaged by the workforce	1.25	1.19	1–27
Firm age	Continuous: captures the total number of months the enterprise has been in operation	6.17	2.92	0–52
Age	Continuous: captures the age of individual (years)	50.84	14.18	20–102
Sex	Binary: measures the gender of the individual; it assumes a value of 1 if 'male' and 0 if otherwise	0.87	0.33	0–1
Married	Binary: measures the marital status of the individual; it assumes a value of 1 if 'married' and zero if otherwise	0.85	0.38	0–1
Credit access	Binary: measures the individual's access to finance; it assumes a value of 1 if individual has access to credit and 0 if otherwise	0.41	0.49	0–1
Effort to get credit	Binary: measures the individual's effort to access credit; it assumes a value of 1 if individual makes an effort to access credit and 0 if otherwise	0.33	0.18	0–1
Source of capital	Categorical: captures the main source of the enterprise's start capital; it assumes a value of 1 if 'Household savings', 2 if 'NGO support', 3 if 'Loan from bank (commercial, microfinance)', 4 if 'Money lender', 5 if 'Esusu/adashi', 6 if 'Other loans', 7 if 'District/town association support', 8 if 'Cooperative/trade association', 9 if 'Remittances from abroad', 10 if 'Proceeds from family farm', 11 if 'Church/mosque assistance', 12 if 'Proceeds from family enterprise', 13 if 'Relatives/friends', 14 if 'Other'	4.09	4.77	1.14
Main source of lighting fuel	Categorical: captures the main source of lighting fuel of the household; it assumes a value of 1 if 'Firewood, 2 if 'Grass', 3 if 'Kerosine', 4 if 'Electricity', 5 if 'Gas', 6 if 'Battery/dry cell torch', 7 if 'Candles', and 8 if 'Other'	5.06	1.78	1–8
Main source of cooking fuel	Categorical: captures the main source of cooking fuel of the household; it assumes a value of 1 if 'Firewood, 2 if 'Grass', 3 if 'Coal', 4 if 'Kerosine', 5 if 'Electricity', 6 if 'Gas', 7 if 'Battery/dry cell torch', 8 if 'Candles', and 9 if 'Other'	2.34	1.76	1–9
Electricity access	Binary: captures whether the individual has access to electricity in the dwelling; it assumes a value of 1 if 'yes' and 0 if otherwise	0.54	0.47	0–1

Variable	Description	Mean	SD	Range
Frequency of blackouts	Categorical: measures how frequently the individual experiences local blackouts; it assumes a value of 1 if ‘Never’, 2 if ‘Every day’, 3 if ‘Several times a week’, 4 if ‘Several times a month’, and 5 if ‘Several times a year’	2.56	0.84	1–5
Energy	Categorical: measures the household’s source of energy for lighting during blackouts; it assumes a value of 1 if ‘Firewood’, 2 if ‘Kerosene’, 3 if ‘Rechargeable lamp’, 4 if ‘Generator’, and 5 if ‘Other’	2.76	9.22	1–5
Market	Categorical: measures whom the enterprise sells products or services to; it assumes a value of 1 if ‘Final consumers’, 2 if ‘Traders’, 3 if ‘Other small businesses’, 4 if ‘Large established businesses’, 5 if ‘Institutions (schools, hospitals, govt)’, 6 if ‘Manufacturers’, 8 if ‘Other’	1.57	1.01	1–8
Health	Binary: measures whether the individual consulted a health practitioner in the past two weeks; it assumes a value of 1 if ‘yes’ and 0 if ‘no’	0.12	0.32	0–1
Operation location	Categorical: captures where the enterprise operates; it assumes a value of 1 if ‘Home (inside residence)’ 2 if ‘Home (outside residence)’ 3 if ‘Industrial site’, 4 if ‘Traditional market’, 5 if ‘Commercial area shop’, 6 if ‘Roadside’; 7 if ‘Mobile/no fixed location’; 8 if ‘Other fixed place’; 9 if ‘Other (specify)’	3.68	2.41	1–9
Urban	Binary: measures the locality of the enterprise; it assumes a value of 1 if ‘urban’ and 0 if otherwise	0.35	0.47	0–1
Distance to get drinking water	Continuous: measures the time (in minutes) taken to get to drinking water and back to dwelling	14.26	84.72	0-4000
Power	Continuous: measures the number of hours in a typical month an average household receives power from main public system	37.03	40.40	0–450
States	Categorical: captures the 34 States in Nigeria included in our sample	18.33	9.12	1–34
Sector	Categorical: captures the sector to which data relate; it assumes a value of:			
	1 if ‘Construction’	0.04	0.21	0–1
	2 if ‘Manufacturing’	0.14	0.35	0–1
	3 if ‘Mining and Quarrying’	0.01	0.05	0–1
	4 if ‘Non-Tradable Services’	0.11	0.47	0–1
	5 if ‘Tradable Services’,	0.01	0.07	0–1
	6 if ‘Others’			
DT history 1	Binary: measures the number of years the individual had access to DT; it assumes a value of 1 if ‘individual had access to DT in 2010 and 2012’ and 0 if otherwise	0.78	0.41	0–1
DT history 2	Binary: measures the number of years the individual had access to DT; it assumes a value of 1 if ‘individual had no access to DT in 2010 and 2012’ and 0 if otherwise	0.04	0.20	0–1
DT history 3	Binary: the measures number of years the individual had access to DT; it assumes a value of 1 if ‘individual had access to DT either in 2010 or 2012’ and 0 if otherwise	0.17	0.37	0–1
DT history 4	Binary: measures the number of years the individual had access to DT; it assumes a value of 1 if ‘individual had access to DT in 2010 but not 2012’ and 0 if otherwise	0.08	0.28	0–1
Lightning strike intensity	Continuous: measures the intensity of lightning strikes in a local government area	1.25	0.63	0.02–4.24

Variable	Description	Mean	SD	Range
Mobile phone coverage	Continuous: measures mobile phone coverage in a local government area weighted by population	1.07	0.76	0–2
DT intensity	Continuous: measures total number of mobile phones, computers, and internet subscriptions in a local government area	22.96	19.63	0–109
Distance to nearest major road	Continuous: measures total distance (km) to the nearest major road	9.08	13.44	0–102.1
Distance to nearest population centre	Continuous: measures total distance (km) to the nearest major population centre (>20,000 inhabitants)	18.68	17.88	0.06–109
Distance to nearest market	Continuous: measures total distance (km) to the nearest market	69.63	43.82	0.37–214
Distance to nearest border crossing	Continuous: measures total distance (km) to the nearest border crossing	275.20	166.98	1.6–641
Distance to capital of state of residence	Continuous: measures total distance (km) to the capital of the state of residence	62.54	51.97	0.18–259

Source: authors' calculations based on described data.