UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2023/99

Mobile technologies and firm formalization

Evidence from Uganda

Marina Dodlova,[1] Krisztina Kis-Katos,[2] Anna Kochanova,[3] and Olivia Wirth[4]

August 2023

Abstract: We investigate how the arrival and expansion of mobile network access in Uganda influences firm tax behaviour. Access to mobile technologies could broaden government revenues from corporate income tax through the extensive margin: by reducing the costs of formalization, it could increase the number of firms filing corporate income tax. If these newly formalizing firms are also economically successful, they will contribute to the expansion of the tax base. Moreover, mobile technologies could also enhance firm performance directly, resulting in further increases in the tax base. Among the possible channels of the relationship between the use of mobile technology and tax liabilities, we examine changes in firm performance as higher firm productivity will broaden the tax base. We assess the effects of mobile technologies on firm formalization and tax outcomes using administrative tax records provided by the Uganda Revenue Authority from 2013 to 2020. We link subcounty-sector-level aggregate outcomes to the roll-out of the 3G mobile network and contrast two identification strategies to assess causal effects. We extend two-way fixed effects models by a shift–share instrumental variable strategy that predicts the local roll-out of mobile technologies based on the costs of network maintenance, proxied by local exposure to lightning. We complement those results with staggered difference-in-differences estimates. We find that the roll-out of mobile technologies increases the number of firms reporting to the tax authorities, as well as overall tax revenues. While increased formalization results in more overall sales recorded in the formal economy, firms also report substantially larger costs and deductions, which lead to higher losses that are carried forward to the next fiscal year, reducing next year's tax base.

Key words: mobile technologies, tax behaviour, administrative records, staggered difference-in-differences, shift–share instruments, firm performance, Uganda

JEL classification: H25, H71, O17

Acknowledgements: This study uses the Uganda Revenue Authority Firm Panel data developed in collaboration with the Uganda Revenue Authority (URA) and UNU-WIDER. The authors would like to thank URA and UNU-WIDER for providing the opportunity to access the data, as well as Brian Arinaitwe, Gerald Agaba, Nicholas Musoke, and Tina Barugahara from the URA, and Maria Jouste from UNU-WIDER for their incredible support throughout the project. The authors also thank Michael Grimm, Maria Jouste, Rose Camille Vincent, Pietro Bomprezzi, Florian Anselm Münch, and the audience at the EPCS 2023, GDEC 2023, and ICDE 2023 for their helpful comments on earlier versions of the paper. The project was funded by UNU-WIDER as well as the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—project number 443732464.

This study has received ethical approval by the Joint Ethical Review Board of the United Nations University (Ref No: 202104/01) on 11 May 2021.

---

[1] University of Passau, Passau, Germany, and CESifo, Munich, Germany; [2] University of Göttingen, Göttingen, Germany; [3] Cardiff Business School, Cardiff, UK; [4] University of Passau, Passau, Germany; corresponding author: Marina.Dodlova@Uni-Passau.De

This study has been prepared within the UNU-WIDER project Building up efficient and fair tax systems – lessons based on administrative tax data, which is part of the Domestic Revenue Mobilization programme. The programme is financed through specific contributions by the Norwegian Agency for Development Cooperation (Norad).

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-407-6

https://doi.org/10.35188/UNU-WIDER/2023/407-6

Typescript prepared by Gary Smith.

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 and Sweden, as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Access to digital technologies, such as mobile phones and the internet, has been quickly and successfully expanding in Africa, spurring a much-needed development process (Aker and Mbiti 2010). Mobile technologies have been shown to increase household welfare and employment (Bahia et al. 2023; Hjort and Poulsen 2019), improve the targeting of anti-poverty programmes (Aiken et al. 2022) and outcomes in the domains of health (Busso et al. 2022; Mensah et al. 2022) and education (Aker and Ksoll 2019; Aker et al. 2012; Porter et al. 2016), and foster economic development in general (Mensah 2021). From the firms' perspective, studies also document substantial scope for improvements in firm productivity with mobile technology access (Bertschek and Niebel 2016; D'Andrea and Limodio 2019; Islam et al. 2018; Konte and Tetteh 2022; Zuo 2021). However, there is yet no empirical evidence on how the access to mobile technologies affects firm tax behaviour. This is of high policy relevance as firm formalization and improvements in the overall tax base are expected to play a crucial role in developing state capacity in the Global South (e.g. OECD 2007 and Sustainable Development Goal (SDG) 17:1).

This paper aims at closing this gap by investigating how the arrival and expansion of third-generation (3G) mobile technology affected firm formalization and payments of corporate income tax (CIT) in Uganda.

The availability of mobile technologies may affect firm formalization and tax behaviour through various channels. Mobile internet can help to reduce communication and transaction costs, simplify the bureaucratic procedures of filing taxes, and increase transparency. The expansion of mobile technologies may also provide tax authorities with better instruments to identify unofficial firms (Santoro et al. 2023). Recording economic transactions digitally increases the probability of their detection by tax authorities. This could stimulate firms to register and report their activities as well as to pay taxes, raising tax compliance in the local economy. Therefore, we expect higher mobile internet coverage to lead to a higher rate of firm formalization. Mobile technologies could also give firms cheap and quick access to all information regarding their tax liabilities and might be used to simplify the payment processes (e.g. through the use of mobile money). The lowering of information and administrative frictions could result in higher tax compliance, increasing not only the number of reporting firms, but also the overall volume of tax liabilities (Benzarti 2020; Besley et al. 2021; Slemrod 2019). Moreover, the expansion of mobile technologies could additionally improve the local tax base by enhancing firm performance. Better access to borrowing, improved patterns of investment, and an overall increasing liquidity could broaden the local tax base by enhancing firm productivity. Such expected increases in the local tax base and in total taxes paid might not materialize, however, if the mass of newly formalized firms is not profitable or if firms deliberately adjust their reporting behaviour. Strategic changes in reporting might lead to firms recording higher costs, which would mean that tax liabilities do not rise despite increasing sales (Almunia et al. 2022).

Uganda offers the ideal testing ground to study the effects of mobile technologies on firm formalization and tax behaviour. The country started to transform its telecommunication sector with the Communications Act of 2013,[1] which triggered an expansion of 3G mobile technologies that were rolled out across the country, with sharp increases in coverage in 2014 and 2020.[2] This broadband internet technology was the first to allow free web browsing and streaming and facilitated a variety of new online services with the potential to lead to an expansion of the tax base.

[1] The Act created a single regulatory body for telecommunications, broadcasting, radio, data and postal communications, and infrastructure; it expanded the coverage and variety of communications services available in the country, reduced the role of government, and encouraged competition in the communications sector.

[2] Fourth-generation (4G) mobile coverage arrived in Uganda starting in 2020, allowing for substantial increases in the speed and reliability of mobile services. In this paper we focus on 3G coverage only as our sample ends in 2020.

In this paper, we use administrative data from firm tax records from the Uganda Revenue Authority (URA 2022) and study the aggregate yearly dynamics of firm formalization and tax outcomes at the level of subcounty-sector cells. We define firm formalization as the increase in the number of firms filing CIT reports.[3] Our other main outcome variables include the total tax base, total tax liabilities, and a number of aggregated firm performance measures from profit and loss and balance sheet account statements, which we capture at the level of 22 economic sectors and 829 administrative regions (subcounties). We link this data to the spatial roll-out of 3G mobile coverage weighted by population density and aggregated at the subcounty level.

The speed of mobile coverage expansion within a country is not exogenous, as 3G coverage most likely first expanded in places where mobile technologies offer the largest net benefits, and hence where the demand for digital services is higher and the costs of their expansion are lower. Economic activities, thus, might not only respond to mobile network access, but also trigger its expansion. To overcome such endogeneity concerns, our regression analysis controls for nighttime light intensity as a proxy of local economic activity and all results are conditional on a full set of district–year and sector–year fixed effects. In addition, we instrument the local expansion of mobile coverage by a shift–share-like instrument (following Bartik 1991). The 'shift' part of this instrument captures the national expansion of mobile coverage over time, which is interacted with the initial intensity of lightning strikes in each subcounty ('share') that proxies the costs of maintaining and expanding a mobile network (Andersen et al. 2011, 2012). The product between the 'shift' and 'share' variables results in a continuous shift–share instrumental treatment variable (SSIV) (Borusyak et al. 2022), which helps to isolate the exogenous component in the local expansion of 3G network coverage over time. We complement this strategy with the results from an event study analysis using a staggered difference-in-differences (DID) design and applying an estimator developed by De Chaisemartin and d'Haultfoeuille (2020).

Our results provide evidence that mobile internet expansion is associated with firm formalization, increasing the number of firms reporting to the tax authorities. We document increases among firms reporting zero activities in all years (nil-reporting firms) but also among firms reporting actual non-zero information. More importantly, total tax liabilities as well as the tax base increase with the expansion of mobile internet, leading to improvements in the overall state capacity. However, these results offer only indirect evidence for higher tax compliance as they could theoretically reflect both firm formalization and market entry of new firms. Moreover, the documented increase in deferred tax liabilities might indicate strategic delaying of tax payments and therefore lower tax compliance. We also find that the roll-out of mobile technologies is associated with larger overall sales recorded in the formal economy as well as with substantially larger costs, losses and deductions reported by CIT filing firms.

Our paper makes several contributions to the literature. To the best of our knowledge, this is the first study to provide causal empirical evidence on the role of mobile coverage expansion on tax revenue collection using detailed CIT data in a developing country context. Previous studies either present cross-country evaluations (Jacolin et al. 2021) or document the positive impact on taxation of electronic filing and other digital technologies (e.g., Bellon et al. 2022; Kochanova et al. 2020). We enrich this strand of the literature by focusing exclusively on mobile internet access and thus generalizing the available findings on specific technological innovations to a wider range of digital technologies that are linked to internet connectivity (Hjort and Tian 2023). Further, our paper uncovers an important determinant of firm formalization in developing countries and its link with tax compliance (e.g., Bruhn and McKenzie

[3] Formalization is commonly used to refer to a variety of different processes, such as titling of property, accessing public services, registering for taxpayer identification numbers or business licences, and paying taxes (Floridi et al. 2020; Gallien and van den Boogaard 2021). Our definition of formalization constitutes a lower bound of actual formalization, as it is likely that more firms are registered with the URA (having attained a Taxpayer Identification Number, TIN), but are not submitting tax reports. In this study, we focus on the number of firms filing their CIT as this reflects the state's capacity to collect tax revenues.

2014; Devas and Kelly 2001; Floridi et al. 2020). We also add to the nascent literature on the impact of mobile technologies on firm performance (e.g., Bloom and Van Reenen 2007; Bloom et al. 2013; Cariolle and Le Goff 2023; Casaburi et al. 2014; Commander et al. 2011; De Mel et al. 2009; Paunov and Rollo 2016). Therefore, the results of our study are highly relevant for policy-makers aiming to improve both firm performance and tax collection in developing countries.

The remainder of the paper is structured as follows. Section 2 outlines the related literature and briefly describes the institutional environment of taxation and mobile technologies in Uganda. Section 3 introduces the data, while Section 4 outlines the empirical methodology. Section 5 presents our empirical results. Section 6 concludes.

2 Background

2.1 Firm formalization and taxation

Tax revenues from corporate and personal income play an important role as a source of government revenue in African countries (in accordance with SDG 17:1). In Uganda, like in many low- and middle-income countries (LMICs), a large share of the firms, however, engages in the informal economy and is not registered with the official tax authorities. These firms do not pay income tax and hence do not directly contribute to the national tax revenues. Against this backdrop, various policies target firm formalization with the goal to improve the state's tax capacity, to increase the overall fairness of the tax system, and ultimately to achieve more inclusive growth (Gallien and van den Boogaard 2021). Beyond its effects on tax revenues, formalization is expected to yield a wide range of further positive effects. At the societal scale, the resulting increase in state capacity is expected to enhance access to public services and to promote civic and political participation. Formalization might bring further direct advantages to firms by reducing extortion, facilitating adoption of professional practices and technology, increasing the utilization of formal financial services, and ultimately improving firms' investment and business performance.

Policies that effectively reduce the costs associated with filing taxes can be expected to have a positive impact on firm formalization (Devas and Kelly 2001). But whether tax compliance actually increases will depend on whether firms receive tangible benefits from formalization (Dom et al. 2022). General equilibrium models of firm dynamics that incorporate various types of financial frictions illustrate that despite the costs associated with taxes and regulations, operating within the formal sector enables firms to access more efficient credit markets and to obtain capital at lower interest rates (D'Erasmo and Boedo 2012; Lopez-Martin 2019). Empirical evidence emphasizes the importance of private incentives to formalize, including improved firm performance upon formalization (Boly 2018; Demenet et al. 2016; Rand and Torm 2012). By contrast, the cross-country study of Bruhn and McKenzie (2014) and the meta-analysis by Floridi et al. (2020) show that merely simplifying the formalization process is insufficient if the perceived benefits are minimal. Experimental findings document an extremely low willingness to formalize despite reductions in compliance costs in Brazil, indicating that formalization might offer limited private advantages in some contexts (De Andrade et al. 2014).

The literature on the link between firm tax formalization and the increase of total tax revenues collected by the state is mixed. While various studies find increases in tax revenues with firm formalization (Araujo-Bonjean and Chambas 2004; Brockmeyer et al. 2019; De Giorgi et al. 2018; De Mel et al. 2013), a growing body of literature shows that an increase in registered taxpayers does not automatically lead to a proportional increase in tax revenues (Lediga et al. 2020; Mascagni and Mengistu 2016; Mayega et al. 2019; Moore 2023). Rather, formally registered firms often still engage in tax evasion by underreporting their revenues, indicating only partial compliance with formal tax rules (Almunia et al. 2022;

Ulyssea 2020). Further, studies point to the fact that a large proportion of registered taxpayers do not file their returns or nil-file and thus do not contribute to tax revenues after all (Mascagni et al. 2022; Moore 2020; Santoro 2021).

2.2 The role of digital technologies in firm formalization and taxation

With the emergence of digital technologies, governments and tax administrations have started to adopt new forms of electronic tax filing systems, and enabled electronic tax payment via digital financial services. Such advances in information technology have the potential to decrease the costs of tax registration, filing, and payment, while at the same time easing tax monitoring. By reducing the need for in-person interactions, they can save costs directly, both for the firms and the state. Digital transactions can also lead to enhanced transparency and improve the government's access to information. Moreover, digital transactions can limit the opportunities for extortion, thereby reducing corruption. Hence, the widespread adoption of digital technologies offers the potential to substantially improve tax collection in African countries (Okunogbe and Santoro 2023).

The presence of mobile technologies can increase total tax liabilities by triggering firm formalization and enhancing overall firm productivity. For example, drawing on a sample of 101 developing countries, Jacolin et al. (2021) show that the adoption of mobile financial services decreases the relative size of the informal sector. They conclude that economic formalization, and with that the mobilization of domestic resources, can be enhanced through the expansion of mobile money, mobile credit, and savings. Mobile internet access can also increase the national tax base by reducing the costs of tax filing and collection for already formalized firms, for instance through electronic filing (e-filing) and payment of taxes.

A growing body of empirical literature investigates the economic benefits of e-filing interventions around the world. New e-filing systems implemented in LMICs significantly increased the number of small business taxpayers and presumptive tax revenues, reported firm sales, purchases, and VAT liabilities (Bellon et al. 2022; Jouste et al. 2021). Furthermore, e-filing systems reduce tax compliance costs, such as the time spent on tax-related activities, and boost state revenues by increasing the transparency of business transactions (in the Republic of Korea: Lee 2016; in Tajikistan: Okunogbe and Pouliquen 2022). In a cross-country sample of firms from 198 countries, Kochanova et al. (2020) show that e-filing decreases the time needed to prepare and to pay taxes, the probability and frequency of firms being visited by tax officials, and the perception of tax administrations as an obstacle to firms' operation and growth.

Not all improvements in taxation outcomes require mobile technologies. For instance, offline transaction recording systems play a crucial role in the case of value added tax (VAT). Adoption of a new digitized invoice encryption technology made it more difficult for Chinese manufacturing firms to falsify detectable claims, thereby reducing tax evasion and increasing VAT payments (Fan et al. 2018). In a similar vein, the adoption of electronic sales register machines (ESRMs) increased the volume of collected VAT in Ethiopia (Ali et al. 2021; Mascagni et al. 2021). However, the empirical literature also raises the issue of an eroding tax base if there is lower entry of new firms (Ali et al. 2021) or if existing firms become less productive (Fan et al. 2018) or start over-reporting their costs (Mascagni et al. 2021). Inefficient organization of national ICT systems might pose additional challenges to tax collection (Ligomeka 2020; Mayega et al. 2019).

2.3 Mobile internet and taxation in Uganda

In 2013, Uganda's President Yoweri Museveni launched a 30-year development master plan and set off an ambitious agenda of digital transformation. Its goal was to spur income growth by transforming a predominantly rural country into a competitive and market-driven economy (Okeleke 2019). A cornerstone of this development strategy consisted of providing access to affordable, reliable, and fast digital services. This led to a relatively quick roll-out of internet access across the country. Theoretically, large-scale coverage could have been achieved either by expanding the fixed-line networks, the mobile access networks, or both. As fixed-line networks were barely present in Uganda before and their expansion would have come at relatively higher cost, mobile networks were chosen as the more viable and more popular solution to provide internet access (Okeleke 2019). According to the Uganda Communications Commission, fixed-line internet access remains low, and expanded between 2014 and 2018 from 40,000 to 173,600 subscriptions, whereas mobile internet use increased from 5.694 to 9.855 million subscriptions. Thus, mobile technologies had a much larger scope to influence the local economy than fixed-line internet access.

In our study, we link the expansion of mobile networks to firms' CIT returns, relying on administrative records from the URA. The CIT is paid by non-individuals, as well as self-employed individuals who choose not to file personal income tax returns. In Uganda, a CIT rate of 30 per cent is levied on the taxable income of firms and individuals, which is based on profits. Uganda's CIT rate lies slightly above the average in the region of sub-Saharan Africa (28.7 per cent) (OECD et al. 2021). Income from mining is charged at a 25 per cent tax rate, and repatriated branch profits a 15 per cent tax rate. Firms either pay CIT, or—in the case of smaller firms with annual turnover of less than 150 million UGX (Ugandan shillings)—a presumptive tax based on their turnover. While the CIT requires a firm to keep detailed accounts, it allows them to request tax deductions and exemptions in return. A non-individual taxpayer must file their final income tax returns electronically within six months of the end of the firm's accounting period.

Our analysis is not confounded by a parallel introduction of electronic tax filing and payment systems. Electronic systems were introduced in Uganda in 2010, and were largely adopted by 2011. Thus, starting in 2013, our analysis focuses on the expansion of mobile coverage in a time period in which electronic tax and payment systems were already operational in Uganda and could facilitate the filing and payment of CIT in all places with internet access.

3 Data

3.1 Mobile coverage

We extract high-resolution data on the spatio-temporal variation in mobile internet coverage across Uganda from the Global System for Mobile Communications Association (GSMA). The information on signal availability is provided by mobile operators with the purpose of creating worldwide roaming maps. The GSMA is compiled by Collins Bartholomew at a 260×260 m spatial resolution, and released every year in January.[4] The data allows us to measure yearly variation in mobile coverage at a high spatial resolution and has been widely used in recent years to assess the impact of mobile technologies on various socio-economic and political outcomes (Guriev et al. 2021; Manacorda and Tesei 2020; Mensah et al. 2022). As the data distinguishes between two types of signal strength, weak and strong,[5] we combine them in one population-weighted 3G coverage index. For that, we first assign the value of 0.5 to each cell covered by a weak signal and the value of 1 to each cell covered by a strong signal. We then calculate the population-density weighted average of mobile signal coverage across all grid cells within a subcounty. The population-density maps are extracted from the CIESIN 2018 (GPW), v4.11.[6]

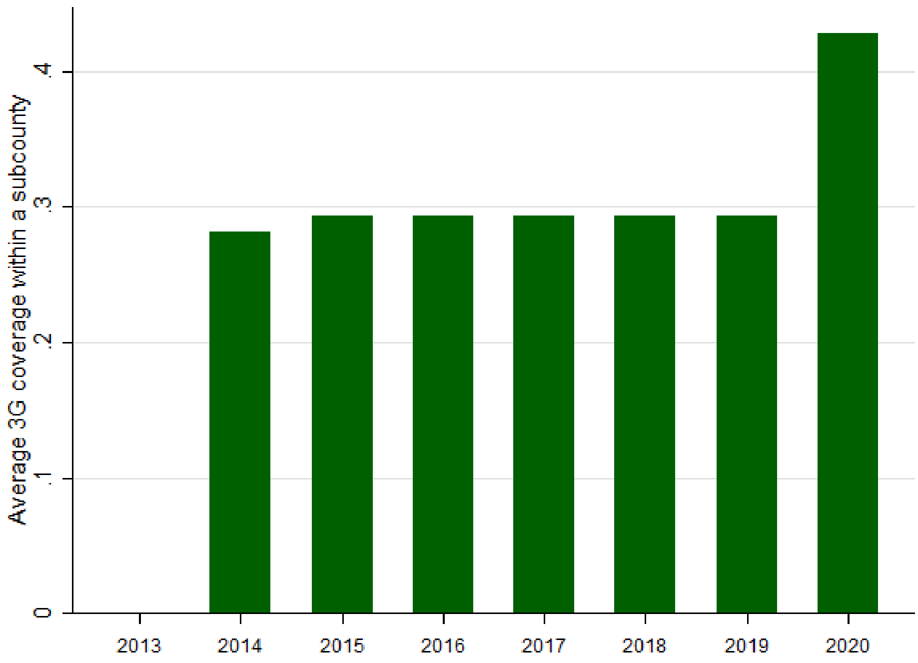
[4] This means, for example, that the data from 2014 was released in January 2014 and in fact depicts the extent of mobile coverage at the end of 2013.

[5] A signal is defined as strong within the 3G system when it is greater than –92 decibel-milliwatts (dBm).

[6] The data is available from https://sedac.ciesin.columbia.edu/data/collection/gpw-v4, accessed April 2023.

Figure 1 shows the expansion of the population-weighted 3G coverage across all subcounties. By the beginning of 2014, the newly expanded 3G network coverage reached about 0.3. This corresponds to strong (weak) mobile internet coverage of about 30 (60) per cent of the local population. Internet penetration increased slightly over the following years, and there was a second wave of expansion in 2020. By the end of our time period, more than 40 per cent of each subcounty's population had access to a strong (or about 90 per cent to a weak) 3G signal.

Figure 1: The expansion of population-weighted 3G coverage across Ugandan subcounties (2013–20)



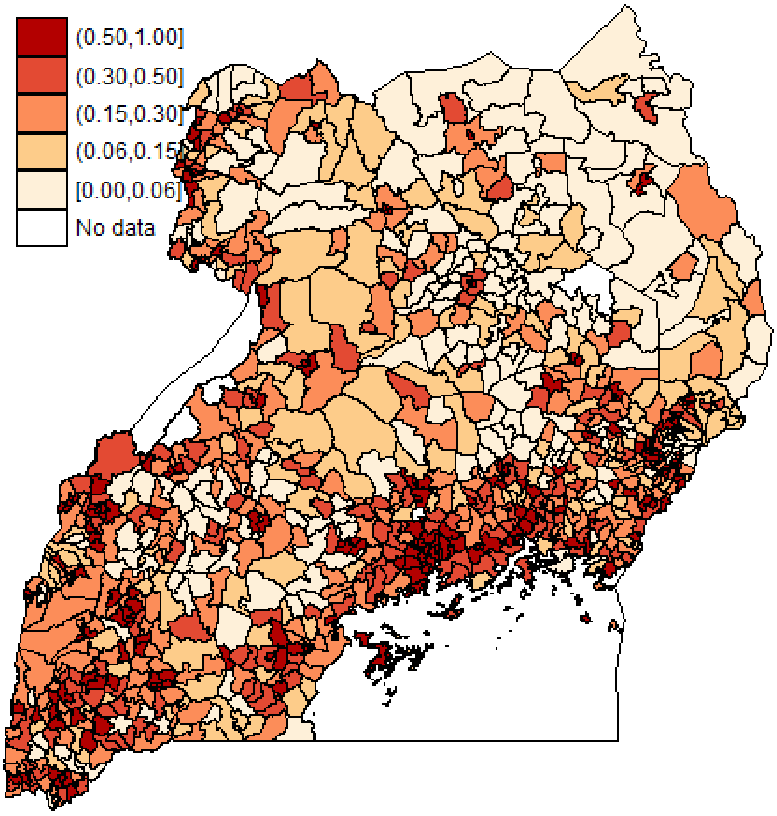
Note: the bars represent an average 3G coverage index measured at the beginning of each year.
Source: authors' compilation based on GSMA data.

The map in Figure 2 shows the spatial distribution of the averages of population-weighted 3G coverage over 2014–20. The coverage by 3G signal varies substantially across Ugandan subcounties, with higher coverage around cities (especially the capital Kampala), as well as in border regions. In the dark red areas at least 50 per cent of the population was covered by sufficient mobile internet.

For the instrumental variable approach we calculate the national roll-out of mobile coverage for every year since 2013. For a subcounty, the national roll-out of mobile coverage is computed as an average across the country, excluding a subcounty itself ('leave-one-out instrument', see Bartik 1991). This gives a shift variable in a shift–share instrument. Our share component is based on the population-weighted average frequency of lightning in each subcounty in 2010, computed using the TRIMM LIS Science Data (Blakeslee 2021) and provided by Kaplan and Lau (2021) as the WGLC v2022.0.0 data set.[7] We introduce an indicator that equals 1 if the population-weighted average frequency of lightning in a subcounty is above the median across the country, and 0 otherwise. This indicator is our share component. The interaction between the shift and the share components is our instrument.

[7] The lightning data is measured at the spatial resolution of 10×10 km and is aggregated to the subcounty level using population-density weights. The WGLC (v2022.0.0) data set provided by Kaplan and Lau (2021) is available at https://zenodo.org/record/6007052#.Y0deYXZBxPa, accessed April 2023.

Figure 2: Spatial distribution of 3G coverage (2014–20)



Note: the map presents the spatial distribution of the averages of population-weighted 3G coverage in Uganda (averaged over 2014–20).

Source: authors' illustration based on GSMA data.

In all regressions, we also control for the population-weighted 2G network coverage, which is used for phone calls but does not provide internet access. The data for 2G coverage comes from the GSMA; the 2G coverage variable is computed in the same way as the 3G coverage variable, combining weak and strong signals and using population-density weights. Moreover, in all regressions we control for the strength of local demand and income dynamics by including population-weighted nighttime lights aggregated at the subcounty level. We extract this data from the annual global VIIRS V2 data set (D'Elvidge et al. 2021).[8]

3.2 Tax records data

To explore firms' formalization and tax behaviour, we rely on novel firm-level panel data provided by the URA, a product of collaboration between the URA and UNU-WIDER. The data comes from administrative CIT returns between 2013 and 2020. The data is rich and contains information on firms' locations, balance sheets and profit–loss account items, as well as details on firms' tax bases (McNabb et al. 2022).[9]

In Uganda, most firms fill in tax forms according to the fiscal year schedule, which goes from 1 July to 30 June of the next year. We assign these records always to the earlier year of the two. For example, a reporting period from 1 July 2014 to 30 June 2015 is assigned to 2014. However, there are also firms

[8] The VIIRS V2 data set is available at https://eogdata.mines.edu/products/vnl/, accessed April 2023.

[9] To the best of our knowledge, in the context of sub-Saharan Africa, only the South African Revenue Service–National Treasury Panel (see Pieterse et al. 2018) and the Ethiopian Revenue and Customs Authority (see Mascagni et al. 2021) would be comparable to the data employed in this study.

that file their tax records based on irregular fiscal years (McNabb et al. 2022). In these cases, we assign each filed record to the year in which the firm reports more months of business activity. For example, we assign a record to the year 2014 if a firm reported more months in 2014 than in 2013 or 2015.[10]

We perform a series of further data-cleaning steps. First, we exclude outliers at the firm level.[11] Second, we identify all firms that report zero turnover and zero cost of sales in all years of tax records—nil-filing firms, which amounts to 27.1 per cent of the firms from our sample. We drop these firms in the tax behaviour analysis, but we still consider their dynamics in the formalization analysis in Section 5.1. Third, we normalize all flow variables (such as sales and profits) to yearly numbers,[12] and deflate all variables using the World Bank GDP deflators for Uganda obtained from the WDI (2021) database. After data cleaning, we aggregate the number of firms, firm performance, and tax behaviour variables at the subcounty-sector level for 2013–20. We distinguish between 22 sectors—including agriculture, mining, manufacturing, 18 service sectors, and 'others', a category that includes firms that were not reporting their sector—and 829 subcounties. This results in a balanced sample of 18,238 subcounty–sector pair observations for each year of our analysis.

We capture firm formalization by recording the yearly number of all firms that file their CIT returns and that actually report meaningful information to the URA within each sector and subcounty.[13] We complement this information with the number of nil-filing firms to further shed light on firm formalization patterns. Beyond studying firm formalization, we also capture firm tax behaviour more broadly by computing subcounty–sector-level sums of the tax base, tax liabilities, deferred tax liabilities, and the effective tax ratio. In particular, the total tax base and total tax liabilities allow us to assess changes in total government revenues from CIT. Deferred tax liabilities and the effective tax ratio, computed as total tax liabilities over profit before tax, are well-established proxies for firms' tax burden and tax compliance (Hanlon and Heitzman 2010; Koivisto et al. 2021; Mascagni and Mengistu 2019).

We complement our analysis of firm tax behaviour by looking at a series of variables, relying only on the firms that report their activities as business and professional (according to Schedule 1).[14] In particular, we assess not only the profits before tax but also the sizes of deductions, losses, and additional benefits that together define the tax base.[15] Finally, we assess the dynamics of firm balance sheet and profit and loss account variables that might highlight the changes in aggregate firm performance. We report aggregate sales, cost of sales, capital, equity, wages, profits before tax, investments, loans, long-term liabilities, and current net assets. The complete list of outcomes and their definitions are given in Table A1 in the Appendix.

Table 1 presents summary statistics of our main variables for 2013–20. All nominal values of subcounty–sector variables, except for the number of firms, are deflated and presented in millions of Ugandan

[10] We dropped 42 duplicate records that were linked to the same year and firm.

[11] To define an outlier, we run a regression on the inverse hyperbolic sine transformation of gross profits controlling for the inverse hyperbolic sine transformation of turnover and year, district and sector fixed effects. We predict errors from this regression and treat the top and bottom 1 per cent of the error distribution as observation outliers. If a firm has observation outliers for more than half of the years of being in the sample, this whole firm is treated as an outlier. This procedure reduces the sample by 2.2 per cent.

[12] For example, if a firm reports flow outcomes only for two months of activity, we multiply them by 6 to make it comparable with other firms that report for 12 months. We do not adjust stock outcomes or balance sheet items in this way.

[13] Figure A1 in the Appendix depicts the number of firms included in the administrative CIT data, including nil-filing firms, and the number of actually reporting firms.

[14] The other three schedules refer to short-term insurance business, mining, and repatriated branch profits. Very few firms report according to these schedules.

[15] According to the accounting rules, if yearly profits before tax after all additional benefits, deductions, and losses from the previous year are negative, the tax base for a firm is set to zero, while the loss this year is set to a positive number.

shillings. In the empirical analysis, all subcounty–sector variables are transformed using the inverse hyperbolic sine transformation, which approximates the natural logarithm transformation but retains negative values and zeros. Our analysis relies on a balanced panel of 145,904 sector–subcounty–year observations. On average, there are 2.2 firms in each subcounty and sector. The average total tax liabilities per subcounty–sector pair constitute 6.794 million UGX. The maximum effective tax rate of 28 per cent is close to the 30 per cent statutory corporate tax rate in Uganda.

Table 1: Summary statistics of the main variables, 2013–20

	Mean	SD	Min.	Max.
Variables at subcounty–sector level				
Total number of firms	2.232	38.858	0	5,073
Number of nil-reporting firms	0.602	8.481	0	1,006
Number of actually reporting firms	1.629	30.886	0	4,096
Total tax liabilities	6.794	182.4	0	20,400
Total tax base	22.732	609	0	67,900
Deferred tax liabilities	28.965	2014	0	388,000
Effective tax ratio (%)	1.797	6.661	0	28.129
Profit before tax	0.806	2,850	–208,000	932,000
Sales	770.5	19,000	0	2,260,000
Cost of sales	626.3	16,400	0	2,010,000
Operational costs	52.863	1,028	0	102,000
Other costs	92.684	1,660	0	209,000
Administrative costs	74.029	1,281	0	149,000
Financial costs	50.843	3,121	0	658,000
Long-term liabilities	645.7	31,710	0	6,070,000
Equity	365	9,486	0	1,470,000
Net current assets	379.2	9,658	0	1,130,000
Wages	51.439	986.8	0	149,000
Capital	728.1	19,500	0	3,040,000
Investments	206.7	28,890	0	6,110,000
Loans	38.137	1,070	0	110,000
Tax base (Schedule 1)	22.299	602.1	0	67,900
Deductions (Schedule 1)	120.7	5,663	–3,880	1300000
Chargeable income (Schedule 1)	–26.540	1,564	–327,000	185,000
Loss previous year (Schedule 1)	3,208	1,175,000	–172	449,000,000
Loss this year (Schedule 1)	3,258	1,175,000	0	449,000,000
Variables at subcounty level				
3G coverage	0.298	0.302	0	1
2G coverage	0.906	0.189	0	1
Nighttime lights	0.271	1.264	0	23.368

Note: summary statistics refer to the key outcome variables from the CIT panel in 2013–20. The number of observations is N = 145,904 for all subcounty–sector variables, except for the effective tax ratio (N = 145,893), and N = 7,461 for subcounty variables. All nominal values of subcounty–sector variables, excluding the number of firms, are deflated and presented in millions of UGX. Nil reporters and outliers are excluded. All variables are described in Appendix Table A1.

Source: authors' compilation based on URA CIT data and further sources.

4 Empirical framework

Our empirical analysis links location and sector variation in firm formalization and tax behaviour over time to the spatio-temporal roll-out of mobile coverage in the country. Identifying a causal impact of mobile internet access on firm formalization, taxation, or firm performance is challenging. The spread of mobile internet coverage might not be random as it reflects private operators' decisions on where and when to install the technology, but also whether to report this information to the GSMA (Buys et al. 2009). Quickly developing and urban markets are more likely to get access to mobile technologies earlier and, hence, the crucial challenge of any empirical strategy is to ensure that the measured effects are due to the spread of mobile coverage and not to underlying spatial differences in the development dynamics of firms. We address the problem of causal identification as follows.

All our regressions are conditional on district–year and sector–year fixed effects as well as on local fluctuations in nighttime light intensity. By that we isolate the effects of the expansion of mobile networks that are not directly driven by the variation in the strength of local demand for mobile technologies. However, to the extent that nighttime lights provide an imperfect proxy for local demand, our simple two-way fixed effect results would remain biased as they might reflect not only the effects of mobile access on firm formalization but also those of a general increase in local demand. Further, the GSMA data might suffer from reporting errors because the information is provided by the operators. To deal with these remaining concerns, we rely on two distinct empirical strategies.

First, we implement a SSIV technique (à la Bartik 1991; Borusyak et al. 2022) to isolate a plausibly exogenous component in the expansion of the 3G mobile network in space and time. The SSIV combines the national average speed of the roll-out of the 3G mobile network with the variation in the local costs of setting up and especially maintaining such a network. We follow the literature by assuming that the costs of mobile network expansion depend on how strongly a place is exposed to lightning (Guriev et al. 2021; Manacorda and Tesei 2020). Electrostatic waves from lightning may cause voltage surges and lead to the destruction of electrical components of digital infrastructure. Africa has the highest lightning activity in the world, with 17.3 strikes/km^2 per year in comparison to 2.9 strikes/km^2 globally. Thus, we expect a lower speed of network expansion in areas with a high frequency of lightning strikes (Andersen et al. 2011, 2012).

Our SSIV is computed as follows:

$$SSIV_{rt} = \text{Coverage}_{\bar{r}t} \times \text{Lightning Exposure}_{r0} \tag{1}$$

where the 'shift' variable for subcounty r, $\text{Coverage}_{\bar{r}t}$, captures the average population-weighted coverage share of the 3G network in year t at the national level, but excluding subcounty r itself ('leave-one-out instrument', Bartik 1991). The 'share' variable, $\text{Lightning Exposure}_{r0}$ is an indicator that equals 1 if the population-weighted frequency of lightning in subcounty r in year 2010 is above the national median, and 0 otherwise. Our identifying assumption in this instrumental variable approach relies on the conditional exogeneity of the coverage shares: conditional on subcounty fixed effects and additional controls for local economic activity and the 2G network expansion, we do not expect that historical exposure to lightning strikes explains firm formalization and tax payment dynamics through other channels than the roll-out of the mobile phone network.

Our instrumental variable strategy consists of the following first- and second-stage equations:

$$\begin{aligned} \text{Coverage}_{rt} &= \lambda SSIV_{rt} + X_{rt}'\beta + \mu_r + \gamma_{dt} + \varepsilon_{rt} && (2) \\ Y_{srt} &= \alpha \widehat{\text{Coverage}}_{rt} + X_{rt}'\beta + \mu_r + \gamma_{dt} + \tau_{st} + \nu_{srt} && (3) \end{aligned}$$

where Y_{srt} denotes an aggregate outcome of sector s and subcounty r in year t. Coverage_{rt} is the treatment variable indicating the population-weighted share of 3G mobile internet coverage in subcounty r and in year t. $SSIV_{rt}$ is the shift–share instrument, specified in Equation (1). $\widehat{\text{Coverage}}_{rt}$ is the predicted treatment variable from the first stage (2). The vector of control variables X_{rt} includes the population-weighted 2G coverage and nighttime light intensity to proxy for GDP and demand fluctuations. Subcounty fixed effects, μ_r, capture all time-invariant sources of spatial variation, whereas the district–year fixed effects, γ_{dt}, absorb broader regional development dynamics. This tightens our identifying variation to the comparison of subcounties located within the same district that experience a different expansion of the mobile network.[16] Finally, the sector–year fixed effects, τ_{st}, included in the second stage, capture

[16] Subcounties are fully nested within districts, and we omit the subscript d from all other variables to ease notation.

yearly sectoral dynamics in Uganda, reflecting fluctuations in global and national demand, or the overall business cycle. For every outcome, we contrast the SSIV approach with a simple two-way fixed effect model that estimates Equation (3), without taking the endogeneity of the mobile network expansion process into account. In all regressions we cluster standard errors at the subcounty level.

Second, we exploit variation over time and space in the expansion of the 3G network coverage between 2013 and 2020 by employing an event study design that relies on the staggered DID estimator developed by De Chaisemartin and d'Haultfoeuille (2020). This estimator addresses the concern that the standard two-way fixed effect (TWFE) regression analysis produces inconsistent and biased results when the treatment is varying across time and the treatment effect is heterogeneous (Goodman-Bacon 2021) by using only never-treated or not-yet-treated units as control groups. The method provides time-specific average treatment effects on the treated (ATTs) for each post-treatment period and then aggregates the effects across time periods to deliver event study coefficients of interest for all post-treatment periods, allowing us to observe an evolution of the treatment effect over time. It also allows testing for pre-trends by comparing locations with early and late access to mobile coverage prior to the treatment.

We estimate treatment effects and pre-trends using the following event study framework:

$$Y_{rst} = \sum_{n=-3}^{6} \alpha_n M_{rt+n} + X_{rt}'\beta + \mu_r + \gamma_{dt} + \tau_{st} + \varepsilon_{rst} \tag{4}$$

where Y_{rst} is an aggregate outcome of sector s and subcounty r in year t, as in Equation (3). The main explanatory variable M_{rt} is a treatment indicator that equals 1 if the population-weighted share of 3G mobile internet coverage in subcounty r in year t exceeds 0.5, which implies that at least 50 per cent of all population in a subcounty is covered by strong 3G signal (or 100 per cent of all population in a subcounty is covered by weak 3G signal, or any linear combination in between). This corresponds to the upper quartile (75th percentile) of the continuous treatment variable.[17] As before, μ_r, γ_{dt}, and τ_{st} denote subcounty, district–year, and sector–year fixed effects, respectively, and X_{rt} denotes the vector of control variables including the population-weighted 2G coverage and nighttime light intensity. Standard errors are clustered at the subcounty level.

We use the results from the staggered DID approach only as suggestive evidence to complement our SSIV estimates. Our event study design is based on more limited time variation: early-treated locations have only one pre-treatment year (2013), while the late-treated locations provide only one post-treatment year (2020). In both approaches we estimate the effect of introducing mobile internet on the number of tax-paying firms and on a range of aggregated firm tax behaviour and performance outcomes at the subcounty–sector level.[18]

5 Results

Our empirical results present the impact of mobile internet expansion on firm formalization, total tax liabilities, and other aggregate measures of localized firm performance at the subcounty–sector level. First, we study formalization by assessing how mobile internet changes the number of firms that file

[17] An alternative specification defining treatment above the median of the population-weighted share of 3G mobile internet coverage that equals 0.2 yields comparable results.

[18] In TWFE, SSIV, and staggered DID models we control for subcounty fixed effects but not for subcounty–sector fixed effects, as our treatment varies at the subcounty level. Thus, we look at within-subcounty variation after controlling for district–year and sector–year fixed effects. Accounting for subcounty–sector fixed effects produces identical results in TWFE and SSIV models. We expect to obtain the same results in staggered DID models if we would control for subcounty–sector fixed effects, but STATA command did_multiplegt was extremely slow in producing any results.

their CIT with the Ugandan tax authority. In a second step, we explore how mobile internet influences firms' tax behaviour by assessing changes in the total tax base and liabilities and additional variables from the CIT returns used to calculate the tax base. Finally, we assess whether changes in firms' tax base are more likely to come from composition effects, or whether they point to changes in tax reporting behaviour.

5.1 Mobile internet and tax compliance on the extensive margin

Our first hypothesis is that mobile internet access may facilitate the filing of tax declarations and by that reduce the costs of formalization, resulting in more firms submitting their tax records over time. Therefore, our first outcome of interest is the number of firms filing CIT with the URA within each subcounty–sector cell.

Table 2 reports our baseline formalization results, contrasting TWFE estimates that rely on Equation (3) with SSIV estimates that correct for the endogeneity of the network roll-out by first estimating Equation (2) and only utilizing the exogenous part of the predicted 3G coverage at the second stage. Both the TWFE and the second-stage SSIV regressions control for population-weighted coverage of the 2G mobile network and nighttime lights, as well as subcounty, district–year, and sector–year fixed effects. We contrast three measures of the number of filing firms: the total number of firms filing CIT returns with the URA and its decomposition into firms that report nil activities in all years of their reporting period and firms that report actual non-zero information. The number of firms that file actual tax reports reflects firm formalization dynamics more precisely as nil-reporting firms are either not fully operational or not fully engaging with the tax authorities.

Table 2: Mobile network coverage and firm formalization

	asinh Number of firms					
	By reporting status			Actual reporting by sector		
	All	Nil-reporting	Actual reporting	Primary	Secondary	Tertiary
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: TWFE						
3G coverage	0.262***	0.181***	0.212***	0.185***	0.267***	0.230***
	(0.018)	(0.016)	(0.017)	(0.029)	(0.032)	(0.019)
Panel B: SSIV						
3G coverage	0.641***	0.462***	0.488***	0.538*	0.361	0.620***
	(0.168)	(0.145)	(0.142)	(0.303)	(0.340)	(0.161)
F-stat. first stage	12.68	12.68	12.68	11.87	11.87	12.66
No. observations	145,904	145,904	145,904	13,264	13,264	112,744
No. subcounties	829	829	829	829	829	829
Controls	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓

Note: results from TWFE and the second stage of SSIV regressions are reported. The dependent variables measure the inverse hyperbolic sine of the number of total reporting firms by subcounty–sector and year (column 1), the number of firms that report nil information (zero turnover and zero costs of sales for all years of being in the sample, column 2), and the number of firms actually reporting to the URA (excluding nil reporters and outliers, column 3). This latter group is split by sectors in columns 4–6. All specifications control for 2G coverage and nighttime light intensity, and include subcounty, district–year, and sector–year fixed effects. Standard errors, clustered at the subcounty level, are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations based on GSMA and URA CIT data.

The TWFE results show significant increases in firm formalization along all three dimensions. With the expansion in mobile coverage, the number of firms filing CIT returns to the tax authority significantly increases (column 1 of Panel A). At the same time, the number of firms that file nil results on their CIT returns increases to a comparable extent (column 2 of Panel A). Thus, not all firms improve their tax filing behaviour, despite engaging more with tax authorities. Nonetheless, the increase in the number of firms filing actual reports with the URA (column 3 of Panel A) is economically relevant. For instance, column 3 shows that a 10 point increase in 3G mobile coverage is linked to an about 2.4 per cent increase in the number of firms that actually report activities to the URA.[19] This type of firm formalization is also not restricted to a specific sector, but is present in primary, secondary, and tertiary sectors (columns 4–6).[20] The TWFE regressions explicitly control for local economic development, the main determinant of mobile coverage: by conditioning all results on local nighttime light intensity, we capture the fact that the expansion of mobile coverage responds strongly to local demand and economic activities. The TWFE regressions also control for 2G coverage to exclude the potential effects of the 2G network expansion. Nonetheless, the concern remains that further local unobserved characteristics might drive both mobile network expansion and firm formalization.

To deal with this endogeneity concern, we instrument the population-weighted local 3G signal coverage by the indicator that the population-weighted frequency of lightning is greater than the national median interacted with the national roll-out of network coverage. The first stage of SSIV regression (see Table A2 in the Appendix) confirms that the spatial roll-out of the 3G network was significantly slower in places that were historically more prone to lightning strikes and hence faced higher relative costs of network expansion and prospective maintenance. The F-statistic of about 12 passes the minimum requirements for instrument validity. The SSIV estimates presented in Panel B of Table 2 confirm the TWFE results: they are positive and statistically significant for all types of reporting firms. The increase in coefficient magnitudes (the SSIV coefficients are two to three times larger than the TWFE coefficients) could, among others, reflect a local average treatment effect (LATE) interpretation: mobile access expansion could especially be relevant for firm formalization in places that suffer less from electricity surges. A 10 point increase in 3G mobile coverage leads to an about 6.3 per cent increase in the number of firms that actually report activities to the URA (column 3 of Panel B). Particularly, the tertiary sector demonstrates strong firm formalization in column 6 of Panel B. The increase in the number of actually reporting firms in the secondary sector is not statistically significant (column 5 of Panel B) and in the primary sector is only marginally significant (column 4 of Panel B).

Table 3 reports the results of the heterogeneity analysis, in which we split the sample by economic activity proxied by nighttime light intensity above and below the median (columns 1–2), by high and low wealth proxied by nighttime light intensity per capita above and below the median (columns 3–4), and by the number of firms per subcounty–sector above and below the median of three (columns 5–6). The positive trend in firm formalization with the expansion of mobile internet is driven by economically active and wealthier subcounties as well as by subcounty–sector pairs with at least three firms. The weak F-statistics from the first stage of SSIV regression also indicate that the national expansion trends do not have predictive power for the local expansion of the mobile network in the less developed and more rural places, not even in combination with proxies of the relative costs of expansion. Taken together, our results confirm that firm formalization occurs in populated areas with higher economic activity.

The staggered DID estimates (De Chaisemartin and d'Haultfoeuille 2020) reported in Figure 3 confirm our SSIV results. Upon receiving the treatment, the number of firms actually reporting economic activity increases on average by 20 per cent.[21] This effect is stable for all post-treatment periods. Thus both estimation approaches offer consistent results that show an increased firm formalization on the extensive

---

[19] Our measure of mobile signal strength converts weak mobile signal coverage into strong signal coverage at the rate of 1:2. Thus, a 10 point increase in signal coverage can be achieved either by 10 percentage points more people covered by a strong 3G signal within the subcounty, or a 20 percentage point increase in the population covered by a weak 3G signal, or a combination of the two. The 2.4 percent increase is calculated as $[(\exp(0.212) - 1)/10] \times 100\%$.

[20] The primary sector includes agriculture and mining, the secondary sector includes manufacturing and construction, and the tertiary sector includes the remaining industries, but excludes cases where the firms were not reporting their sector.

[21] Recall that the treatment is discretized to take value 1 when 3G coverage reaches the upper quartile (75th percentile) of our sample's distribution—that is, when at least 50 per cent of the local population are covered by a strong signal.

margin as a result of expanded mobile internet access. We, however, cannot distinguish whether this result is driven by the formalization of existing firms or by new formal firms entering the market.

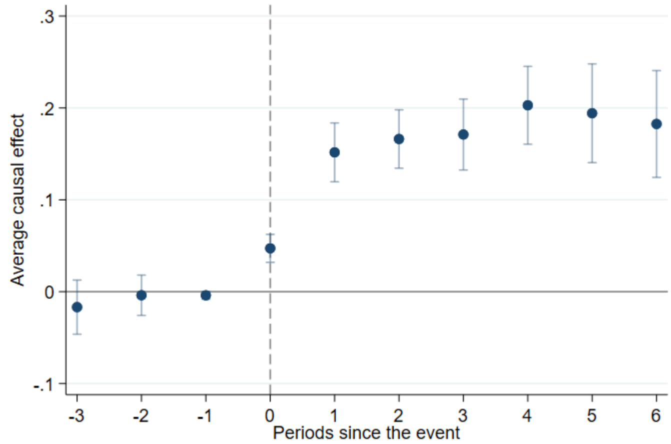
Table 3: Heterogeneous effects of mobile coverage on firm formalization

	asinh Number of firms					
	By economic activity		By wealth		By no. firms per subcounty	
	High	Low	High	Low	≥ 3	< 3
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: TWFE						
3G coverage	0.276***	0.004	0.271***	0.005	0.272***	0.012
	(0.024)	(0.009)	(0.024)	(0.009)	(0.022)	(0.008)
Panel B: SSIV						
3G coverage	0.584***	–0.131	0.564***	–0.188	0.730***	–0.957
	(0.221)	(0.354)	(0.202)	(0.428)	(0.214)	(2.613)
F-stat. first stage	8.84	0.54	9.80	0.48	9.84	0.14
No. observations	73,040	72,864	73,040	72,864	79,552	66,352
No. subcounties	415	414	415	414	452	377
Controls	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓

Note: results from TWFE and the second stage of SSIV regressions are reported. Columns 1–2 split the sample by economic activity (proxied by nighttime lights above and below the median), columns 3–4 split the sample by high and low wealth (proxied by nighttime lights per capita above and below the median), and columns 5–6 split the sample by the number of firms per subcounty–sector (above and below the median of three). The dependent variable is the inverse hyperbolic sine of the total number of firms actually reporting to the URA by subcounty–sector and year. All specifications control for 2G coverage and nighttime light intensity, and include subcounty, district–year, and sector–year fixed effects. Standard errors, clustered at the subcounty level, are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations based on GSMA and URA CIT data.

Figure 3: Event study DID effects of 3G coverage on firm formalization: total number of actual reporting firms



Note: staggered DID estimates (De Chaisemartin and d'Haultfoeuille 2020). The dependent variable captures the inverse hyperbolic sine of the total number of firms actually reporting to the URA. The control variables include 2G coverage and nighttime light intensity as well as subcounty, district–year, and sector–year fixed effects.

Source: authors' illustration based on URA and GSMA data.

5.2 Mobile internet and tax compliance on the intensive margin

Ex ante it is not clear whether a higher rate of firm formalization also results in higher aggregate tax revenues from CIT as the newly formalizing firms might still fail to report positive profits. Thus, we test whether the expansion of mobile coverage also leads to more tax revenues. Using the administrative records, we can assess its effects on state tax revenues by looking at the tax base, tax liabilities, and deferred tax liabilities. The total tax base is defined as the total chargeable income added up from all tax schedules (business and profession, short-term insurance business, mining and repatriated profits).

The total tax liabilities combine tax payments from all tax schedules. We compute it as the sum of chargeable incomes from each schedule multiplied by the respective tax rate: 30 per cent for business and profession and insurance business, 25 per cent for mining, and 15 per cent for repatriated branch profits. These two variables directly measure the total tax revenues from CIT that the state can collect at the subcounty–sector level.

Additionally, we report two other indicators of firm tax behaviour that are considered in the literature as proxies for tax compliance (Hanlon and Heitzman 2010): deferred tax liabilities and the effective tax ratio (ETR). Deferred tax liabilities include all taxes that have not yet been paid but are due in the future. The justifications for changes in deferred tax liabilities are difficult to verify as the amounts and timing of future chargeable income depend on managers' expectations of earnings for many years into the future. This provides space for manipulation (Miller and Skinner 1998). Thus, higher deferred tax liabilities might also signal lower tax compliance in the future. They can also signal aggressive tax planning and even potential tax risks if firms are strategically trying to reduce their current tax liabilities, which will result in higher tax obligations. The ETR is computed as the ratio of the total tax liabilities to total profit before tax at the firm level, and then averaged at the subcounty–sector level. Thus, it captures the average rate of tax per unit of income per year. Higher ETR implies higher tax compliance. Calculating an annual ETR can be difficult due to the diverse nature of tax liabilities within a fiscal year. These liabilities may pertain to the current year, represent delayed payments from previous fiscal years, encompass interest or penalties, or even involve prepayments for future years (Koivisto et al. 2021). However, it can be assumed that all these components of future or late tax payments balance out over the eight fiscal years. In addition to the dynamics of the tax base, we consider total reported profit before tax, which also incorporates the losses reported by the firms and represents a ground for the tax base calculations.

The full sample results in both panels of Table 4 show an increase in total tax liabilities as well as in the total tax base with expanding 3G mobile coverage in both models (columns 1–2). The increases in both variables are very large and at least marginally significant even in the SSIV models. A 10 percentage point increase in mobile coverage results in a more than 50 per cent expansion of the total tax base (column 1) and in a more than 47 per cent increase of total tax liabilities in the SSIV model (column 2). Thus, mobile coverage triggers not only firm formalization, but also results in more taxable income being reported and an overall increase of state capacity through the expanding tax base and tax liabilities.

At the same time, we observe the other changes in firms' tax behaviour that do not contribute to higher tax compliance on the intensive margin. In particular, Table 4 reports a statistically significant increase in the deferred tax liabilities, which might result from 'tax engineering' (column 3). Firms might strategically delay tax liabilities that reduce current revenues collected by the tax authorities. A 10 percentage point increase in mobile coverage is associated with a 15 per cent increase of deferred tax liabilities. Overall, higher tax liabilities lead to a higher ETR per subcounty and sector (column 4), which is statistically significant in the TWFE but is not in the SSIV model. Thus, we do not find robust evidence of higher tax compliance on the intensive margin, which might be explained by the simultaneous increase in tax liabilities and deferred tax liabilities, leading to no changes in the current average rate of tax per unit of income. Equally, the ETR might have distinct patterns for firms with different characteristics. For example, Bachas et al. (2023) find that the ETRs differ by firm size. Because we trace the ETR averaged from all types of firms at the subcounty–sector level, such trends might not be visible.[22]

[22] We also provide the results for the subcounty–sector ETR computed as the ratio of the subcounty–sector aggregated tax liabilities to the subcounty–sector aggregated profits before tax, but they are statistically insignificant too.

Table 4: The effects of mobile coverage on firm tax outcomes

	Full sample				
	asinh Total tax base	*asinh* Total tax liabilities	*asinh* Deferred tax liabilities	Effective tax ratio	*asinh* Profit before tax
	(1)	(2)	(3)	(4)	(5)
Panel A: TWFE					
3G coverage	1.261***	1.172***	0.331***	0.009***	0.195
	(0.115)	(0.106)	(0.060)	(0.002)	(0.188)
Panel B: SSIV					
3G coverage	1.874*	1.744*	0.938*	0.011	0.437
	(1.034)	(0.954)	(0.549)	(0.015)	(1.725)
No. observations	145,904	145,904	145,904	145,893	145,904
No. subcounties	829	829	829	829	829
	Balanced subsample				
Panel C: TWFE					
3G coverage	0.254**	0.236**	–0.006	0.005*	0.164
	(0.101)	(0.093)	(0.046)	(0.003)	(0.219)
Panel D: SSIV					
3G coverage	–1.148	–1.071	0.035	0,018	2.016
	(0.821)	(0.761)	(0.468)	(0.022)	(2.295)
No. observations	64,240	64,240	64,240	64,240	64,240
No. subcounties	365	365	365	365	365
Controls	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓

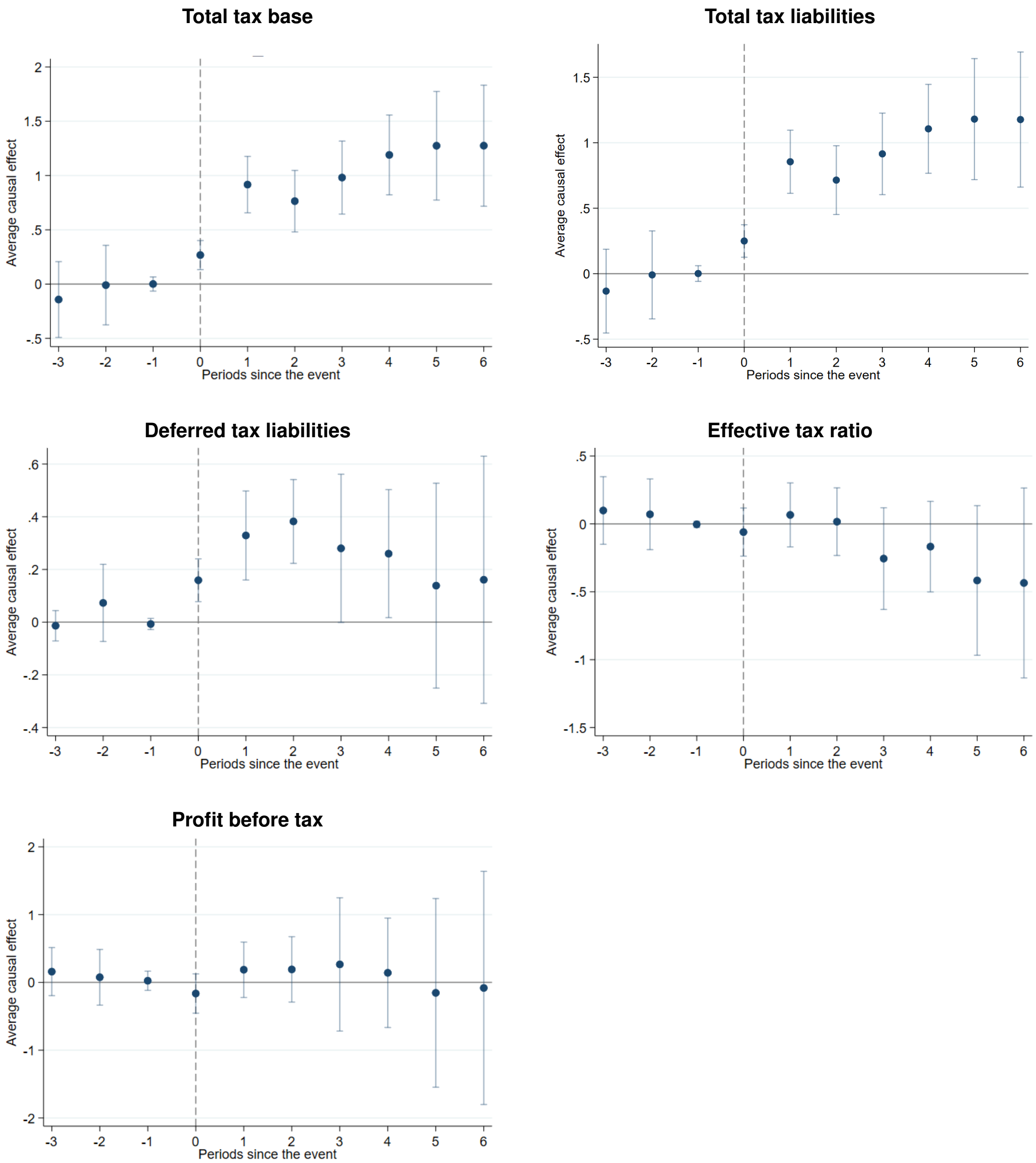
Note: results from TWFE and the second stage of SSIV regressions rely on the sample of firms actually reporting to the URA (excluding nil reporters and outliers). The dependent variables include the inverse hyperbolic sine of total tax liabilities, of the total tax base, of deferred tax liabilities, of total profits before tax, and the ETR for each subcounty–sector and year combination. All specifications control for the share of 2G coverage and nighttime light intensity and include subcounty, district–year, and sector–year fixed effects. The full sample (Panels A and B) stands for aggregates based on all firms in the data. The balanced subsample (Panels C and D) stands for aggregates based on the firms that report to the URA from 2013 to 2019/20. The F-statistic for the first stage of the SSIV for the full sample is 12.68 and for the balanced sample is 10.94. Standard errors, clustered at the subcounty level, are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations based on GSMA and URA CIT data.

We do not observe statistically significant increases in the total amount of profits aggregated at the subcounty–sector level (column 5). In Section 5.3 we explore in more detail the components of the profit before tax to understand its insignificant change. In this section, to shed light on the opposite results for the tax base and for the profit before tax, we investigate the tax base and its components for firms reporting business and professional activity (Schedule 1) only. The firm tax base is calculated by subtracting the losses of the previous year from the current chargeable income. If the losses are higher than the chargeable income, then the firm tax base is set to zero and the negative difference is recorded as the losses to carry forward to the next reporting period. Table A3 in the Appendix presents the TWFE and SSIV results for the aggregate subcounty–sector tax base, losses this year to be carried forward, chargeable income from profits and gains, losses from the previous year, and deductions. Although the aggregate tax base is increasing with higher 3G coverage (as in Table 4), we also observe increases in the aggregate losses to carry forward that confirms that firms are diverse in their tax behaviours. Some firms report fewer losses and thus contribute to the increase of the aggregate tax base. At the same time, the aggregate losses and deductions from the current year increase, while chargeable income remains unchanged. This could signal that most firms learn how to strategically adjust their reporting behaviour to minimize their tax burden in a respective year.

The staggered DID estimates reported in Figure 4 confirm the TWFE and SSIV results. They show marked increases in the tax base and total tax liabilities upon the roll-out of mobile network coverage, whereas deferred tax liabilities increase only temporarily but decline again and become insignificant five years after the roll-out of the network. By contrast, we see no statistically significant adjustment in the ETRs. While the size of the formalized economy increases, we do not see an increase in the average tax burden charged on reported profits.

Figure 4: Event study DID effects of 3G coverage on tax outcomes



Note: staggered DID estimates (De Chaisemartin and d'Haultfoeuille 2020). The dependent variable captures the inverse hyperbolic sine of total tax base, total tax liabilities, deferred tax liabilities, effective tax ratio, and profit before tax. The control variables include the local share of population-weighted 2G coverage and nighttime light intensity as well as subcounty, district–year, and sector–year fixed effects.

Source: authors' illustration based on URA and GSMA data.

The results presented in Panels A and B of Table 4 are based on subcounty–sector aggregates that include all actually reporting firms in the sample. Due to this aggregation, we cannot immediately see whether the tax base increases because of formalization or because of better firm performance in response to mobile coverage expansion. While insignificant results for aggregate profits before tax might favour the former explanation, we repeat the same analysis for a balanced subsample of reporting firms that remain in the sample from 2013 till 2019 or 2020. These results are based on a substantially smaller number of firms. As the number of firms reporting to URA is increasing over time, those that enter the sample after 2013 are not considered. The results are presented in Panels C and D of Table 4. The magnitudes of the coefficients of interest for the total tax base and tax liabilities are substantially smaller in the TWFE models in comparison to those presented in Panel A, but remain significant at the 5 per cent level (columns 1 and 2). In the SSIV models, however, they lose their significance and even turn negative. The estimates for other outcomes are insignificant in both models. Comparing aggregates based on the unbalanced and balanced sample of firms suggests that after controlling for demand factors (nighttime lights) and fixed effects, 3G expansion improves government capacity by stimulating firm formalization rather than firm performance. We confirm this conclusion by looking at other aggregate performance outcomes in Section 5.3.

5.3 Mobile internet and aggregate firm outcomes

As hypothesized before, the expansion of mobile technologies could also contribute to the local tax base by enhancing firm performance. If digital services result in increasing liquidity and by that improve firms' access to borrowing and enable new investments, this will result in higher firm productivity and potentially also higher tax payments. The results reported in columns 1 and 5 of Table 4 document indeed an increase in the total tax base in response to mobile coverage expansion, but no improvements in total profits reported by firms in the local economy. Such discrepancies could suggest that many non-profitable firms are formalizing with the mobile network expansion and/or that strategic misreporting takes place. The information from balance sheet and profit and loss account items reported to the tax authorities can provide further insights on the structural changes that might be contributing to such dynamics. We focus on the local dynamics of the formal economy by aggregating some firm outcomes at the subcounty–sector level.

Table 5 reports changes in aggregate outcomes of tax filing firms in response to the expansion of 3G coverage. In Panels A and B these outcomes are computed at the subcounty–sector level using all firms, while in Panels C and D they are computed for a subsample of firms that report during the whole time period of 2013–19/20. We focus on sales, wage bill, investments, and other asset accumulation components. The results of TWFE show substantial increases in all these variables (Panel A). Although increases in total investment, net current assets, and loans are at a smaller scale, the estimated coefficients have substantial magnitudes and show from 5 (in investments) to 34 (in sales) per cent increases with a 10 point increase in local mobile network coverage. These positive dynamics are also confirmed by the staggered DID estimators presented in Figure A2 in the Appendix. Only for the wage bill does the effect become insignificant six years after the local introduction of the mobile network. By contrast, the SSIV results reported in Panel B of Table 5 are statistically weaker, as they show only marginally significant improvements for sales, wage bill, and long-term liabilities, but no significant effects for other outcomes. Overall, these results show growth of the aggregate performance of the formal sector along several dimensions.

To assess whether the improvements in the aggregate firm performance outcomes are due to the growth of firms that were part of the formal economy from the onset or arose because of new firms becoming part of the formal economy, we repeat the TWFE and SSIV models for the same aggregate firm performance outcomes for a balanced subsample of firms (Panels C and D). The estimates from the TWFE models are much smaller than those presented in Panel A, and significant only for sales, capital, and loans. The estimates from the SSIV models are not significant, besides investments. The results from Panels C and D suggest that the positive effects of the expansion of the mobile network presented in Panels A and B must be mainly due to new formal firms appearing on the market, but not due to improvements in the performance of existing firms.

Table 5: The effects of 3G coverage on aggregate firm performance

	asinh Sales (1)	*asinh* Wages (2)	*asinh* Investment (3)	*asinh* Long-term liabilities (4)	*asinh* Equity (5)	*asinh* Capital (6)	*asinh* Net current assets (7)	*asinh* Loans (8)
	Full sample							
Panel A: TWFE								
3G coverage	1.486***	1.138***	0.389***	1.132***	1.273***	1.265***	0.752***	0.758***
	(0.154)	(0.122)	(0.070)	(0.136)	(0.140)	(0.143)	(0.175)	(0.108)
Panel B: SSIV								
3G coverage	2.525*	1.782*	0.672	1.988*	1.284	1.779	-0.318	1.259
	(1.418)	(1.057)	(0.575)	(1.187)	(1.255)	(1.287)	(1.419)	(0.954)
No. observations	145,904	145,904	145,904	145,904	145,904	145,904	145,904	145,904
No. subcounties	829	829	829	829	829	829	829	829
	Balanced subsample							
Panel C: TWFE								
3G coverage	0.292***	0.133	0.024	–0.015	0.107	0.178*	–0.205	0.271***
	(0.100)	(0.095)	(0.059)	(0.120)	(0.095)	(0.096)	(0.180)	(0.093)
Panel D: SSIV								
3G coverage	–0.552	–0.562	1.048*	–0.353	–1.506	–1.009	–0.248	–*0.099
	(0.743)	(0.827)	(0.540)	(0.884)	(0.964)	(0.819)	(1.651)	(0.815)
No. observations	64,240	64,240	64,240	64,240	64,240	64,240	64,240	64,240
No. subcounties	365	365	365	365	365	365	365	365
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Note: results from TWFE and the second stage of SSIV regressions rely on the sample of firms actually reporting to the URA (excluding nil reporters and outliers). The dependent variables include the inverse hyperbolic sine of total sales, wage bill, cost of investments, long-term liabilities, equity, capital, net current assets, and loans for each sector–subcounty pair and year. All specifications control for the share of 2G coverage and nighttime light intensity and include subcounty, district–year, and sector–year fixed effects. The full sample (Panels A and B) stands for aggregates based on all firms in the data. The balanced subsample (Panels C and D) stands for aggregates based on the firms that report to the URA from 2013 to 2019/20. The F-statistic for the first stage of the SSIV for the full sample is 12.68 and for the balanced sample is 10.94. Standard errors, clustered at the subcounty level, are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations based on GSMA and URA CIT data.

The results in Table 6 zoom into the firm cost structure instead. As before, we compute aggregate outcomes at the subcounty–sector level using the whole sample of firms (Panels A and B), and using only firms that stay in the sample for the whole period (Panel C and D). In Panels A and B we document substantial increases in the main reported cost components with the mobile network expansion. The TWFE results are all positive and highly significant, showing substantial increases in aggregate firm costs reported to the tax authorities. These results are supported by the staggered DID event study estimates reported in Figure A3 in the Appendix. Moreover, the SSIV results also show significant increases in aggregate costs. Significant increases in the costs along with the increases in sales with the expansion of 3G coverage explain our insignificant results for profit before tax (column 5, Table 4)

Focusing on the aggregate cost variables computed using the balanced subsample of firms and the results from the TWFE models (Panels C), we observe that the reported costs increase also among the firms that were filing taxes from the beginning. The SSIV estimates, in contrast, are negative and not significant (Panel D). Our SSIV approach thus cannot confirm the conclusions from the TWFE models, suggesting no cost adjustments among firms that existed in the whole time period from 2013 to 2019/20.

Table 6: The effect of 3G coverage on total reported costs

	asinh Cost of sales (1)	*asinh* Operational costs (2)	*asinh* Admin. costs (3)	*asinh* Financial costs (4)	*asinh* Other costs (5)
			Full sample		
Panel A: TWFE					
3G coverage	1.357***	1.332***	1.267***	1.110***	1.336***
	(0.131)	(0.135)	(0.130)	(0.112)	(0.140)
Panel B: SSIV					
3G coverage	1.842	2.372*	1.951*	1.884*	2.540**
	(1.297)	(1.247)	(1.183)	(0.987)	(1.248)
No. observations	145,904	145,904	145,904	145,904	145,904
No. subcounties	829	829	829	829	829
			Balanced subsample		
Panel C: TWFE					
3G coverage	0.310***	0.216***	0.232***	0.149**	0.250***
	(0.088)	(0.074)	(0.082)	(0.075)	(0.082)
Panel D: SSIV					
3G coverage	–1.162	–0.763	–1.021	–0.787	–1.122
	(0.925)	(0.692)	(0.811)	(0.678)	(0.779)
No. observations	64,240	64,240	64,240	64,240	64,240
No. subcounties	365	365	365	365	365
Controls	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓

Note: results from TWFE and the second stage of SSIV regressions are reported, relying on the sample of firms actually reporting to the URA (excluding nil reporters and outliers). The dependent variables include the inverse hyperbolic sine of costs of sales, operational costs, administrative costs, financial costs, and other costs for each sector–subcounty pair and year. All specifications control for the share of 2G coverage and nighttime light intensity and include subcounty, district–year, and sector–year fixed effects. The full sample (Panels A and B) stands for aggregates based on all firms in the data. The balanced subsample (Panels C and D) stands for aggregates based on the firms that report to the URA from 2013 to 2019/20. The F-statistic for the first stage of the SSIV for the full sample is 12.68 and for the balanced sample is 10.94. Standard errors, clustered at the subcounty level, are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations based on GSMA and URA CIT data.

6 Conclusion

Our study documents positive effects of the expansion of 3G mobile network coverage on firm formalization and tax compliance in Uganda between 2013 and 2020. During this period, the 3G mobile network was introduced and rolled out across the country. Our results provide evidence that the local expansion of 3G mobile internet is associated with a higher number of firms being included in CIT records. This trend signals increased firm formalization, which offers us an indirect proxy for higher tax compliance. However, it is unclear whether this trend is driven by a number of new formal firms entering the market, or by firms moving from the informal to the formal sector. At the same time, we also document an increase in the number of nil-filing firms linked to the improved mobile network access, which shows that being registered with the URA does not yet translate directly into better reporting behaviour.

Our results, in addition, show that firm formalization leads to an increase in the local tax base as well as in higher tax liabilities in the post-treatment years. However, we see no comparable increase in the sum of firm profits, which suggests that the number of firms reporting losses is increasing as well. To examine the potential channels behind this relationship, we make use of the rich set of profit and loss account variables taken from the CIT returns. We find that firm formalization leads to increased sales in the formal economy, but also to larger total costs including wages, operational costs, administrative costs, financial costs, and other costs. For this reason, the sum of total profits before tax is not changing, as it is calculated by subtracting all costs from sales and other income. A more detailed investigation of the tax base components suggests that firms report larger losses and deductions, especially those carried forward to the next year, which decreases their tax burden in the respective year. This finding, together with the significant increase in deferred tax liabilities, might indirectly signal lower tax compliance. Finally, using outcomes aggregated for the subsample of firms that remain in the data during the whole period (2013–19/20), we provide suggestive evidence that 3G mobile coverage expansion does not improve firm performance.

We conclude that the arrival of mobile internet leads to an overall increase in the size of the formal economy, also increasing the local tax base. The increase of fiscal resources generated by CIT bears in turn the potential to translate into improved state capacity. These findings highlight that access to mobile technologies is a meaningful prerequisite for employing different policies aiming at improving local tax collection (e.g., through e-filing, mobile apps, online payment options, and others) and for increasing the overall size of the formal economy. These findings add gains in taxation as one additional important dimension to our understanding of the socio-economic effects of mobile technologies.

References

Aiken, E., S. Bellue, D. Karlan, C. Udry, and J.E. Blumenstock (2022). 'Machine Learning and Phone Data Can Improve Targeting of Humanitarian Aid'. *Nature*, 603(7903): 864–70. https://doi.org/10.1038/s41586-022-04484-9

Aker, J.C., and C. Ksoll (2019). 'Call Me Educated: Evidence from a Mobile Phone Experiment in Niger'. *Economics of Education Review*, 72(C): 239–57. https://doi.org/10.1016/j.econedurev.2019.05.001

Aker, J.C., C. Ksoll, and T.J. Lybbert (2012). 'Can Mobile Phones Improve Learning? Evidence from a Field Experiment in Niger'. *American Economic Journal: Applied Economics*, 4(4): 94–120. https://doi.org/10.1257/app.4.4.94

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

Ali, M., A.B. Shifa, A. Shimeles, and F. Woldeyes (2021). ‘Building Fiscal Capacity in Developing Countries: Evidence on the Role of Information Technology’. *National Tax Journal*, 74(3): 591–620. https://doi.org/10.1086/715511

Almunia, M., J. Hjort, J. Knebelmann, and L. Tian (2022). ‘Strategic or Confused Firms? Evidence from ‘Missing’ Transactions in Uganda’. *Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_01180

Andersen, T.B., J. Bentzen, C.-J. Dalgaard, and P. Selaya (2011). ‘Does the Internet Reduce Corruption? Evidence from US States and Across Countries’. *The World Bank Economic Review*, 25(3): 387–417. https://doi.org/10.1093/wber/lhr025

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

Araujo-Bonjean, C., and G. Chambas (2004). ‘Taxing the Urban Unrecorded Economy in Sub-Saharan Africa’. In J. Alm, J. Martinez-Vazquez, and S. Wallace (eds), *Contributions to Economic Analysis: Taxing the Hard-to Tax: Lessons from Theory and Practice*. Amsterdam: Elsevier. https://doi.org/10.1016/S0573-8555(04)68815-8

Bachas, P., A. Brockmeyer, R. Dom, and C. Semelet (2023). ‘Effective Tax Rates and Firm Size’. World Bank Policy Research Working Paper 10312. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-10312

Bahia, K., P. Castells, G. Cruz, T. Masaki, C. Rodríguez-Castelán, and V. Sanfelice (2023). ‘Mobile Broadband, Poverty, and Labor Outcomes in Tanzania’. *World Bank Economic Review*, 37(2): 235–56. https://doi.org/10.1093/wber/lhad003

Bartik, T.J. (1991). *Who Benefits from State and Local Economic Development Policies?*. Kalamazoo, MI: Upjohn Press. https://doi.org/10.17848/9780585223940

Bellon, M., E. Dabla-Norris, S. Khalid, and F. Lima (2022). ‘Digitalization to Improve Tax Compliance: Evidence from VAT e-Invoicing in Peru’. *Journal of Public Economics*, 210: 104661. https://doi.org/10.1016/j.jpubeco.2022.104661

Benzarti, Y. (2020). ‘How Taxing Is Tax Filing? Using Revealed Preferences to Estimate Compliance Costs’. *American Economic Journal: Economic Policy*, 12(4): 38–57. https://doi.org/10.1257/pol.20180664

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

Besley, T., A. Jensen, and T. Persson (2021). ‘Norms, Enforcement, and Tax Evasion’. *Review of Economics and Statistics*, 105(4): 1–10. https://doi.org/10.1162/rest_a_01123

Blakeslee, R.J. (2021). ‘Lightning Imaging Sensor (LIS) on TRMM Science Data’. Version 4. Washington, DC: NASA.

Bloom, N., and J. Van Reenen (2007). ‘Measuring and Explaining Management Practices Across Firms and Countries’. *The Quarterly Journal of Economics*, 122(4): 1351–408. https://doi.org/10.1162/qjec.2007.122.4.1351

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013). ‘Does Management Matter? Evidence from India’. *The Quarterly Journal of Economics*, 128(1): 1–51. https://doi.org/10.1093/qje/qjs044

Boly, A. (2018). ‘On the Short- and Medium-Term Effects of Formalisation: Panel Evidence from Vietnam’. *The Journal of Development Studies*, 54(4): 641–56. https://doi.org/10.1080/00220388.2017.1342817

Borusyak, K., P. Hull, and X. Jaravel (2022). ‘Quasi-Experimental Shift–Share Research Designs’. *The Review of Economic Studies*, 89(1): 181–213. https://doi.org/10.1093/restud/rdab030

Brockmeyer, A., S. Smith, M. Hernandez, and S. Kettle (2019). ‘Casting a Wider Tax Net: Experimental Evidence from Costa Rica’. *American Economic Journal: Economic Policy*, 11(3): 55–87. https://doi.org/10.1257/pol.20160589

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

Busso, M., M.P. Gonzalez, and C. Scartascini (2022). 'On the Demand for Telemedicine: Evidence from the COVID-19 Pandemic'. *Health Economics*, 31(7): 1289–521. https://doi.org/10.1002/hec.4523

Buys, P., S. Dasgupta, T.S. Thomas, 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

Cariolle, J., and M. Le Goff (2023). 'Spatial Internet Spillovers in Manufacturing'. *The Journal of Development Studies*. https://doi.org/10.1080/00220388.2023.2204177

Casaburi, L., M. Kremer, S. Mullainathan, and R. Ramrattan (2014). 'Harnessing ICT to Increase Agricultural Production: Evidence from Kenya'. Working Paper. Cambridge, MA: Harvard University.

CIESIN (2018). 'Gridded Population of the World' Version 4 (GPWV4): Population Density, Revision 11. New York: NASA and Columbia University.

Commander, S., R. Harrison, and N. Menezes-Filho (2011). 'ICT and Productivity in Developing Countries: New Firm-Level Evidence from Brazil and India'. *Review of Economics and Statistics*, 93(2): 528–41. https://doi.org/10.1162/REST_a_00080

D'Andrea, A., and N. Limodio (2019). 'High-Speed Internet, Financial Technology and Banking in Africa'. BAFFI CAREFIN Centre Research Paper 124. Milan: Bocconi University. https://doi.org/10.2139/ssrn.3480373

De Andrade, G.H., M. Bruhn, and D. McKenzie (2014). 'A Helping Hand or the Long Arm of the Law? Experimental Evidence on What Governments Can Do To Formalize Firms'. *The World Bank Economic Review*, 30(1): 24–54. https://doi.org/10.1093/wber/lhu008

De Chaisemartin, C., and X. d'Haultfoeuille (2020). 'Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects'. *American Economic Review*, 110(9): 2964–96. https://doi.org/10.1257/aer.20181169

De Giorgi, G., M. Ploenzke, and A. Rahman (2018). 'Small Firms' Formalisation: The Stick Treatment'. *The Journal of Development Studies*, 54(6): 983–1001. https://doi.org/10.1080/00220388.2017.1327660

De Mel, S., D. McKenzie, and C. Woodruff (2009). 'Returns to Capital in Microenterprises: Evidence from a Field Experiment'. *The Quarterly Journal of Economics*, 124(1): 423.

De Mel, S., D. McKenzie, and C. Woodruff (2013). 'The Demand For, and Consequences of, Formalization among Informal Firms in Sri Lanka'. *American Economic Journal: Applied Economics*, 5(2): 122–50. https://doi.org/10.1257/app.5.2.122

D'Elvidge, C., M. Zhizhin, T. Ghosh, F.-C. Hsu, and J. Taneja (2021). 'Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019'. *Remote Sensing*, 13(5): 922. https://doi.org/10.3390/rs13050922

Demenet, A., M. Razafindrakoto, and F. Roubaud (2016). 'Do Informal Businesses Gain from Registration and How? Panel Data Evidence from Vietnam'. *World Development*, 84: 326–41. https://doi.org/10.1016/j.worlddev.2015.09.002

D'Erasmo, P.N., and H.J.M. Boedo (2012). 'Financial Structure, Informality and Development'. *Journal of Monetary Economics*, 59(3): 286–302. https://doi.org/10.1016/j.jmoneco.2012.03.003

Devas, N., and R. Kelly (2001). 'Regulation or Revenues? An Analysis of Local Business Licences, with a Case Study of the Single Business Permit Reform in Kenya'. *Public Administration and Development: The International Journal of Management Research and Practice*, 21(5): 381–91. https://doi.org/10.1002/pad.195

Dom, R., A. Custers, S. Davenport, and W. Prichard (2022). *Innovations in Tax Compliance: Building Trust, Navigating Politics, and Tailoring Reform*. Washington, DC: World Bank Publications. https://doi.org/10.1596/978-1-4648-1755-7

Fan, H., Y. Liu, N. Qian, and J. Wen (2018). 'Computerizing VAT Invoices in China'. Working Paper 24414. Cambridge, MA: NBER. https://doi.org/10.3386/w24414

Floridi, A., B.A. Demena, and N. Wagner (2020). ‘Shedding Light on the Shadows of Informality: A Meta-analysis of Formalization Interventions Targeted at Informal Firms’. *Labour Economics*, 67(C): 101925. https://doi.org/10.1016/j.labeco.2020.101925

Gallien, M., and V. van den Boogaard (2021). ‘Rethinking Formalisation: A Conceptual Critique and Research Agenda’. ICTD Working Paper 127. Brighton: Institute of Development Studies.

Goodman-Bacon, A. (2021). ‘Difference-in-Differences with Variation in Treatment Timing’. *Journal of Econometrics*, 225(2): 254–77. https://doi.org/10.1016/j.jeconom.2021.03.014

Guriev, S., N. Melnikov, and E. Zhuravskaya (2021). ‘3G Internet and Confidence in Government’. *The Quarterly Journal of Economics*, 136(4): 2533–613. https://doi.org/10.1093/qje/qjaa040

Hanlon, M., and S. Heitzman (2010). ‘A Review of Tax Research’. *Journal of Accounting and Economics*, 50(2–3): 127–78. https://doi.org/10.1016/j.jacceco.2010.09.002

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

Hjort, J., and L. Tian (2023). ‘The Economic Impact of Internet Connectivity in Developing Countries’. Working Paper. London: CEPR.

Islam, A., S. Muzi, and J.L. Rodriguez Meza (2018). ‘Does Mobile Money Use Increase Firms’ Investment? Evidence from Enterprise Surveys in Kenya, Uganda, and Tanzania’. *Small Business Economics*, 51(3): 687–708. https://doi.org/10.1007/s11187-017-9951-x

Jacolin, L., J. Keneck Massil, and A. Noah (2021). ‘Informal Sector and Mobile Financial Services in Emerging and Developing Countries: Does Financial Innovation Matter?’ *The World Economy*, 44(9): 2703–37. https://doi.org/10.1111/twec.13093

Jouste, M., M.I. Nalukwago, and R. Waiswa (2021). ‘Do Tax Administrative Interventions Targeted at Small Businesses Improve Tax Compliance and Revenue Collection? Evidence from Ugandan Administrative Tax Data’. WIDER Working Paper 17/2021. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/951-8

Kaplan, J.O., and K.H.-K. Lau (2021). ‘The WGLC Global Gridded Lightning Climatology and Timeseries (WGLC)’. *Earth System Science Data*, 13(7): 3219–37. https://doi.org/10.5194/essd-13-3219-2021

Kochanova, A., Z. Hasnain, and B. Larson (2020). ‘Does e-Government Improve Government Capacity? Evidence from Tax Compliance Costs, Tax Revenue, and Public Procurement Competitiveness’. *The World Bank Economic Review*, 34(1): 101–20. https://doi.org/10.1093/wber/lhx024

Koivisto, A., N. Musoke, D. Nakyambadde, and C. Schimanski (2021). ‘The Case of Taxing Multinational Corporations in Uganda: Do Multinational Corporations Face Lower Effective Tax Rates and Is There Evidence for Profit Shifting?’ WIDER Working Paper 51/2021. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/989-1

Konte, M., and G.K. Tetteh (2022). ‘Mobile Money, Traditional Financial Services and Firm Productivity in Africa’. *Small Business Economics*, 60: 745–69. https://doi.org/10.1007/s11187-022-00613-w

Lediga, C., N. Riedel, and K. Strohmaier (2020). ‘What You Do (and What You Don’t) Get When Expanding the Net-Evidence from Forced Taxpayer Registrations in South Africa’. In *Proceedings of the Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*.

Lee, H.C. (2016). ‘Can Electronic Tax Invoicing Improve Tax Compliance? A Case Study of the Republic of Korea’s Electronic Tax Invoicing for Value-Added Tax’. World Bank Policy Research Working Paper 7592. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-7592

Ligomeka, W. (2020). ‘Assessing the Performance of African Tax Administrations: A Malawian Puzzle’. ICTD Working Paper 15365. Brighton: Institute of Development Studies.

Lopez-Martin, B. (2019). ‘Informal Sector Misallocation’. *Macroeconomic Dynamics*, 23(8): 3065–98. https://doi.org/10.1017/S1365100517001055

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

Mascagni, G., and A. Mengistu (2016). ‘The Corporate Tax Burden in Ethiopia: Evidence from Anonymised Tax Returns’. Working Paper 13524. Brighton: Institute of Development Studies. https://doi.org/10.2139/ssrn.2776570

Mascagni, G., and A. Mengistu (2019). ‘Effective Tax Rates and Firm Size in Ethiopia’. *Development Policy Review*, 37(S2): 248–73. https://doi.org/10.1111/dpr.12400

Mascagni, G., A.T. Mengistu, and F.B. Woldeyes (2021). ‘Can ICTs Increase Tax Compliance? Evidence on Taxpayer Responses to Technological Innovation in Ethiopia’. *Journal of Economic Behavior & Organization*, 189: 172–93. https://doi.org/10.1016/j.jebo.2021.06.007

Mascagni, G., F. Santoro, D. Mukama, J. Karangwa, and N. Hakizimana (2022). ‘Active Ghosts: Nil-Filing in Rwanda’. *World Development*, 152(2002): 105806. https://doi.org/10.1016/j.worlddev.2021.105806

Mayega, J., R. Ssuuna, M. Mubajje, M. I Nalukwago, and L. Muwonge (2019). ‘How Clean Is Our Taxpayer Register? Data Management in the Uganda Revenue Authority’. Brighton: Institute of Development Studies.

McNabb, K., D. Nakyambadde, M. Jouste, and S. Kavuma (2022). ‘The Uganda Revenue Authority Firm Panel’. WIDER Technical Note 2. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/WTN/2022-2

Mensah, J.T. (2021). ‘Mobile Phones and Local Economic Development: A Global Evidence’. SSRN Working Paper 3811765. https://doi.org/10.2139/ssrn.3811765

Mensah, J.T., K. Tafere, and K.A. Abay (2022). ‘Saving Lives Through Technology: Mobile Phones and Infant Mortality’. Policy Research Working Paper 9978. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-9978

Miller, G.S., and D.J. Skinner (1998). ‘Determinants of the Valuation Allowance for Deferred Tax Assets Under SFAS No. 109’. *The Accounting Review*, 73(2): 213.

Moore, M. (2020). ‘What Is Wrong with African Tax Administration?’ ICTD Working Paper 15661. Brighton: Institute of Development Studies.

Moore, M. (2023). ‘Tax Obsessions: Taxpayer Registration and the “Informal Sector” in Sub-Saharan Africa’. *Development Policy Review*, 41(1): e12649. https://doi.org/10.1111/dpr.12649

OECD (2007). *Implementing Competition Policy in Developing Countries*. Paris: OECD.

OECD, African Union Commission, and African Tax Administration (2021). *Revenue Statistics in Africa 2021*. Paris: OECD.

Okeleke, K. (2019). ‘Uganda: Driving Inclusive Socio-Economic Progress Through Mobile-Enabled Digital Transformation’. London: GSM Association.

Okunogbe, O., and V. Pouliquen (2022). ‘Technology, Taxation, and Corruption: Evidence from the Introduction of Electronic Tax Filing’. *American Economic Journal: Economic Policy*, 14(1): 341–72. https://doi.org/10.1257/pol.20200123

Okunogbe, O., and F. Santoro (2023). ‘Increasing Tax Collection in African Countries: The Role of Information Technology’. *Journal of African Economies*, 32(1): i57–i83. https://doi.org/10.1093/jae/ejac036

Paunov, C., and V. Rollo (2016). ‘Has the Internet Fostered Inclusive Innovation in the Developing World?’ *World Development*, 78: 587–609. https://doi.org/10.1016/j.worlddev.2015.10.029

Pieterse, D., E. Gavin, and C.F. Kreuser (2018). ‘Introduction to the South African Revenue Service and National Treasury Firm-Level Panel’. *South African Journal of Economics*, 86(C): 6–39. https://doi.org/10.1111/saje.12156

Porter, G., K. Hampshire, J. Milner, A. Munthali, E. Robson, A. De Lannoy, A. Bango, N. Gunguluza, M. Mashiri, and A. Tanle (2016). ‘Mobile Phones and Education in Sub-Saharan Africa: From Youth Practice to Public Policy’. *Journal of International Development*, 28(1): 22–39. https://doi.org/10.1002/jid.3116

Rand, J., and N. Torm (2012). 'The Benefits of Formalization: Evidence from Vietnamese Manufacturing SMEs'. *World Development*, 40(5): 983–98. https://doi.org/10.1016/j.worlddev.2011.09.004

Santoro, F. (2021). 'To File or Not to File? Another Dimension of Tax Compliance: The Eswatini Taxpayers' Survey'. *Journal of Behavioral and Experimental Economics*, 95: 101760. https://doi.org/10.1016/j.socec.2021.101760

Santoro, F., A. Lees, M. Carreras, T. Mukamana, N. Hakizimana, and Y. Nsengyiumva (2023). 'Technology and Tax: Adoption and Impacts of e-Services in Rwanda'. ICTD Working Paper 153. Brighton: Institute of Development Studies.

Slemrod, J. (2019). 'Tax Compliance and Enforcement'. *Journal of Economic Literature*, 57(4): 904–54. https://doi.org/10.1257/jel.20181437

Uganda Revenue Authority (2022). 'The Uganda Revenue Authority Firm Panel Data'. Kampala: Uganda Revenue Authority.

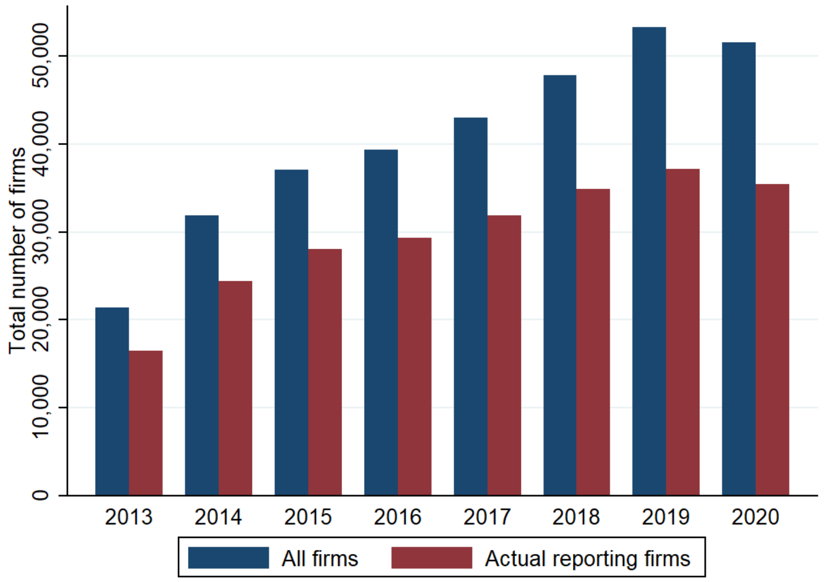
Ulyssea, G. (2020). 'Informality: Causes and Consequences for Development'. *Annual Review of Economics*, 12: 525–46. https://doi.org/10.1146/annurev-economics-082119-121914

WDI (2021). 'World Development Indicators: GDP Deflators—Uganda'. Washington, DC: World Bank.

Zuo, G.W. (2021). 'Wired and Hired: Employment Effects of Subsidized Broadband Internet for Low-Income Americans'. *American Economic Journal: Economic Policy*, 13(3): 447–82. https://doi.org/10.1257/pol.20190648

Appendix A

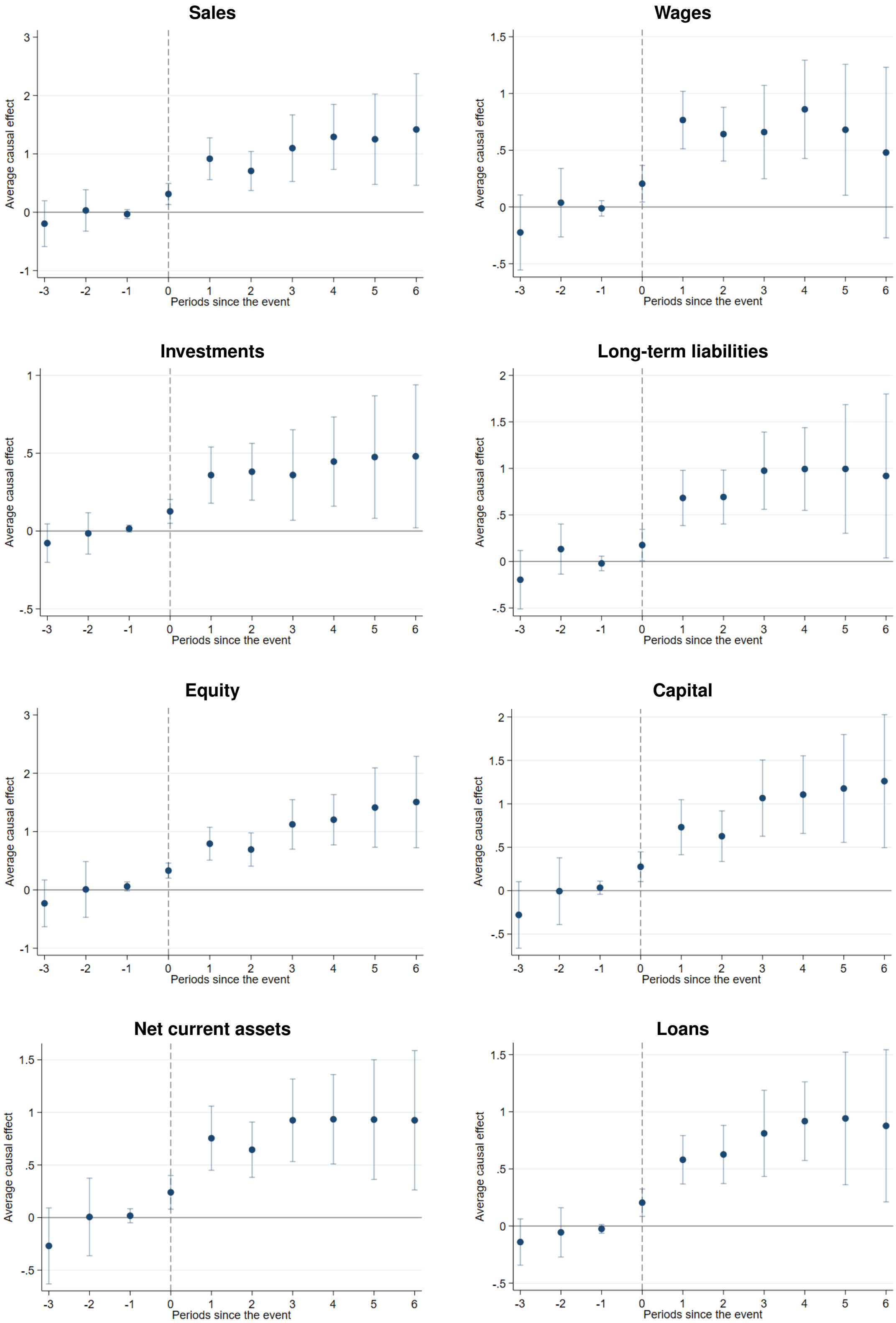
Figure A1: Number of firms filing CIT with the URA



Note: the number of firms filing CIT in 2013–20.The blue bars indicate the total number of all firms that file CIT with the URA. The red bars display the total number of firms that report positive turnover and costs of sales for at least one year over the considered time span 2013–20.

Source: authors' compilation based on URA CIT data.

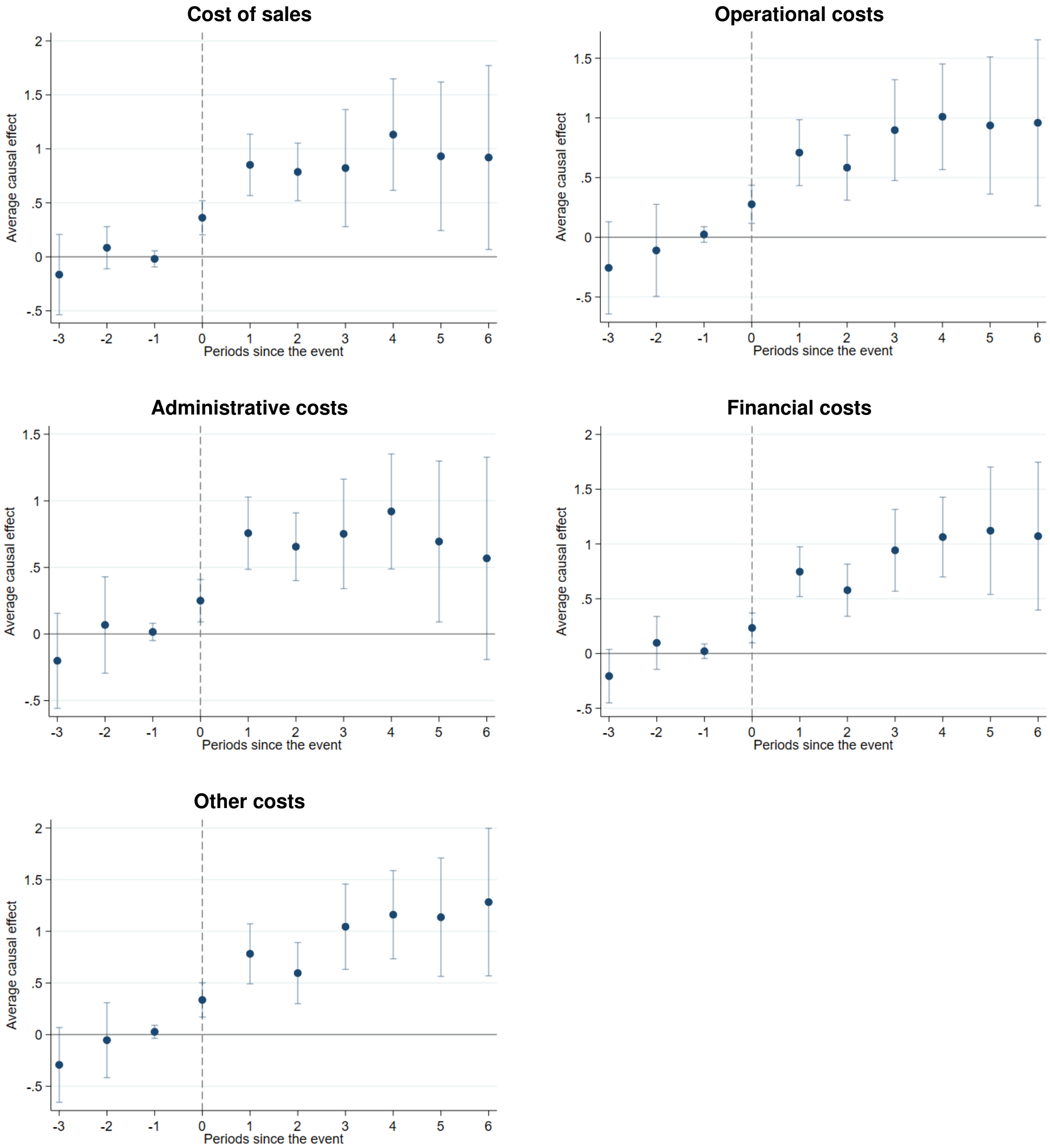
Figure A2: Event study estimates of 3G coverage on the aggregate performance of formalized firms



Note: staggered DID estimates (De Chaisemartin and d'Haultfoeuille 2020). The dependent variables are the inverse hyperbolic sine of the variables specified in the titles of the graphs; all are subcounty–sector aggregates. The control variables include the local share of 2G coverage and nighttime light intensity as well as district–year and sector–year fixed effects.

Source: authors' illustration based on URA and GSMA data.

Figure A3: Event study estimates of 3G coverage on total reported costs



Note: the figures report event study (staggered DID) estimates (De Chaisemartin and d'Haultfoeuille 2020). The dependent variables are the inverse hyperbolic sine of the variables specified in the titles of the graphs; all are subcounty–sector aggregates. The control variables include the local share of 2G coverage and nighttime light intensity as well as subcounty, district–year, and sector–year fixed effects.

Source: authors' illustration based on URA and GSMA data.

Table A1: Definitions of variables

Variable name	Notation
Firm performance	
Sales	Total sales
Wages	Total employment expenses (e.g. salaries, wages, bonus, and contribution to retirement fund)
Capital	Total fixed assets (e.g. land and buildings, plants and machinery)
Equity	Total shareholders' funds
Investments	Shares, debentures, fixed deposits, and government securities
Loans	Total loans and advances
Long-term liabilities	Secured and unsecured liabilities
Current net assets	Current assets minus current liabilities
Costs	
Cost of sales	Input usage (e.g. materials, intermediate goods, labour costs)
Operating costs	e.g. advertisement, audit expenses, commission, donations, entertainment, rent, repairs, startup costs
Administrative costs	e.g. depreciation, loss on disposal of assets, management fees, research and employment expenses
Financial costs	e.g. interest expenses, bank charges, commitment fees, insurance
Costs from insurance	Expenses attributable to short-term insurance income
Other costs	Other expenses
Schedule 1	
Deductions	Total allowable deductions
Chargeable income	Chargeable income: profits and gains from business and profession
Loss previous year	Brought forward loss of previous year from business activity to be set off against current year of income
Tax base	Taxable income equals profit before tax after adjustment for depreciation and capital allowance plus Income from capital gains minus capital losses
Loss this year	Loss to be carried forward to next year
Tax outcomes	
Profit before tax	Gross profit and other income minus all costs
Tax base total	Total taxable income added up from all tax schedules
Tax liabilities total	Sum of tax bases multiplied with respective tax rate: Schedule 1/business and profession: 30%; insurance business: 30%; mining: 25%; repatriated branch profits: 15%
Deferred tax liability	Deferred tax liability
Effective tax ratio	Total tax liabilities divided by profit before tax

Source: authors' descriptions and calculations based on URA CIT data.

Table A2: Mobile network coverage and firm formalization

	By reporting status			Actual reporting by sector		
	All (1)	Nil-reporting (2)	Actual reporting (3)	Primary (4)	Secondary (5)	Tertiary (6)
Panel A: SSIV first stage	3G Coverage					
Lightning exposure × coverage	–0.216***	–0.216***	–0.216***	–0.216***	–0.216***	–0.216***
	(0.061)	(0.061)	(0.061)	(0.062)	(0.062)	(0.062)
Nighttime lights	0.062***	0.062***	0.062***	0.062***	0.062***	0.062***
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)
2G coverage	–0.026	–0.026	–0.026	–0.026	–0.026	–0.026
	(0.027)	(0.027)	(0.027)	(0.028)	(0.028)	(0.027)
F-stat. first stage	12.68	12.68	12.68	11.87	11.87	12.66
Panel B: SSIV second stage	*asinh* Number of firms					
3G coverage	0.641***	0.462***	0.488***	0.538*	0.361	0.620***
	(0.168)	(0.145)	(0.142)	(0.303)	(0.340)	(0.161)
No. observations	145,904	145,904	145,904	13,264	13,264	112,744
No. subcounties	829	829	829	829	829	829
Controls	✓	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓	✓

Note: the SSIV first stage and SSIV regressions are reported. The dependent variable in Panel A is 3G coverage. The dependent variables in Panel B measure the inverse hyperbolic sine of the number of total reporting firms by subcounty–sector and year (column 1), the number of firms that report nil information (zero turnover and zero costs of sales for all years being in the sample, column 2), and the number of firms actually reporting to the URA (excluding nil reporters and outliers, column 3). This latter group is split by sector in columns 4–6. All specifications control for the local share of 2G coverage and nighttime light intensity, and include subcounty, district–year, and sector–year fixed effects. Standard errors, clustered at the subcounty level, are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations based on GSMA and URA CIT data.

Table A3: Mobile coverage, tax base, and its components from Schedule 1

	Full sample				
	asinh Tax base (1)	*asinh* Loss this year (2)	*asinh* Chargeable income (3)	*asinh* Loss previous year (4)	*asinh* Deductions (5)
Panel A: TWFE					
3G coverage	1.260***	1.124***	0.081	0.782***	1.297***
	(0.115)	(0.139)	(0.179)	(0.114)	(0.130)
Panel B: SSIV					
3G coverage	1.874*	1.967*	-0.540	2.004**	2.402**
	(1.034)	(1.134)	(1.518)	(0.978)	(1.190)
No. observations	145,904	145,904	145,904	145,904	145,904
No. subcounties	829	829	829	829	829
Controls	✓	✓	✓	✓	✓
Fixed effects	✓	✓	✓	✓	✓

Note: results from TWFE and the second stage of SSIV regressions are reported, relying on the sample of firms actually reporting to the URA (excluding nil reporters and outliers). Dependent variable (from Schedule 1) are the inverse hyperbolic sine of the tax base from business and profession, loss this year to be carried forward, chargeable income from profits and gains, loss previous year, and deductions. All specifications control for the share of 2G coverage and nighttime lights and include subcounty, district–year, and sector–year fixed effects. The F-statistic for the first stage of the SSIV is 12.68. Standard errors, clustered at the subcounty level, are reported in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: authors' calculations based on GSMA and URA CIT data.