ICTD

African Tax Administration Paper 35

Enhancing Taxpayer Registration with Inter-Institutional Data Sharing – Evidence from Uganda

Celeste Scarpini, Fabrizio Santoro, Ronald Waiswa, Moyo Arewa and Jane Nabuyondo

June 2024

Bill & Melinda Gates foundation

Norad

UK International Development
Partnership | Progress | Prosperity

institute of development studies

ICTD African Tax Administration Paper 35

Enhancing Taxpayer Registration with Inter-Institutional Data Sharing – Evidence from Uganda

Celeste Scarpini, Fabrizio Santoro, Ronald Waiswa, Moyo Arewa and Jane Nabuyondo

June 2024

Enhancing Taxpayer Registration With Inter-Institutional Data Sharing – Evidence From Uganda

Celeste Scarpini, Fabrizio Santoro, Ronald Waiswa, Moyo Arewa and Jane Nabuyondo

ICTD African Tax Administration Paper 35
First published by the Institute of Development Studies in June 2024

ISBN: 978-1-80470-209-3
DOI: 10.19088/ICTD.2024.047

Available from:
The International Centre for Tax and Development at the Institute of Development Studies, Brighton BN1 9RE, UK
Tel: +44 (0) 1273 606261 Fax: +44 (0) 1273 621202
E-mail: info@ictd.ac.uk
Web: ictd.ac/publication
X: twitter.com/ICTDTax
Facebook: facebook.com/ICTDTax
LinkedIn: linkedin.com/company/ICTDTax

IDS is a charitable company limited by guarantee and registered in England
Charity Registration Number 306371
Charitable Company Number 87733

Enhancing Taxpayer Registration with Inter-Institutional Data Sharing – Evidence from Uganda

Celeste Scarpini, Fabrizio Santoro, Ronald Waiswa, Moyo Arewa and Jane Nabuyondo

Summary

Comprehensive population data is often lacking in many developing countries, especially in Africa. This is a critical challenge for tax administrations, who are already grappling with a substantial hidden informal economy. Recent studies highlight the importance of national identification (ID) data for enhancing tax collection efforts. This study looks into the impact of inter-institutional collaboration to share national ID data on tax administration data quality and functions. The Uganda Revenue Authority (URA) has integrated its registration system with that of the National Identification and Registration Authority (NIRA) and Uganda Registration Services Bureau (URSB), which allows it to access ID data for individuals and businesses. The Instant Tax Identification Number (Instant TIN) – an interface pulling this third-party data into the taxpayer registration form – promises a swifter registration process for taxpayers, and better data on taxpayers.

Our mixed methods approach combines qualitative data collection from in-depth interviews with government officials from various agencies, with quantitative analysis of URA administrative data. We present three sets of results. First, Instant TIN registrations show significant differences from the normal registration process. Most Instant TIN registrations are for individual taxpayers, and they are more likely to be female, younger, and to have been informal before registration. This highlights the facilitation function of Instant TIN, bringing in generally more marginalised categories. Second, our analysis highlights Instant TIN's discernible impact on data quality. It significantly reduces duplicate TINs for individual taxpayers, and markedly improves the validity of email addresses – with a notable 8 percentage point reduction in invalid, missing, or duplicate emails. In addition, the proportion of invalid phone numbers among Instant TIN registrations decreases by a substantial 6 percentage points. However, there is a worrying trend with accuracy of sector information, with Instant TIN registrations showing a 12 percentage point increase in the likelihood of invalid or missing sector data. Third, on one hand, Instant TIN streamlines processes, reduces duplication and manual data entry, and saves resources. On the other, we highlight several challenges, including infrequent updates to external datasets, and the absence of

data validation within the Instant TIN interface. These issues increase administrative costs, and hinder URA's ability to effectively engage, educate, and enforce tax laws. Mandatory in-person updates for specific taxpayers and invalid contact details further add to compliance costs.

Keywords: digital ID system; tax administration; formalisation.

Celeste Scarpini is a Research Officer at the International Centre for Tax and Development (ICTD), and a PhD student at the Department of Economics, University of Sussex. Her main research interests relate to tax administration in sub-Saharan Africa, from technology adoption to data management and revenue collection strategies.

Fabrizio Santoro is a Research Fellow at the Institute of Development Studies, and the Research Lead for the second component of the International Centre for Tax and Development (ICTD)'s DIGITAX Research Programme. His main research interests relate to governance, public finance, and taxation, with a strong focus on impact evaluation methodologies and statistical analysis. He holds a PhD in Economics from the University of Sussex.

Ronald Waiswa was formerly Research and Policy Analysis Supervisor at the Uganda Revenue Authority, and is now Specialist in Applied Research and Statistics at the African Tax Administration Forum (ATAF).

Moya Arewa is the Programme Director for the Local Government Revenue Initiative (LoGRI). He was previously the Manager for Strategic Initiatives at the International Centre for Tax and Development (ICTD), and, before then, a Policy Development Officer at the City of Toronto. His tax research has focused on understanding how new technologies impact tax policy, administration, and public service delivery.

Jane Nabuyondo is a supervisor for Business Intelligence with 15 years of experience in tax administration and vast knowledge of government (public) operations. She is also the team lead in development and maintenance of the URA Data Warehouse/Business Intelligence solution, including overseeing the implementation of data governance initiatives in URA that will support the organisation to become a data driven entity.

Contents

Figures

Tables

Acknowledgements

We would like to thank the Uganda Revenue Authority for the support provided for this study. We are particularly grateful to Allen Nassanga and the senior management team for allowing us to perform the study. We would like to thank all anonymous interviewees within URA, URSB, KCCA, NIRA, and NITA-U for agreeing to participate in the interview data collection. We are also grateful to two anonymous referees, whose comments greatly helped to improve the paper, and participants in the Oxford CSAE 2024 conference, where this paper was presented.

Acronyms

API	Application Programming Interface
BRN	Business registration number
GDP	Gross domestic product
ID	Identification
ITAS	Integrated Tax Administration System
KCCA	Kampala Capital City Authority
LIC	Lower-income country
LMIC	Lower- and middle-income country
NIN	National identity number
NIRA	National Identification and Registration Authority
NITA-U	National Information Technology Authority - Uganda
NSSF	National Social Security Fund
OECD	Organisation for Economic Co-operation and Development
PAYE	Pay-As-You-Earn
PSM	Propensity score matching
TADAT	Tax Administration Diagnostic Assessment Tool
TIN	Taxpayer identification number
TREP	Taxpayer Register Expansion Program
UBOS	Uganda Bureau of Statistics
URA	Uganda Revenue Authority
URSB	Uganda Registration Services Bureau
VAT	Value added tax

1. Introduction

Governments in lower-income countries (LICs) struggle to collect and use information about their population adequately. Tax administrations, in particular, are information-intensive, and the availability of taxpayer information is of paramount importance for effective tax collection (Jensen 2019; Kleven *et al.* 2011; Kleven, Kriener and Saez 2016; Naritomi 2019; Pomeranz 2015). The lack of systematic information on the tax base is a crucial challenge in many African countries – data is either not available or, if present, inaccurate (Mayega *et al.* 2021).[1] At the same time, African tax administrations operate in a context of high informality, where accurate information on taxable entities is scarce (Besley and Persson 2014). For this reason, one of their primary objectives is to register a large proportion of the population for tax purposes, and the accuracy of data collected is often a secondary concern (Moore 2022).

In this context, cross-agency data sharing agreements and system integration between institutions, especially for national ID schemes, can significantly strengthen African tax administrations (Santoro *et al.* 2022). Tapping into third-party ID data can help tax administrations in many ways. First, connecting every taxpayer to a unique, foundational national ID can improve the revenue authorities' registration function, and provide more accurate data about who taxpayers actually are. By leveraging ID data, tax administrations can unambiguously identify each taxpayer. Similarly, access to business registers would improve the ability of tax administrations to link businesses to their owners and employees, enabling a more straightforward assessment of tax liability. Second, cross-institutional integrated systems, with a single point of registration across agencies, can reduce compliance costs for taxpayers, sparing them the burden of visiting the tax office and other registration bureaux, interacting with several officials, and filling and submitting multiple paper forms. In addition, good quality, updated data on taxpayers' contact details and business location allows tax administrations to reach them more easily for compliance and sensitisation purposes. Third, ID data and integrated systems can improve the identification of potential evaders who exploit loopholes in registration systems (such as multiple tax identifiers), strengthening tax enforcement. Finally, reliance on integrated systems can improve the management and efficiency of revenue administrations. Better quality data would allow a shift towards a more data-centric approach, and greater reliance on metrics and indicators (e.g. in performance targeting, measurement, and task restructuring). Moreover, the personnel freed up from in-person registration activities can be employed in taxpayer services, an often-

[1] Data from taxpayer registers becomes inefficient when often prioritised low-hanging fruits, like mass registration campaigns, clog registers with inactive taxpayers (Moore 2022).

neglected tax administration area where automation can be less helpful.

All the above benefits depend on essential assumptions that are often not met. First, that the data necessary for tax purposes is available in other government agencies. This is often not the case – for example, with real estate ownership and land registration. Useful information – especially for wealthy individuals – lies outside government and with commercial banks, many of which resist sharing it with tax administrations backed by stringent regulations. Second, that data from third-party institutions, when available, is accurate and captured in a way that fits the needs of the tax administration. In contexts of limited coordination across institutions, government agencies often collect data to fit their needs – not considering what would be useful for tax purposes. Third, that tax administrations have adequate capacity and skills to analyse it and put it to its best use once the data is shared. This is often not the case, as resource-constrained revenue authorities often lack the capacity to systematically analyse the data they hold, which complicates efforts to develop a more data-driven approach to tax administration. Further, that tax administrations are adequately equipped with sophisticated storage and security systems to safely manage data, respecting data privacy regulations.

The Uganda case study is emblematic in this sense. As part of a broader plan to integrate the Uganda Revenue Authority's (URA) systems with many other agencies' platforms, URA recently shifted towards accessing third-party data to identify taxpayers. In July 2021, URA launched a data sharing strategy to integrate data from the national ID system managed by the National Identification and Registration Authority (NIRA) with their register of individual taxpayers. A similar system integration, halted in 2023 by URA to review its implementation, took place with the Uganda Registration Services Bureau (URSB) to obtain business information. From the taxpayer's perspective, this integration resulted in an online, much faster registration process called Instant TIN, where taxpayers enter their national identification number (for individuals) or business registration number (for registered businesses) on the online registration platform. Instantly, data is automatically retrieved from the NIRA and URSB systems, and a unique taxpayer identification number (TIN) is generated. This system integration strategy comes in a context where there are gaps and inaccuracies in URA's taxpayer register (Mayega *et al.* 2021), which prevent unambiguous identification of taxpayers. This has important repercussions on the capacity to both enforce and facilitate compliance, if taxpayers cannot be easily located and contacted (Okunogbe and Santoro 2022). Hence, relying on third-party data at registration was thought to significantly improve the quality of the URA taxpayer register.

Against this background, our study addresses three interrelated questions:

i. How the system integration between URA and other public institutions' databases is taking place, and what are the challenges in the process?;

ii. Does data sharing and registration on the Instant TIN platform improve data quality within URA databases?;

iii. Is the new data from third-party sources helpful in strengthening the capacity of URA to identify taxpayers, enforce and facilitate compliance, and improve overall administrative efficiency?

To address these questions, we followed a mixed methods approach, combining qualitative and quantitative data analysis. First, in collaboration with URA and other government institutions, including NIRA, URSB, Kampala Capital City Authority (KCCA), and National Information Technology Authority (NITA), we carried out 17 in-person, in-depth interviews with government officials to gain a comprehensive understanding of the integration in place, and to gauge the different perspectives of the institutions involved. Second, we ran a more quantitative analysis of URA administrative data. In particular, we explored registration patterns and key correlates with Instant TIN registration using the taxpayer register. We used this data to capture gaps in quality and inaccuracies, trying to understand the role of the new registration solution in limiting them.

The evidence we present is mixed. On one hand, the impact of Instant TIN technology on registrations has been remarkable. In 2022 alone, it accounted for 35 per cent of all registrations. The technology's registration outcomes are closely linked to the Taxpayer Register Expansion Program (TREP), a mass registration campaign of URA. Instant TIN is a crucial tool for TREP field officers to streamline registration when operating door-to-door or at a one-stop shop. We find that 39 per cent of Instant TIN registrations are attributable to in-the-field TREP. Notably, the technology is predominantly utilised by previously informal businesses, particularly those that are young, and, to a lesser extent, owned by women.

However, the picture looks more complex when considering the effect on data quality and URA functions. Duplication of identities is, by design, largely removed. Also, while using Instant TIN reduces inaccuracies in email addresses and phone numbers, it is ineffective in enhancing company data quality. In-depth interviews suggest that these persisting data flaws pose a challenge for URA. The register maintenance team makes significant efforts to rectify inaccuracies caused by using Instant TIN in TREP. Additionally, the poor quality of contact and location details undermines URA's capacity to bring new taxpayers on board and to enforce tax laws, muting the benefits of a strengthened identification function.

This study is particularly relevant for various strands of research on tax administration in LICs. First, it directly contributes to the ongoing debate on the relevance of third-party data to improve tax collection. While evidence on third-party data in high- and middle-income countries is abundant (Fan *et al.* 2018; Mittal and Mahajan 2017; OECD 2017; Slemrod *et al.* 2017; Brockmeyer *et al.* 2019; Carrillo, Pomeranz and Singhal 2017), very little is known about its potential in lower-income contexts, and even less so in Africa. On one hand,

promising evidence from Chile and Brazil shows that tax administrations can effectively exploit third-party data to increase revenue (Naritomi 2019; Pomeranz 2015). Likewise, Brockmeyer *et al.* (2019) show that credible enforcement emails reporting information from third parties increased tax payment among previously non-compliant firms in Costa Rica. Success stories also come from Pakistan and China, thanks to value added tax (VAT) refund validation mechanisms using third-party data (Fan *et al.* 2018; Waseem 2022). Similar positive evidence comes from India (Mittal and Mahajan 2017), where third-party reporting on transactions increased the compliance of the top 1 per cent of firms. On the other hand, more mixed results come from Africa. In South Africa, Lediga, Riedel and Strohmaier (2020) evaluate the synchronisation of the tax register and business register. Similar to Uganda, the authority's strategy was to encourage registration, and produce a cleaner and more comprehensive taxpayer register. Despite the synchronisation sizeably increasing the tax net, revenue did not increase due to poor compliance of newly registered firms, a concern highly likely to apply in Uganda.

Second, this study speaks to the increasing literature around formalisation, specifically registration for tax purposes. When it comes to Africa, recent studies focus on government programmes aiming to include informal entities in the tax net, and, at least potentially, increase tax revenue (Gallien, Moore and van den Boogard 2021; Jouste *et al.* 2021; Lediga *et al.* 2020; Moore 2022). Evidence from this strand of work challenges the idea that expanding the tax base to the informal sector, usually micro- and subsistence-level businesses, generates substantial revenue gains (Benhassine *et al.* 2018; Moore 2022). It suggests that it could eventually hamper the reliability of tax return data, clogging it up with a large number of nil submissions (Mascagni *et al.* 2022).[2] Moreover, registration interventions can have critical unintended consequences, causing confusion among taxpayers, and affecting their perception of the tax system and compliance. As Moore (2022) argues, and as we corroborate in this study, political motivations and policy targets, rather than their practical promise, often tend to drive this ‘registration obsession’. In cases where policy targets prevail over implementation, technology can become a potentially harmful tool. This study contributes to this strand of the literature by exploring the impact of system integration between government entities for registration purposes – an increasingly popular tax registration strategy in Africa.[3]

Finally, we directly connect with increasing evidence around technology in tax

[2] Furthermore, the firm-level benefits to formalising, in terms of profits or investment, are often limited (Bruhn and McKenzie 2014; Ulyssea 2020).

[3] Many other, probably less expensive, policies proved ineffective in raising formalisation rates, especially interventions that provided information, reduced registration costs, or simplified regulation (e.g. Bruhn 2011; de Mel, McKenzie and Woodruff 2013; De Giorgi and Rahman 2013; de Andrade, Bruhn and McKenzie 2014; Rocha, Rachter and Ulyssea 2014).

administration, as reviewed in Okunogbe and Santoro (2022). As the Organisation for Economic Co-operation and Development (OECD) argues, echoed by donors and international organisations, the digitalisation of tax administration could be the most powerful tool for shifting light on the shadow economy (OECD 2017). While most of the literature looking at developing countries has focused on electronic fiscal devices (Ali et al. 2021; Mascagni *et al.* 2022; Hakizimana and Santoro 2023) and electronic filing and payment of taxes (Okunogbe and Pouliquen 2018), we focus on a novel technology boosting registration through the integration of the tax register with national ID information. In this sense, we connect to recent work around IT-enabled solutions to register taxpayers, especially regarding property taxes (Okunogbe 2021). We contribute to this literature by adding evidence around the benefits of data integration and technology for tax administration. We note their high dependence on context, and susceptibility to challenges in the implementation process, design, take-up, and taxpayer behaviour.

This study provides evidence to Ugandan and African policymakers of the benefits and challenges for tax administrations from sharing data across agencies. We acknowledge the ongoing difficulty of merging other institutional datasets with URA's register, and aim to extract valuable lessons from this case for future data sharing initiatives. This evidence is urgent, because several tax agencies are either exploring or implementing similar integration initiatives (Ethiopia, Malawi, Nigeria, and Ghana). In the paper's conclusion, we also suggest practical policy recommendations for URA.

In what follows, we describe the system integration and Instant TIN technology in Section 2, and present the research design in Section 3. Section 4 discusses our key results, and Section 5 concludes.

2. Context

2.1 The Ugandan context

Uganda is a lower-income country in East Africa, which shares the challenge of raising domestic revenue with other LICs (Besley and Persson 2013). These are reflected in its tax-to-GDP ratio of around 15 per cent, and are generally attributed to widespread evasion and informality. The informal sector in Uganda represented about 40 per cent of GDP on average from 2004 to 2015 (Medina and Schneider 2019). Approximately 87 per cent of total employment is informal (UBOS 2019). To expand the tax net, URA has invested heavily on formalisation strategies and the expansion of the tax base to foster compliance (Moore 2022). It is currently doing this by relying on third-party data.

Recent descriptive research shows several gaps and inaccuracies in URA data – for instance, duplicates are quite common. As of mid-2018, Mayega *et al.* (2019) show that as many as 44 per cent of all registered taxpayers have contact details that are identical to those of at least one other registered taxpayer – this translates into more than half a million taxpayers. This was probably because tax agents provided their details, rather than the taxpayers', when registering different clients. This practice might arise because of the requirement to provide an email address to register, which many taxpayers might not have, and because tax agents want control over their clients' filing process to demand more frequent payment (Mayega *et al.* 2019). Duplicates also refer to national identification numbers (16,000) and passport numbers (6,200). The authors also show that about 30,000 individuals possess more than one TIN.

Gaps in TINs have severe repercussions on the quality of tax return data (Mayega *et al.* 2021). In 2014-2018, 85 per cent of tenants' TINs declared by landlords in rental income tax returns were either missing or incorrect, and information on tenants (33 per cent) and property addresses (77 per cent) was inaccurate. Similarly, half of supplier TINs in withholding tax returns in 2020 were wrong, while most employees declared in Pay-As-You-Earn (PAYE) returns have incorrect identification details. The same inconsistencies plague income tax return data. About a third of the TINs of directors of companies are invalid. This finding is alarming, given that directors of companies are also liable to remit individual income taxes, and URA inevitably struggles to link individual directors to their companies (Santoro and Waiswa 2022)

More recently, URA has been at the centre of a broader government effort to promote inter-agency data sharing. It now benefits, with different degrees of consistency, from receiving data from the National Identification and Registration Authority (NIRA), the Uganda Registration Services Bureau (URSB), the Kampala

Capital City Authority (KCCA), the National Social Security Fund (NSSF), and the Uganda Bureau of Statistics (UBOS), among other governmental entities.[4]

NIRA and URSB, respectively, provide the revenue administration with information about individuals and businesses through an Application Programming Interface (API). The API allows the transmission of taxpayer data from NIRA and URSB to URA. Currently, the application does not allow other potentially valuable functions – such as real time data exchange between institutions, the possibility to change taxpayer information if it is updated at any point, or notification of updates in any of the registers. Data transfer takes place each time a new application on the Instant TIN portal triggers a request for information from URA to either NIRA or URSB, following a 'pulling' mechanism.[5] Despite future government plans to fully integrate systems between these institutions, the data is still shared in one direction – from NIRA and URSB to URA.[6]

2.2 Instant TIN technology

In line with the wave of digitisation efforts across African tax administrations, after implementing a robust Integrated Tax Administration System (ITAS) (Occhiali, Akol and Kargbo 2022), URA rolled out Instant TIN – an innovative online registration system that relies on transfer of data from NIRA and URSB. Instant TIN allows individuals and businesses to register for income tax with URA. Unlike the normal registration process, Instant TIN provides taxpayers with a TIN as soon as they register, which can be used immediately for most tax-related activities. Previously, taxpayers had to wait two working days to obtain a TIN.

With data from NIRA and URSB, rapid registration by Instant TIN is possible in its current iteration. Online individual registration through Instant TIN is only possible by submitting a national identity number (NIN) and date of birth at the beginning of the online application. The system then retrieves taxpayer data from NIRA, and automatically populates the name, date of birth, and where the taxpayer was

[4] Data sharing between URA and NIRA is enabled by two laws underpinning the establishment of NIRA in 2015 - the Registration of Persons Act (2015), and Data Protection and Privacy Act (2019). Both Acts mandate NIRA to share data with government entities involved in revenue generation, most notably URA.

[5] Data sharing between URA and other governmental institutions is not at the same level of sophistication as occurs with NIRA and URSB. For example, at the time this paper was written KCCA was sharing data with URA manually every day - sending data on an Excel spreadsheet. The matching of data was also done through Excel. On the other hand, KCCA, which is responsible, among other areas, for registration of trading licences, and administration and collection of local taxes and property tax in Kampala, already uses an API to validate TINs from the URA register.

[6] At the time our interviews were carried out, URA, NIRA, and URSB were involved in a plan to integrate their systems allowing the unique identification of citizens. According to this plan, in future the TIN will be replaced by the national identify number (NIN), and information on citizens and businesses would be shared and updated in real time between the institutions.

born. A similar process occurs when registering a business. The taxpayer provides their business registration number (BRN), which triggers the release of business information from URSB to URA – including business name, registration date, and entity type. The process for business registration was halted by URA in 2023, as it is currently revising and improving its implementation.

Instant TIN could benefit both taxpayers and the tax administration. First, it automates income tax and VAT registration for individuals and businesses, saving them time and money. Before the rollout of Instant TIN, a citizen wanting to register for income tax needed to complete an application form online, and provide information about their identity and details of their economic activity. Similarly, VAT registration required proof of business location, with a URA official visiting the taxpayer's premises. Once URA officers manually checked the information, the taxpayer received their TIN in two working days. Thanks to Instant TIN technology, registering individuals and businesses for these tax heads is automatic and done online.

Second, Instant TIN is widely recognised as a significant improvement to URA's internal processes.[7] It is expected to massively reduce the time and resources spent on registration tasks, and, most importantly, avoid duplicate identities in the taxpayer register. While this feature of Instant TIN technology is undoubtedly an improvement, the broader effects of the technology on other aspects of tax administration (e.g. data quality and broader administrative costs) are not clear.

Notably, Instant TIN only allows taxpayers to obtain a TIN. This makes some activities (like presumptive income tax payments, import and export of goods, and motor vehicle transfer of ownership) immediately available to taxpayers. Others (like VAT registration and filing returns) need the new taxpayer to visit URA to amend and provide further information. These steps are critical in the taxpayers' compliance journey – although they are not always taken, as we discuss more in-depth in Section 4.3.

2.3 Taxpayer Register Expansion Program (TREP)

The implementation of Instant TIN comes in the context of a major effort to encourage citizens and small businesses to register for tax purposes. An important pillar of this process is the Taxpayer Register Expansion Program (TREP), an intervention initially launched in 2013 by the Government of Uganda, together with URA, URSB, KCCA, and Ministry of Local Government. The URA-NIRA integration initiative effectively came out of TREP at a later stage, when system integration started to be considered the primary tool of taxpayer registration (URSB 2018). While currently registration with the agencies involved

[7] Interviewee URA14: ‘Instant TIN is the way to go, because you save a lot of time with paperwork’ (URA14, Kampala, December 2022).

in TREP entails each isolated institution collecting the same information from resident-clients – which is inefficient from both an end user and administrative standpoint – the agencies involved in TREP are working towards streamlining registration and identification, allowing a single touchpoint.

TREP has multiple objectives – registering businesses, educating them about the tax system, reducing compliance costs, and matching the information available to government agencies collecting revenue. Various methods have been used to enhance taxpayers' registration. The only method from TREP's launch in 2013 until 2015 was door-to-door visits, together with one-stop shops where business owners could register their activity with the different authorities, and receive assistance in filing and paying taxes.[8] From 2015, TREP started awareness campaigns, like sensitisation workshops, radio advertisements, and newspaper articles.[9] Door-to-door visits and one-stop shops still continue. As discussed more in-depth in Section 4.1, evidence shows that many previously informal businesses registering with Instant TIN did so through TREP (44 per cent). As discussed in Section 3, we do not know whether field officers made taxpayers register in TREP, or if taxpayers voluntarily sought assistance at one-stop shops.

Since the rollout of Instant TIN, TREP has been leveraging the technology to make the registration process quicker and easier in the field. Registration officers involved in door-to-door campaigns and one-stop shops use Instant TIN to register new taxpayers if they have a NIN or BRN. TREP in-person registrations are carried out by officials temporarily contracted by URA, who are paid based on the number of registrations performed. Anecdotal evidence highlights that these registrations are at least partly inaccurate. There is an incentive to register as many entities as possible, and the temporary officials may have less commitment to URA's mission and standards.[10]

Since the end of 2021, URA has embarked on a new – automated and cheaper – way of registering large numbers of unregistered taxpayers as part of TREP. URA is using third-party data to spot and register business activities that are not yet in the tax net. For this, URA combines different types of datasets, such as those from the NSSF and Land Registry. Appendix Table A1.1 gives the names of registers that URA is using for this stage of TREP. Once an economic activity has been identified, URA registers the new taxpayers by relying on third-party contact

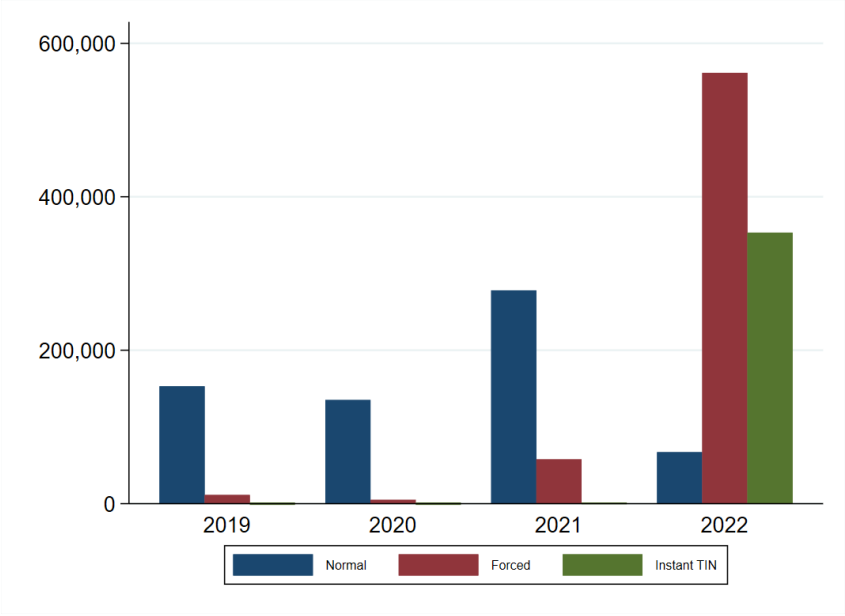
[8] One-stop shops are service centres with at least one officer from various government agencies, including URA, URSB, KCCA, and local government. One-stop shops are typically in municipal and district government offices, but can be located at URA's offices. These service centres are equipped with service desks, computers, and barcode readers to facilitate registration processes (Jouste *et al.* 2021).

[9] TREP awareness campaigns also included stakeholder engagement, tax clinics and workshops, public notices, public meetings, radio talk shows, and press briefings (Jouste *et al.* 2021).

[10] This is suggested by a large proportion of in-person registrations with Instant TIN having invalid email addresses, easily recognisable from the standardised '@trep.it' domain.

information, mainly telephone numbers. If the contact information allows, only after that URA sends an SMS to tell the taxpayer about their new registration. During this phase of TREP, hundreds of thousands of taxpayers have been registered using third-party data. Most registrations in 2022 are of this kind, and the number of normal registrations from taxpayers voluntarily enrolling in the tax system are now the minority. Figure 2.1 shows the trend of registrations over time by modality. The blue line shows what we call normal registrations, where individuals visit URA to go through registration procedures, identification, validation of documents, manual filling of registration forms and provision of biographic data and information of economic activity. The red line shows forced registrations, and includes those automatically registered by URA using third-party data. The green line includes those registered through Instant TIN, whether by themselves or with assistance through TREP.

Figure 2.1 Trend in registrations over time



Source: Own calculations on URA administrative data.
Note: Normal registration is where individuals visit URA to go through registration procedures, identification, validation of documents, manual filling of registration forms, and provision of biographic data and information of economic activity. Forced registration includes those automatically registered by URA using third-party data. Instant TIN includes those registered through Instant TIN, whether by themselves or with assistance through TREP.

3. Data and methodology

3.1 Data sources

We make use of both qualitative and quantitative data sources. On one hand, we refer to qualitative data to understand the context, describe practices and challenges, and provide essential details to our arguments grounded in the local reality under study. On the other, we use administrative data to produce more quantitative results on the drivers of registration, and the potential impact on data quality and accuracy. We collected qualitative data from 17 in-depth interviews with government officials, conducted in-person in Kampala in the second half of 2021 and 2022. Interviewees were mainly URA tax officials from different departments, such as domestic taxes, process management, IT, the TREP unit, and taxpayer register cleaning team. We also spoke to government officials outside URA, including at NIRA, URSB, KCCA, and NITA-U. Interviewing both parties to the data sharing agreement was crucial for gaining a more comprehensive understanding of the integration, and to gauge actors' perspectives. On average interviews lasted one hour. Appendix Table A1.2 summarises the in-depth interviews.

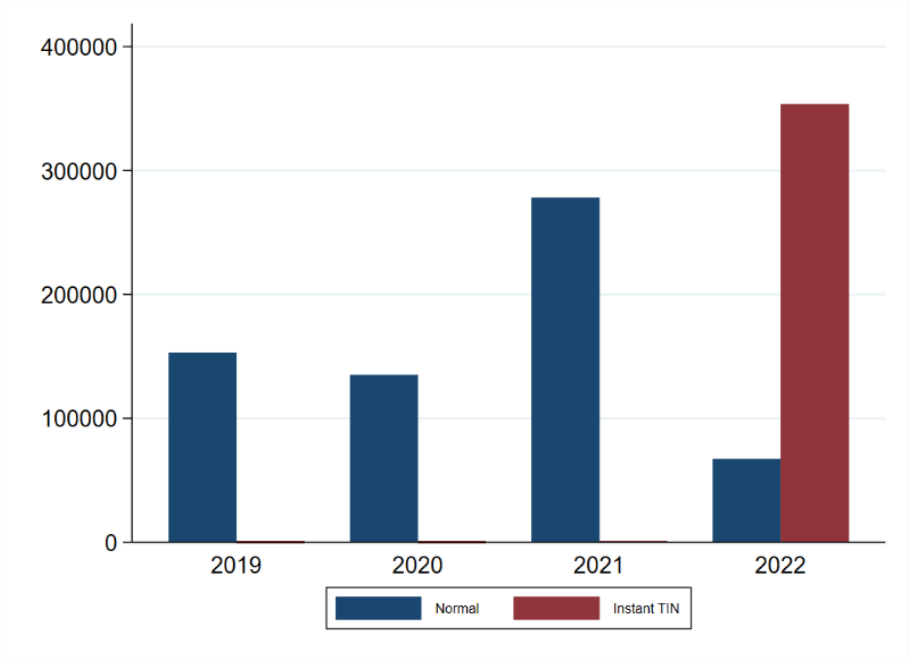
Second, we have access to a wealth of administrative data from URA. We use the whole taxpayer register, including all registrations for tax purposes, as at 6 December 2022. The register also contains information on whether the registration occurred through Instant TIN, the normal process, or forced registration. For ease of analysis, we first restrict our sample of registration to the years 2019-2022, as most registrations took place in this period. There has been a constant rise in registration numbers –around 1 million taxpayers were registered in four years. We restrict our sample to taxpayers who registered through the normal process or Instant TIN, and exclude the automated forced registrations in the latest TREP development. The vast majority of taxpayers in the dataset are individuals (92 per cent), and the rest are companies. While most of the patterns we produce refer to individual taxpayers, we also consider companies – for the sake of completeness, and to highlight potential differences in findings across the two categories.

Figure 3.1a below shows the pattern of registrations, clearly indicating the dramatic fall in normal registrations in 2022, when the new technology was launched and Instant TIN registrations became the majority. About 350,000 taxpayers registered through Instant TIN. This is 85 per cent of registrations in 2022, and more than a third of total registrations in the previous four years, excluding forced registrations. Of all registrations in 2022, 35 per cent were done through Instant TIN. Of these, 39 per cent were part of the in-the-field TREP mass registration effort. There were more Instant TIN registrations in 2022, as shown in Figure 3.1a. Almost all Instant TIN registrations are from individuals (99

per cent), as shown in Figure 3.1b. Instant TIN for companies only started in late 2022, before being halted in 2023, which explains the lower numbers (Figure 3.1b). The analysis below shows results by disaggregating along this dimension. Apart from that, additional detailed information at the taxpayer level is available in the register, from biographic details (name, gender, date, and place of birth) to contact information (phone number and email address), as well as tax-related features, such as registration date, tax type, sector of activity, and tax centre. We use this information in the analytical framework described below. Importantly, and specific to the interaction between TREP and Instant TIN, the register does not show whether the registration is performed by TREP officers in field campaign activities, or voluntarily by taxpayers seeking registration support at TREP one-stop shops.

Figure 3.1 Trend in registrations by year

(a) Aggregated



(b) Disaggregated by taxpayer type

Source: Own calculations on URA administrative data.

3.2 Methodology

We use different methodological approaches based on the data we analyse. First, we applied thematic analysis to the qualitative data from the interviews. We used a specific text analysis software (Nvivo) to code interviews. Given the semi-structured nature of our interviews, we used open coding – we did not have pre-set codes, but developed and modified them during the coding process. The critical dimensions that emerged from the evidence relate to the nature of data sharing between institutions, the impact of Instant TIN technology on both taxpayers' experience and URA's functions and processes, and recommendations and perceptions. Appendix Table A1.3 summarises the thematic analysis.

Secondly, we produce descriptive evidence on uptake of Instant TIN registration and potential repercussions on data quality. We describe the selection into Instant TIN registration with a simple Ordinary Least Squares (OLS) framework, performed on the subpopulations of individual and business taxpayers:

$$Y_i = f(\beta' \, Individ; \, \gamma' \, Business \, ; \, \delta' \, Tax; \, \nu) \qquad (1)$$

Y is the outcome of interest – Instant TIN registration. $Individ$ is a vector of taxpayer-level information derived from the register, including gender, age, and place of birth. By construction, these factors are only available for our estimation on individual taxpayers. $Business$ refers to business-level information – the type of economic activity. Relatedly, we can identify whether the taxpayer was already registered as an official business with other government institutions, such as KCCA for individuals and URSB for companies, or whether they come from the informal sector, a dimension we explore more in-depth in the analysis. Tax

indicates a set of variables referring to the tax profile of the taxpayer – the tax centre to which the taxpayer is assigned at registration, registration date, and whether the taxpayer has been registered under TREP.

As a second exercise, we set up another OLS framework, described below, to understand the correlation between Instant TIN registration, now an explanatory factor, and data quality and accuracy:

$$Y_i = f(\alpha' Instant;\ \beta' Individ;\ \gamma' Business\ ;\ \delta' Tax;\ \nu) \quad (2)$$

In this case, Y is a set of indicators for data quality, namely: (i) a dummy for multiple TINs, indicating if the same taxpayer holds more than one TIN; (ii) a dummy for invalid (including duplicate and missing) email address; (iii) a dummy for invalid (including duplicate and missing) phone number; (iv) a dummy for excessively large or small age (or outliers, measured as larger than twice the normalised difference between age and the overall mean); and (v) a dummy for missing sector of activity.[11] Outcomes (i), (iv) and (v) are only available for individuals, and not for companies. With (i), we consider duplicate entities by looking at the taxpayer's name and taxpayer's date of birth, which can only be extracted for individual taxpayers. For the same reason, the age variable in (iv) is only given for this taxpayer category. For outcome (v), gaps in sector information occur only with individuals, while the sector is specified and present for all companies, regardless of their registration mode. When computing gaps in sector information, we drop taxpayers registered for taxes on employment (PAYE), as this information is not collected for that tax type. Finally, $Instant$ is our key explanatory factor, taking a value of 1 for Instant TIN registration and 0 for normal registrations. The other variables are the same as in (1), and are used here as control variables to make estimations more precise.

We first match Instant TIN and normal registrations with a Kernel-based propensity score matching (PSM) to strengthen our analysis, and make the two groups more comparable. While we are not attempting to claim any causal effect from Instant TIN registrations on data quality, we are prudent in enhancing the similarity across normal and Instant TIN registrations. Self-selection into Instant TIN registration may introduce bias in our estimate – taxpayers choosing Instant TIN registration may be more or less prone to provide accurate data to URA – which we try to address with PSM. At the same time, it is true that taxpayers self-selecting into Instant TIN may do so for the ease and timeliness of the process, and not for in-built data requirements the registration comes with – which directly affects data quality. Thanks to Kernel PSM, we match every observation in the Instant TIN registration group with a weighted average of units from the normal

[11] We decided to pool together duplicates and invalid/missing contact details, given the very low relevance of the latter. Only 1.2 per cent of email addresses are invalid or missing. For phone numbers, 8 per cent are invalid and 3 per cent are missing. Most inaccuracy in contact details comes from duplication – a point we discuss more in detail in Section 4.2.

registration group, based on a propensity score from a logistic regression of the indicator variable for registering with Instant TIN as the outcome.[12] After matching, we run regression 2 on our outcomes of interest.

When it comes to assessing the quality of the matching, Appendix Figure A2.2 shows that PSM is successful in reducing imbalance in the different covariates after the match. For the two main assumptions PSM builds upon – unconfoundedness and overlap in log odds – while the former cannot be directly tested, it can be respected by including all the potential confounding variables available in our dataset. The latter requires that the propensity score distributions of Instant TIN and normal registrations overlap sufficiently, indicating a similar probability to be included in both groups, mimicking a standardised experiment.[13] Appendix Figure A2.3 shows that this assumption is partially satisfied, and about 21 per cent of the taxpayers are dropped. Matching weights for them could not be built, as taxpayers lie out of the common support. This means that the following estimates refer to 80 per cent of taxpayers in the register – a large proportion, which ensures generalisability to the overall taxpayer population.

As a last note, our analysis does not look at the impact on taxpayer compliance behaviour and revenue collection, but focuses on data quality and tax administration functions. The main reason is that most Instant TIN registrations took place in 2022 (Figure 3.1), and new registrations are only required to file their tax returns by the end of December 2023. We decided to give these taxpayers enough time to file, and to focus on other outcomes that matter for the internal functioning of URA. Future research will focus on the impact of Instant TIN registrations on compliance and revenue collection.

[12] We selected 0.06 as the bandwidth in the kernel matching algorithm, as common practice in the literature.

[13] PSM builds upon the assumptions of: (i) unconfoundedness – which requires that all confounding factors be included in the set of covariates used to calculate the propensity score; and (ii) overlap in log odds or common support – which requires that taxpayers with the same score can be either registered with Instant TIN or the normal process, as in a randomised experiment.

4. Results

4.1 Anatomy of Instant TIN registrations

As a first exercise, we try to map the correlates of Instant TIN registrations using taxpayer-level data. This analysis compares Instant TIN registrations from January to November 2022 to normal registrations from 2019 to 2022.[14] First, we compare key taxpayer features by the two groups, and test whether mean differences are statistically significant, as shown in Table 4.1. All mean differences are statistically significant, primarily due to high statistical power from a large sample of about 1 million registrations. Some immediately interesting patterns emerge from the table. First, those registering on Instant TIN are more likely to be born in Kampala, and more likely to be assigned to tax centres out of it. Second, due to their sheer number, almost all Instant TIN registrations (99 per cent) are for individual taxpayers. This is also because this was more feasible for individuals than for companies. Third, female and younger taxpayers are much more likely to register through Instant TIN than the normal process. Fourth, previously informal businesses are much more relevant in the subgroup of Instant TIN registrations (42 per cent) than in normal registrations (4 per cent), where, instead, officially registered businesses with other government institutions are the majority (47 per cent). Finally, the in-the-field arm of TREP is responsible for about 40 per cent of Instant TIN registrations, mostly individuals. A negligible proportion of normal registrations was done through TREP (1.3 per cent).

Putting all these features in a single framework, Figure 4.1 below reports the coefficient plot from the linear OLS estimation in Equation 1, for both individuals (subgraph a) and companies (subgraph b). Again, the coefficients are highly statistically significant due to the large sample size. For individuals (Figure 4.1a), the plot indicates that two key factors strongly correlate with the probability of registering with Instant TIN – being an informal business, and registering through TREP. This is not surprising. TREP officials, performing in-person registrations in door-to-door campaigns and at one-stop shops, mainly targeted potentially active individual businesses that were still informal. They also assisted taxpayers who needed guidance when voluntarily asking for tax registration. Data shows that as many as 44 per cent of previously informal businesses registering with Instant TIN did so through TREP, compared to 2 per cent of those registering with normal procedures.

[14] Results remain consistent if we consider normal registrations in the same period of Instant TIN registrations, January to November 2022, or the same period one year earlier, in 2021. We opted for considering the whole period 2019-2022 to enhance the generalisability of our results.

Table 4.1 Mean differences by type of registration

	Normal 2019-22		Instant TIN 2022		
	Mean	Obs.	Mean	Obs.	Difference
Kampala tax centre	0.36	633,478	0.30	350,672	0.07***
Kampala born	0.08	633,478	0.13	354,600	-0.05***
Individual	0.88	633,478	0.99	354,600	-0.12***
Female	0.31	555,904	0.36	351,983	-0.05***
Income source: employment	0.55	555,904	0.48	351,983	0.07***
Income source: registered business	0.47	630,175	0.03	354,429	0.44***
Income source: assets	0.02	559,594	0.05	354,429	-0.03***
Income source: rental	0.02	557,599	0.03	354,429	-0.01***
Income source: informal business	0.04	633,478	0.42	354,600	-0.38***
<18	0.00	633,478	0.00	354,429	-0.00***
18-24	0.06	633,478	0.13	354,429	-0.08***
25-34	0.43	633,478	0.48	354,429	-0.05***
35-44	0.21	633,478	0.21	354,429	0.00***
45-54	0.09	633,478	0.09	354,429	-0.00
55-64	0.05	633,478	0.04	354,429	0.00***
>=65	0.17	633,478	0.04	354,429	0.13***
TREP reg.	0.01	633,478	0.39	350,672	-0.37***
N	988,078				

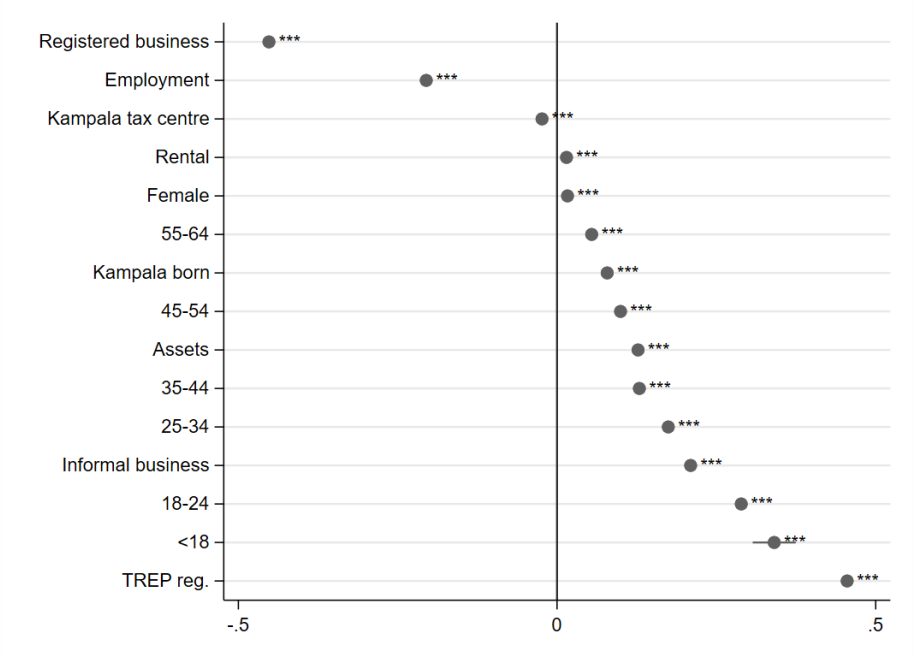
Source: Own calculations on URA administrative data. * p < 0.10, ** p < 0.05, *** p < 0.01.

At the same time, being a formally registered business with other government institutions – an essential prerequisite for registering with URA – is negatively associated with Instant TIN registrations. Although marginal, gender plays a role, as being female is positively associated with the simplified registration procedure. The age distribution also shows a specific pattern in which older taxpayers are much less likely to register with the new technology than younger ones. All the age categories displayed are more likely to register with Instant TIN than the excluded category – the oldest taxpayers (>65). Importantly, the magnitude of the positive correlation with Instant TIN is higher, the younger the taxpayer.

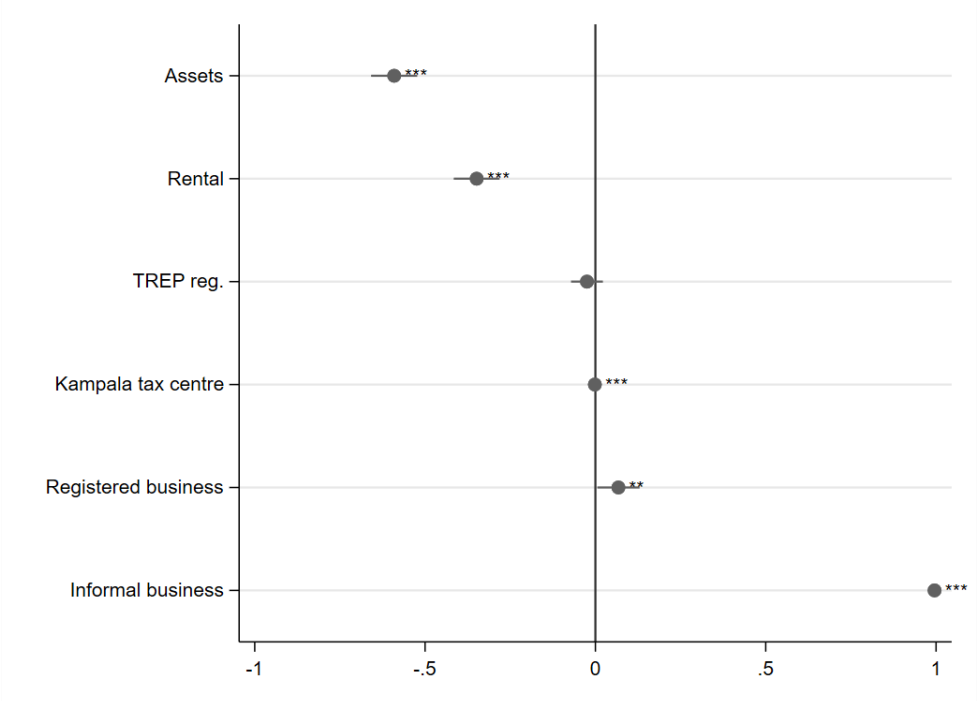
For companies (Figure 4.1b), the availability of information is more limited, hence the regression considers a smaller set of features. Nevertheless, the primary correlate with Instant TIN registration is being previously informal, consistently with what is found for individual taxpayers. TREP is no longer a key factor, as a negligible minority of companies registered through the programme (0.03 per cent).

Figure 4.1 Correlates of Instant TIN registration, OLS framework

(a) Individuals



(b) Companies



Source: Own calculations on URA administrative data.

Some considerations can be made based on the evidence above. First, the new technology, which is simpler and much quicker, can attract different categories of

taxpayers. These are more likely to be women or younger individuals, who might find it easier to register with this technology rather than going through the more burdensome normal process.[15] Female and younger individuals could feel less comfortable navigating complex bureaucratic spaces and be less knowledgeable about the tax system, as shown in the literature (Santoro *et al.* 2023). Female taxpayers might appreciate the online procedure because it is simpler, and you can register without interacting with tax officials. In this sense, a simplified online system like Instant TIN has the potential to make the tax system more inclusive and accessible. In addition, registering taxpayers when they are younger ideally fosters a culture of compliance to be followed throughout their lifetime.

As our interviews showed, Instant TIN is a revolution for how taxpayers can register. At least in theory, and coherent with the positive taxpayer feedback collected by URA, the online procedure of Instant TIN massively reduces the compliance costs of registration for tax purposes. It eliminates the need to visit the URA offices and present numerous documents for identification purposes. Moreover, the instant issue of a TIN relieves taxpayers from having to wait (at least) two days to obtain a TIN, allowing them to start operating immediately.[16]

Second, the role of in-the-field TREP campaigns in boosting Instant TIN registrations seems central. On one hand, this shows how intense tax registration programmes can use technology to reach their targets, ideally more easily and quickly. On the other, due to data limitations discussed in Section 3, the more practical question remains – whether TREP Instant TIN new registrations performed the registration by themselves, if they were assisted, or if a TREP official registered them on their behalf. Whichever way, we know that registrations through Instant TIN during in-the-field TREP campaigns were not performed completely independently by taxpayers. More field research would be needed to fully understand the journey taxpayers undertook when registering in the context of TREP, as this experience could impact new entrants' perceptions, and shape future compliance. The following section will further analyse registrations to understand whether the use of this technology by TREP officials led to gains from the URA side. There is the risk that, because TREP targeted the number of new registrations with little interest in quality, this quick, convenient technology led to somewhat flawed registrations with poor quality taxpayer information.

Third, the new technology seems pivotal in formalising previously informal entities. As further shown in Appendix Figure A2.5, there are many previously

[15] In line with Appendix Figure A2.2, the age distribution of Instant TIN registration shifted to the left compared to normal registrations.

[16] From interviews with URA officials, URA15 said that the new registration technology: ‘is instant and avoids people to wait for days or a longer period before getting a TIN. They are not trapped in this waiting time, but immediately get a TIN. There is a better service cause acquisition is now faster and clear. Taxpayers are happy’ (URA15, Kampala, December 2022).

informal businesses within Instant TIN registrations, and only about 3 per cent chose the normal route to register with URA. This could be due to the simplified process, which may be preferable for less sophisticated, less tax-savvy and financially excluded businesses. At the same time, however, this finding depends on the TREP formalisation strategy, which, by policy intent, primarily targeted informal entities that are more visible and more easy to reach on the ground – 40 per cent of previously informal entities are registered through in-the-field TREP. Critically, the majority of informal taxpayers, 60 per cent, opted for Instant TIN outside of TREP. This may hint again to a strong preference for this simplified digital solution from the informal category, in turn likely to intersect with the previously discussed female and younger new entrants category. It is still unclear, however, if tax compliance after registration will be adequate, especially for entities, such as informal businesses, who did not register through the normal process, and might be less familiar with the requirements and complexities of the tax system. Future research can shed light on this.

4.2 Impact on data quality

Furthermore, we explore the data quality of Instant TIN registrations, and its implications for strategic data used by URA. Table 4.2 reports the OLS coefficients on five different data quality outcomes, as explained in Section 3. All coefficients are highly statistically significant due to the large size of the groups we analyse, which enables precise estimations.

Table 4.2 Correlation between Instant TIN registration and data quality

	(1)	(2)	(3)	(4)	(5)
	Multiple TIN (Indiv. only)	**Invalid email**	**Invalid phone number**	**Age outlier (Indiv. only)**	**Invalid sector (Indiv. only)**
Instant TIN	-0.01***	-0.08***	-0.06***	-0.00***	0.12***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Controls	Yes	Yes	Yes	Yes	Yes
Normal reg Y	0.011	0.193	0.162	0.066	0.070
R-sq.	0.000	0.853	0.271	0.798	0.658
N	691373	693970	693970	691332	250815

Source: Own calculations on URA administrative data.
Note: Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. All regressions are run on a sample of Instant TIN and normal registrations matched through PSM, as described in Section 3. Column 1 refers to an indicator for duplicates in taxpayer identity, whereby taxpayers with the same name and date of birth have different TINs. Column 2 uses an indicator for invalid email addresses, including missing and duplicate ones. The same applies to the outcome in column 3. Age outlier in column 4 is an indicator for age being more than twice the standardised difference from the average population age. Column 5 uses an indicator for invalid sector, including missing. Columns 1, 4 and 5 refers to individuals only, while PAYE taxpayers are also excluded in column 5. More details on the estimation strategy in Section 3.

A first key finding is that Instant TIN registration helps reduce the multiple TINs for individuals (column 1). This is assured by design, since an Instant TIN is generated upon the submission of a NIN – the latter being unique – and only one TIN can be generated for every NIN. While multiple TINs were not a big issue in normal registration procedures (around 1 per cent), they reduced to zero when national identity numbers are used for registration.[17]

Regarding the validity of contact details – column 2 shows that Instant TIN registrations correlate with a significant reduction in the probability of seeing an invalid, missing, or duplicate email address. The reduction of 8 percentage points is also quite sizeable compared to the average (19 per cent) of invalid email addresses among normal registrations.[18] Importantly, Instant TIN registrations present an overall higher share of invalid emails, as shown in Appendix Figure A2.6. However, this negative pattern is primarily due to in-the-field TREP registrations, where tax officials would enter dummy email addresses to proceed with the registration form, as discussed below. Column 2 indicates that when TREP is accounted for as control in the OLS framework, the negative effect of Instant TIN registrations on email validity dissipates, and a positive effect on email accuracy appears. This means that the negative effect of TREP is so strong on the outcome that, when included, it shows the actual positive effect of Instant TIN registration on email address quality.

There is, however, the possibility that the register cleaning unit at URA is responsible for the positive association of Instant TIN with the quality of email data. Our interviewees at URA stressed that one of the main flaws of the Instant TIN system is that it does not validate email addresses or phone numbers.[19] Initially, 'the design of Instant TIN required that an officer contact every newly registered taxpayer, and validate the information, amend, and collect missing' (interview URA16). URA officials also stressed how this system feature initially caused challenges in contacting taxpayers to bring them on board. Because of this and other flaws in data, URA has mobilised internal resources to increase its capacity to clean the register as Instant TIN data comes in.

[17] Only a tiny 0.05 per cent of Instant TIN registrations present multiple TINs.

[18] The significant presence of duplicates among normal registrations could be explained by the practice of tax agents using their email address for all their clients (Mayega *et al.* 2019). The fact that duplicate emails almost disappear for Instant TIN registrations might suggest a very marginal role of tax agents, and a more direct presence of the taxpayer themselves at the point of registration, which is simpler and needs less external support.

[19] The Instant TIN system does not validate email addresses, which is why taxpayers' reliance on agents or other intermediaries for registration led to numerous email duplicates, as confirmed in interview URA12: 'Email - there might be duplicates of it because there might be agents helping people filing or paying'.

This said, a similar positive effect on data quality is found on mobile numbers (column 3), with a large decrease of 6 percentage points in invalid, missing, or duplicate numbers in Instant TIN registrations, or about 36 per cent reduction compared to the control mean. Most invalid phone numbers in normal registrations are due to duplicates, now reduced to a negligible share (Appendix Figure A2.7). These results must be understood similarly to the above on email addresses. The Instant TIN technology does not allow validation of phone numbers, leading to invalid entries – either due to the hasty way TREP officials seem to have performed their registration task, or to the (voluntary or involuntary) inaccuracy of data entered by taxpayers. As gathered from our interviews, this invalid information has been partly addressed with a register cleaning exercise.

Furthermore, no significant effect is found on age outliers and the accuracy of age information (column 4), despite Instant TIN registrations being significantly younger than normal ones (Appendix Figure A2.4). Finally, some worrying evidence on the accuracy of sector information emerges from column 5, indicating that registering with Instant TIN increases the likelihood of having an invalid or missing sector record. The 12 percentage point increase is almost double the average of invalid sectors in normal registrations. This result is probably due to the fact that providing sector information is not mandatory when registering as an individual taxpayer – which it is when registering as a company, hence the absence of gaps in this information for incorporated taxpayers.[20] The information about the sector of economic activity for companies is directly retrieved from URSB data, which is why the sector field is more accurate and always present for this category of taxpayers. Despite this, our interviews with URSB officials highlighted how the reported business sector in URSB data might still not reflect current business activity. In Uganda, there is no regulation requiring businesses to update their business sector with URSB, even if the nature of business activity substantially changes – which it often does.

Additional interesting evidence arises when disaggregating by individual and company registrations, referring to the separate NIRA and URSB data sharing processes, respectively (see Section 2). Table 4.3 below shows that most benefits in data quality and accuracy come from individual Instant TIN registrations. Companies registering with Instant TIN do not provide better quality email addresses, and, even more alarmingly, are more likely to have an invalid phone number. This might be because Instant TIN registrations for companies were enabled later than for individuals – as reflected in the low number of Instant TIN registrations of companies. At the time of analysis, the register cleaning performed for individual new registrations had not been carried out for

[20] We removed companies from the regression as the outcome has no variation for them, hence it is not sensible to run a regression including these units.

businesses. As mentioned above, URA has recently halted the Instant TIN platform for businesses, as they are working on improving it.

Table 4.3 Correlation between Instant TIN registration and data quality by taxpayer type

	(1) Invalid email indiv.	(2) Invalid email company	(3) Invalid phone number indiv.	(4) Invalid phone number company
Instant TIN	-0.09***	-0.00	-0.09***	0.31***
	(0.00)	(0.04)	(0.00)	(0.04)
Controls	Yes	Yes	Yes	Yes
Normal reg Y	0.161	0.416	0.115	0.503
R-sq.	0.858	0.002	0.013	0.043
N	691373	2597	691373	2597

Source: Own calculations on URA administrative data.
Note: Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. All regressions are run on a sample of Instant TIN and normal registrations matched through PSM, as described in Section 3. Columns 1 and 2 refer to an indicator for invalid email addresses, including missing and duplicate ones. The same applies to the outcome in columns 3 and 4. More details on the estimation strategy in Section 3.

As a last piece of evidence, results from Table 4.2 do not change if we restrict the comparison group to normal registrations in 2022, in parallel to Instant TIN registrations. Appendix Table A1.4 reports the OLS results. No change is found on multiple TINs, age, and invalid sector information. Results on accuracy of contact details keep the same direction, with Instant TIN reducing inaccuracies – if anything, they get smaller in magnitude. Likewise, results do not change if we perform a naïve OLS regression without imposing any PSM between Instant TIN and normal registrations. Appendix Table A1.5 indicates that coefficient estimates keep the same direction and significance level. If anything, naïve estimates are larger in magnitude, most likely because the two groups are quite different. This implies that the smaller PSM-based coefficients are more conservative and realistic after removing extreme imbalances across groups through the matching algorithm.

4.3 Impact on tax administration's core functions

What does this evidence mean for the potential of Instant TIN – enabled by inter-institutional data sharing – to improve the core functions of URA? Our findings clearly suggest important repercussions on how URA performs its duties, specifically with registration and identification, enforcement, and taxpayer assistance. We discuss both the immediate benefits and persisting challenges

from integration, with the support of qualitative information from in-depth interviews.

There is a widespread view among URA, NIRA, and URSB officials that integrating registers, foundational to the Instant TIN platform, is the way forward for tax administration and the other agencies. From URA's point of view, the current point-to-point data sharing is beneficial – both the integration with NIRA, to make it easier to identify individuals, and with URSB, to make it easier to identify companies.

The Instant TIN platform has served two primary purposes – taxpayer identification and registration. On one side, it represented the main channel of voluntary registration for income tax and VAT. Conversely, Instant TIN has been a powerful tool for TREP, helping URA contractors reach very high registration targets. The benefits of Instant TIN for these functions of tax administration come primarily in avoiding duplication. An individual or company can only register once with a NIN or BRN, reducing error-prone manual entry of data, and time and resources saved in the registration process.[21]

Instant TIN also seems to have significantly improved the facilitation function of URA. The data shows that the success of Instant TIN technology as a registration tool is especially for individual taxpayers who were previously informal, younger, and female. URA officials say this is driven by reduced compliance costs for the less tax-savvy, less established businesses. At least in the first registration stage, Instant TIN saves individuals from visiting URA to go through the registration procedures, identification, validation of documents, manual filling of registration forms and provision of biographic data and information on the economic activity.[22] In addition, it has reduced the waiting time for a TIN, allowing businesses to start operating immediately.

Both uses of Instant TIN – online voluntary registration and the TREP-facilitated one – seem to have brought, along with the benefits, some challenges to the revenue administration. As shown in the section above, the data collected from this technology is not always of good quality. As we could extrapolate from our interviews, there are at least three reasons for this. First, the data stored in NIRA and URSB datasets is not updated as frequently as it should be. Thus, when URA uses information from these sources (location, business sector, etc.), it populates

[21] URA11: 'The objectives of the integration initiatives are three-fold: to increase the taxpayer base, to make URA's registration touchpoint less error-prone and improve the quality of data in its taxpayer register, and to improve taxpayers' experience of URA registration and filing'. And: 'the register was growing, but the data has been so bad. The only way they could resolve the data issue in terms of the individuals was to look at the mother registration agency: NIRA'.

[22] From interview URA11, regarding the usual registration modality: 'the overwhelming taxpayer feedback is that the existing registration system is much too cumbersome, doubly so because taxpayers must provide the same information at multiple touchpoints', and interview URA14: 'Instant TIN is the way to go because you save a lot of time with paperwork'.

its register with mistakes. Second, the Instant TIN portal does not validate information entered by the taxpayer. This applies to practically all fields, starting with the taxpayer's identity. While the system can check if the NIN or BRN exists, it cannot validate this information against who is making the application. In addition, it does not authenticate contact information or any other field. If the taxpayers – deliberately or not – enter the wrong data, the system accepts it. Third, TREP officers are external officials with short-term contracts, paid based on the number of registrations performed, independently of their quality.[23] This suggests – and seems to be confirmed by the administrative data – that the registrations performed during TREP were not of good quality.[24] A critical trade-off arises between quick registration targets, and collecting accurate taxpayer information.[25] This trade-off is conceptually like that recently documented. It explains the prevalence of nil filing in Africa, and the urgency of increasing registration numbers, without much focus on the quality of tax return information submitted by new entrants (Moore 2022).

The poor quality of the data entered through Instant TIN has repercussions for URA for administrative costs and the potential for mobilising revenue. For the former, they are both the additional resources spent on register maintenance, which has increased since registration has been done through Instant TIN, and the cost of contacting each taxpayer after registration via SMS or a phone call to ask them to update their details in person at the URA offices.[26] One URA official explained: ‘The register update and cleaning function is done every year. The challenge is the numbers, because the speed of new registrations means that for each taxpayer for which you update information, you register many with inaccurate information. This links to the registration drive’ (URA16). In addition: ‘Cleaning URA's register is also going to be a critical challenge, not least because information is captured differently in each participating register. The only way to solve this is to get taxpayers to come back and update their information’ (URA11).

Costs in terms of the potential for revenue mobilisation manifest for two reasons. First, because storing wrong taxpayer information in the register does not allow

[23] The TREP target for the financial years 2021/2022 and 2022/2023 were, respectively, 800,000 and 900,000 new registrations, staggering numbers when compared to the roughly 1.5 million registered in the 15 years previous to 2021 (interviews URA23, URA16).

[24] Descriptive evidence from administrative data suggests that 40 per cent of Instant TIN registrations had an invalid email address generated by officials during the registration process.

[25] URA16: ‘With Instant TIN they take the risk of assuming the taxpayer will give correct information, but of course this doesn’t always happen. There is a trade-off between giving a quick service and accuracy of data’. Again: ‘This is to get a fast service. However, you also have to assume that people are not willing to provide right information to the TA. Because the tax culture is very low. The traditional registration process comes with an audit of the correctness of information. Instant TIN does not’.

[26] As of December 2022, about 60 staff were involved in the register maintenance team, of which 20 are short-term positions.

the revenue authority to bring these taxpayers on board, educate them and eventually enforce the tax law on them. For example, if the business address or contact information is invalid, URA cannot reach the taxpayer and offer education programmes and other support initiatives. Similarly, if taxpayers provide invalid sector information, or the system retrieves an outdated sector code from URSB, URA can only know the correct information if the taxpayer comes in to provide it. Second, once registered with URA on the Instant TIN platform, taxpayers must visit URA to update their contact and location information. This step is mandatory for compliance with VAT. According to our interlocutors at URA, roughly 50 per cent of Instant TIN registered taxpayers come to amend their information and start the compliance journey. Since many taxpayers have used Instant TIN for VAT purposes, these gaps in following up on the registration process may imply significant compliance issues for a crucial tax in the country. Allowing validation of information online would probably reduce compliance costs for some of these taxpayers, increasing URA's capacity to mobilise revenue.

Likewise, not being able to reach taxpayers due to invalid contact details or business locations has a negative implication for the enforcement function of URA. As with taxpayer education and support, the revenue authority cannot contact taxpayers to enforce the tax law if there is no way to locate them.

5. Conclusion and policy recommendations

This study is a preliminary assessment of Instant TIN, an innovative taxpayer registration strategy launched in Uganda in 2022. Instant TIN relies on other government agencies sharing data – the NIRA and URSB. This data identifies the applicant automatically, and instantly generates a TIN. Registration with URA through Instant TIN is available to individuals and businesses, and has been deployed with the promise of facilitating taxpayer registration and boosting the tax administration's functions. Data quality was also thought to benefit in a context like Uganda, historically characterised by gaps and inaccuracies in tax administration data (Section 2).

We present a mixed picture of the findings, using administrative data from URA and several in-depth interviews (Section 3). On one hand, the first-stage impact on the sheer number of registrations is impressive. Instant TIN technology dramatically increased registration numbers, amounting to 85 per cent of all registrations in 2022 alone – when excluding forced registrations – practically replacing normal voluntary registrations. We show that the new technology is strongly connected to the TREP programme – it is a critical tool field officers use to simplify their activities and reach higher registration numbers, as per the programme's mandate. Not surprisingly, previously informal businesses seem to use Instant TIN technology more, and are a large proportion of the total – although we do not know whether they voluntarily opted for it, or have been made to register by TREP officers in their campaigns. Being younger, and, albeit only marginally, female, also correlates with using the technology.

On the other hand, a more complex picture emerges when understanding the impact on data quality and URA functions. Analysis of administrative data indicates that Instant TIN technology reduces inaccuracies in email addresses and phone numbers, yet it worsens the quality of business sector information. The new tool is also ineffective in improving data quality for companies. Furthermore, evidence from in-depth interviews suggest the cost for URA of these flaws in the data obtained through Instant TIN. This is mainly the massive data cleaning effort for the register maintenance team to correct the inaccuracies brought in by using Instant TIN in TREP, and potential revenue losses due to it being impossible to make use of the information – mostly contact and location details – to bring taxpayers on board, or enforce the tax law.

We make some policy recommendations based on the evidence above. These can be useful in informing future data sharing initiatives, where other institutional datasets could eventually be merged with URA's register.

First, Instant TIN technology could significantly improve its capacity to validate third-party information. While having access to individuals' and companies' identification data from third parties is already a remarkable success, the URA system could be upgraded along different dimensions. The Instant TIN portal should be able to validate the taxpayer's email address and phone number. For example, the contact details provided during registration could be validated by sending a one-time passcode, which taxpayers must enter into the Instant TIN platform. This would guarantee the tax administration control over the quality of data that populates its taxpayer register, resulting in fewer wrong contact details. Such an improvement would imply more effective facilitation and enforcement functions at URA, making it possible to reach the totality of new taxpayers to offer support or enforce the tax law. Better quality data in the register would also spare the revenue authority from employing significant resources to nudge taxpayers to update their information, and clean the taxpayer register of invalid taxpayers' records. This is happening, especially for TREP registrations.

Second, the revenue authority should find a way to authenticate the identity of the person registering through Instant TIN, to prevent people from registering newborns or the deceased for tax purposes to decrease their tax liability. While the problem could be partly solved if NIRA incentivised people to update their information and notify the death of their family members,[27] using a facial or biometric recognition technology at the time of registration will also help the URA registration function.[28]

The above point relates to our second policy recommendation. URA would benefit from shifting from the current point-to-point data sharing practice with NIRA and URSB to a data integration system, if only with these two agencies. With such a shift, URA could operate with up-to-date information. The current system does not notify URA (or NIRA and URSB) when citizens update their information with other agencies. In theory, allowing one-touch registration to the agencies involved and real-time information updates could reduce the registers' vulnerability to error-prone manual entry, or voluntary entry of different information.[29] Other benefits of

[27] Relatedly, an interviewee at URA told us: ‘URA needs to build strategic partnerships with other entities to have unique identifiers and impose on taxpayers to enter reliable data. For example, link social security number, TIN, bank account’ (URA12).

[28] From interviewing NIRA, we gathered that there is a plan to implement a fully integrated system between URA and NIRA, which would be set up using VPN over the internet, accessible through digital signatures on both sides. The data sharing would be two-way, allowing information (name, date of birth, residence, contact, gender, other data, etc.) to be shared in real time between the two registers. This – and other system integration – is expected to be facilitated by a digital ID, which would be the key identifier across multiple registers: ‘NIN should be integrated with taxation, financial services, public safety, etc. It needs to permeate the entire economy, especially the informal economy, for it to be truly beneficial for tax’ (NIRA11).

[29] URA16 said: ‘Because the government has to make sure the person provides data at only one point, to avoid duplicating data on the same person across agencies. The information on address is a challenge also, because citizens decide how to give their data depending on what they need from a certain agency. Again, a single point of registration would solve the problem’ (URA16). And: ‘Companies annual tax returns have to

such an integration would be savings in registration costs for all the institutions and for taxpayers, who would be spared from visiting multiple offices.[30]

Apart from benefitting the URA registration function, data integration would improve URA monitoring and enforcement functions, as outlined in the URA11 interview:

> Previously, businesses could operate informally and effectively out of reach of URA. With the URSB integration, URA has real-time access to BRN data and therefore knows immediately when a business is registered, which means it can follow up directly with the business owners or indirectly with the local licensing authority to determine the entity's tax obligations. If businesses know URA has access to this data, they are more likely to comply with their tax obligations.
> (URA 11)

Third, more thought should be devoted to the policy targets and intent around registration and formalisation. Relatedly, more consideration should be given to how to measure the success of technological innovations, and which performance metrics to look at. The increasing attention to tax registration in Uganda is in line with unprecedented mass registration campaigns seen in other African countries (Moore 2022). Our evidence indicates that technologies like Instant TIN can help dramatically improve registration numbers. However, some concerns remain on how the functions of the tax administration can benefit from the great registration efforts, of which Instant TIN – and technological innovations in general – have become a prominent tool. It is particularly unclear whether URA has the capacity in personnel and resources to adequately assist and monitor exponentially increasing numbers of new entrants every year, especially when the data collected on new taxpayers is of poor quality. Evidence from other contexts shows that mass registration drives do not always bring the expected benefits in terms of revenue to tax administrations, and that complications in managing registers bloated with inactive taxpayers could be important (Lediga *et al.* 2020;

be done with URSB and URA. They should speak to each other. For example: when they file annual tax returns with URSB they give info about shareholders, shared capital, etc. But URA doesn't receive this info, and it should because it has tax implications. URA gets only partly this information at the moment of filing, because citizens know which information, they want to give to URA and which not. Companies and individuals do not correct information to NIRA, like the address. The process is long, manual, and tedious. If there was a benefit attached to this, they would go and update. For example, free healthcare in specific healthcare centres based on the address could be an incentive to update address' (URA16).

[30] Interviewee from URSB: 'Why does URA do business registration when its function is revenue collection? They could take the information from URSB. And if it does not have enough information, they can increase the amount of information collected at registration at URSB, so the TIN could be issued automatically. Or,use the BRN as a TIN. The costs and process would just improve. Also the quality of the data' (URSB11).

Mascagni *et al.* 2022; Moore 2022). A solution could be to consider whether to prioritise entities already in the tax net, with higher potential value, to maximise their revenue potential and improve compliance behaviour, which is thought to be, at best, suboptimal. An example is given by URA's remarkable efforts in targeting high net worth individuals. These high-value entities have been successfully identified and registered, but their tax compliance remains difficult to encourage and improve due to a mix of political and technical challenges (Santoro and Waiswa 2022).

Our policy recommendations are relevant beyond the Ugandan context. As mentioned, URA is joining a broader trend of revenue authorities using third-party identification data to register new taxpayers. For example, a parallel study that we are running in Ghana reveals how the Ghana Revenue Authority is expanding its income taxpayer register by migrating data from the National Identification Authority, with the risk of bloating the register with inaccurate – thus unusable – information. As explained above, if the risks of obtaining flawed data are not minimised, the costs of mass registration campaigns could be significant to the revenue authorities, and the tax administration functions might not benefit from the process.

Further research is needed on this subject. Three critical limitations of this study imply that the evidence we produce is still preliminary and descriptive. As a first limitation, this study can only partially isolate the impact of data sharing on the quality of data, given the register cleaning process that URA continuously performs internally. The results of this paper are still valuable, since we gathered that data cleaning activities are performed mainly for the information obtained through TREP registration. Further investigation would be necessary to shed light on the exact functioning of these internal processes.

Second, this research primarily revolves around the internal functioning of the tax administration, exploring implications for data quality and core URA functions. Very little can be gauged around which taxpayer profiles opt for Instant TIN registrations, how taxpayers directly experience the new process, and how this process ultimately affects their perceptions. As stated above, especially for informal businesses, it is unclear if their use of Instant TIN is driven by a voluntary choice to register, or imposed by TREP officers. These two different experiences can have a divergent impact on future compliance. As of now, anecdotal evidence suggests that taxpayers appreciate the simpler, quicker registration process, but more detailed and nationally representative survey data would be helpful in fully understanding the taxpayers' experience. For instance, it is unclear how taxpayers perceive the requirement to visit URA for income tax or VAT registration after Instant TIN registration. Given mixed evidence on taxpayers' visits to URA after registration (Section 4.3), we speculate this is an extra burden for them.

Third, and related, this study is not suited to measuring the impact on taxpayers' compliance behaviour, which we leave to future research. This is due to the timing of the intervention – 2022. It is a recent innovation, and it is too early to investigate the impact on filing and payment behaviour. We plan to run this analysis in the second half of 2023, after the 30 June 2023 filing deadline. After that, we can observe the filing response of Instant TIN registrations for the year from July 2022 to June 2023, and understand whether compliance improves when compared to normal registrations. A very preliminary inspection of administrative data indicates a positive pattern of payment behaviour, which follows a different timeline and is often disconnected from filing behaviour (Santoro and Waiswa 2022). As shown in Appendix Figure A2.8a, the probability of paying any tax in 2022 is slightly higher for Instant TIN registrations than normal ones – although it is remarkably low in general. However, amounts paid by Instant TIN registrations seem significantly smaller (Appendix Figure A2.8b). The latter fact is probably due to the higher incidence of presumptive taxpayers among Instant TIN registrations (16 per cent) than among normal ones (10 per cent). By design, presumptive taxpayers are smaller and pay a fixed amount per year. This also reinforces the above evidence on previously informal businesses driving most informal entities' Instant TIN registrations. These informal businesses are arguably smaller, hence remitting less tax. In sum, considering tax registration as a process, rather than a mere stand-alone action, future research should explore how a given registration experience reverberates into repeated tax compliance decisions throughout a taxpayer's life.

Appendices

Appendix 1 Tables

Table A1.1 Datasets used in last phase of TREP

Datasets used in last phase of TREP	
1.	Architects' registration data
2.	Hotel sector
3.	IFMS (Integrated Financial Management System)
4.	IRAS (Integrated Revenue Administration System)
5.	KCCA directors' data
6.	KCCA property details
7.	KCCA trade licence data
8.	KCCA TREP registration data (two datasets)
9.	Landlord data
10.	Landlord in Kasese
11.	Local government of Arua City
12.	Local governments
13.	Local Governments' Trading Licence Register
14.	Ministry of Education
15.	National Water
16.	NSSF
17.	Pharmacists' registration data
18.	Private sector schools
19.	Property location
20.	UBOS
21.	UMEME Limited

Source: In-depth interviews with URA.
Note: The table gives the names of the registers that URA is using to carry out the last phase of TREP, not the name of the institutions that provided them.

Table A1.2 In-depth interviews

	Institution	Department	Date
1	URA	Business Intelligence	2 June 2021
2	NIRA	Management	17 November 2021
3	KCCA	IT	28 September 2022
4	URSB	Business Processes	3 October 2022
5	URSB	Planning	3 October 2022
6	URA	Business Intelligence	3 October 2022
7	URA	Enterprise Architecture	4 October 2022
8	URA	Service Management	4 October 2022
9	URA	IT	5 October 2022
10	URA	Business Analysis, Domestic Tax	5 October 2022
11	URA	Warehouse of data	5 October 2022
12	URA	TREP officers supervision	5 October 2022
13	URA	Register Cleaning, TREP	5 December 2022
14	URA	Service Management, TREP	7 December 2022
15	KCCA	Compliance and Risk Management	8 December 2022
16	URSB	Business Processes	12 December 2022
17	NITA-U	UG Hub team	30 January 2023
Roles of interviewees: executive directors, managers, supervisors, and officers.			

Source: In-depth interviews with URA.

Table A1.3 Codebook of thematic analysis

Codes	Description	Files	References
Functions	Functions of URA enhanced by the integration (identification, facilitation, enforcement, monitoring)	2	3
Impact data	Impact of Instant TIN on data quality	2	3
Email	Impact of Instant TIN on email address section of register	4	5
Location	Issues from the data about location	2	3
Phone	Impact of Instant TIN on the quality of phone number information	1	1
Processes and strategy	What impact the new technology and integration has on internal processes	1	2
Sector	Info on sector coming from Instant TIN registration	2	2
TIN duplication	Impact of Instant TIN on TIN duplication	4	5
Integration	Ways of integrating dataset across institutions	0	0
KCCA-URA	Integration between KCCA and URA data	4	10
KCCA-URSB-NIRA	Integration between KCCA, URSB and URA data	5	7
NIRA-URA	Integration between NIRA and URA data	4	9
URA-NITAU	Integration between NITAU and URA data	1	3
URA-URSB	Integration and data sharing between URSB and URA	5	8
Instant TIN - companies and individuals	Any information on the difference in uptake of Instant TIN or normal registration across different categories of taxpayers (companies or individuals)	3	3
Instant TIN - formalisation	Role of Instant TIN in improve formalisation	1	1
Instant TIN - info	General and relevant information about Instant TIN	4	6
Perspectives on implementation	Perspectives of other agencies on URA and integration with them, etc.		

Codes	Description	Files	References
Benefits	Benefits of integration from different agencies' perspectives	6	8
Challenges	Challenges of implementation of integration from each agency's perspective	8	15
Policy recommendations	challenges and corresponding policy recommendations	7	14
TREP role	Role of TREP in uptake of Instant TIN - what function did it have	6	9
Utility IT	Utility of Instant TIN to taxpayers	10	15

Source: Authors' own elaboration.

Table A1.4 Correlation between Instant TIN registration and data quality, 2022 registrations only

	(1) Multiple TIN (indiv. only)	(2) Invalid email	(3) Invalid phone number	(4) Age outlier (indiv. only)	(5) Invalid sector (indiv. only)
Instant TIN	-0.01***	-0.03***	-0.02**	-0.00	0.13***
	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
Controls	Yes	Yes	Yes	Yes	Yes
Normal reg Y	0.011	0.379	0.271	0.069	0.063
R-sq.	0.000	0.853	0.272	0.799	0.659
N	375,526	378,010	378,010	375,524	194,802

Source: own calculations on URA administrative data.
Note: Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. All regressions are run on a sample of Instant TIN and normal registrations matched through PSM, as described in Section 3. Column 1 refers to an indicator for duplicates in taxpayer identity, whereby taxpayers with the same name and date of birth hold different TINs. Column 2 uses an indicator for invalid email addresses, including missing and duplicate ones. The same applies to the outcome in column 3. Age outlier in column 4 is an indicator for age being larger than twice the standardised difference from the average population age. Column 5 uses an indicator for invalid sector, including missing. Col. 1, 4 and 5 refers to individuals only, while PAYE taxpayers are also excluded in column 5. More details on the estimation strategy in Section 3.

Table A1.5 Correlation between Instant TIN registration and data quality, without PSM

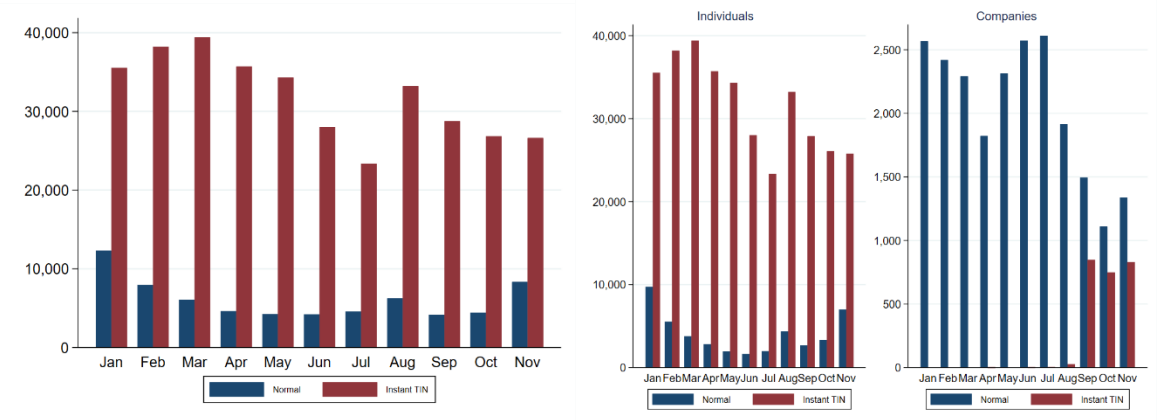
	(1) Multiple TIN (indiv. only)	(2) Invalid email	(3) Invalid phone number	(4) Age outlier (indiv. only)	(5) Invalid sector (indiv. only)
Instant TIN	-0.01***	-0.09***	-0.10***	-0.00***	0.40***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Controls	Yes	Yes	Yes	Yes	Yes
Normal reg Y	0.011	0.193	0.162	0.066	0.070
R-sq.	0.005	0.491	0.200	0.821	0.927
N	907,887	988,078	988,078	907,787	514,810

Source: own calculations on URA administrative data
Note: Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Such naïve regressions are performed without PSM. Column 1 refers to an indicator for duplicates in taxpayer identity, whereby taxpayers with the same name and date of birth hold different TINs. Column 2 uses an indicator for invalid email addresses, including missing and duplicate ones. The same applies to the outcome in column 3. Age outlier in column 4 is an indicator for age being larger than twice the standardised difference from the average population age. Column 5 uses an indicator for invalid sector, including missing. Columns 1, 4 and 5 refers to individuals only, while PAYE taxpayers are also excluded in column 5. More details on the estimation strategy in Section 3.

Appendix 2 Figures

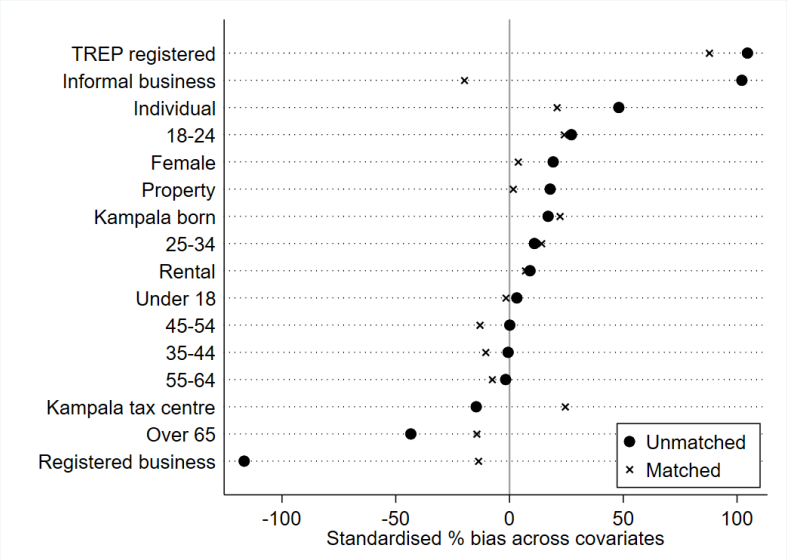
Figure A2.1 Trend in registrations by month

(a) Aggregated **(b) Disaggregated by taxpayer type**



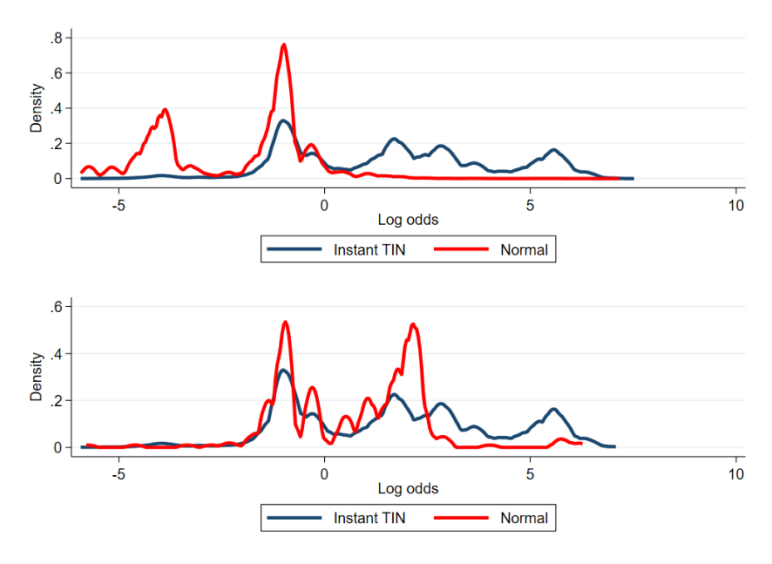
Source: Own calculations on URA administrative data.

Figure A2.2 Matching balance



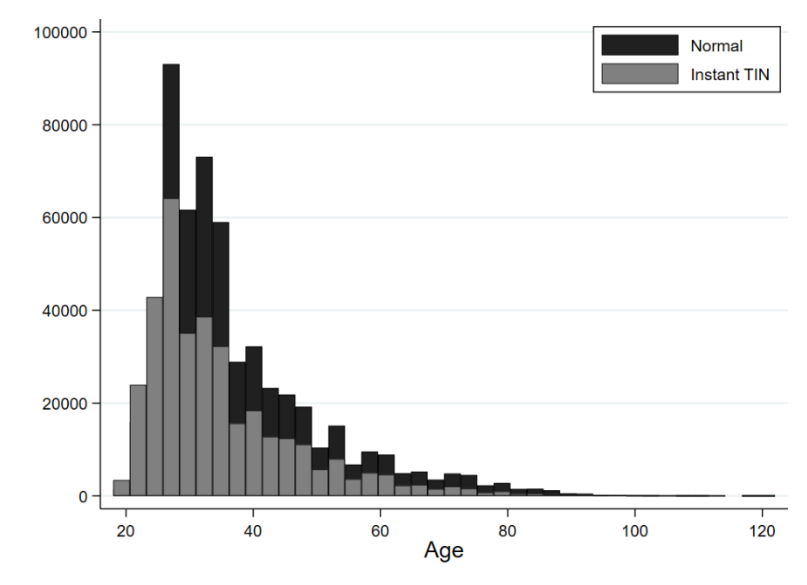
Source: Own calculations on URA administrative data.

Figure A2.3 Distribution of matching log odds



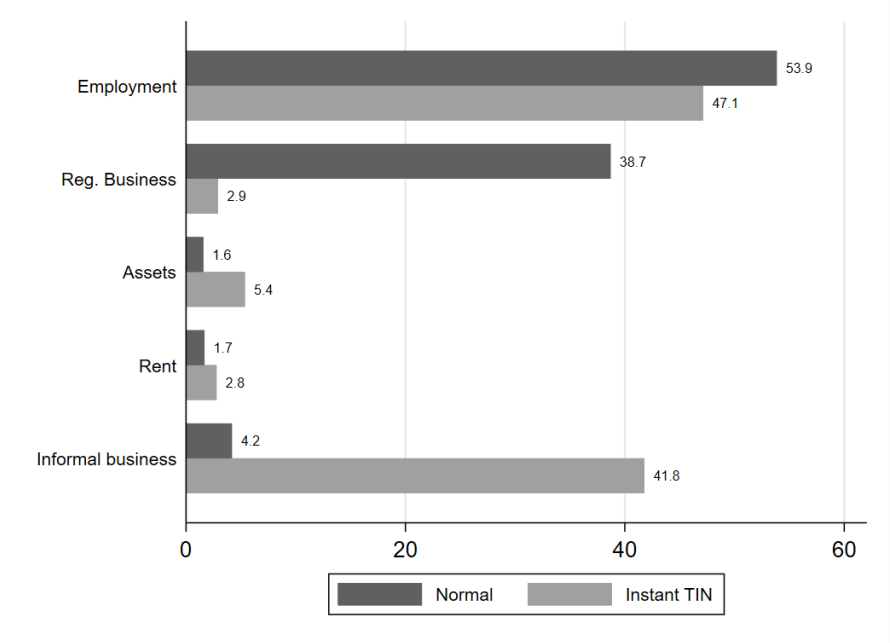
Source: Own calculations on URA administrative data.

Figure A2.4 Distribution of age by type of registration



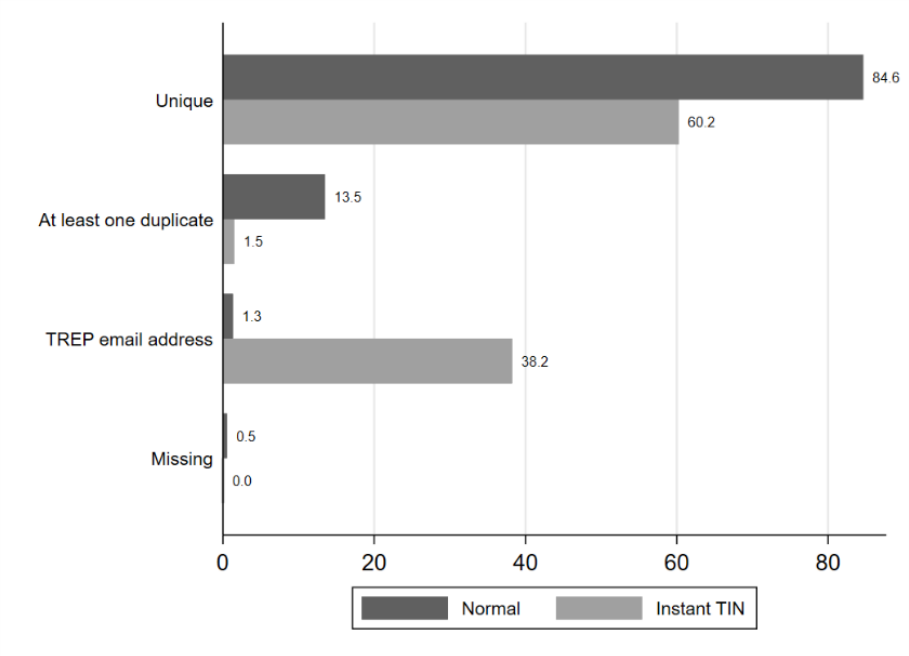
Source: Own calculations on URA administrative data.

Figure A2.5 Distribution of income sources by type of registration



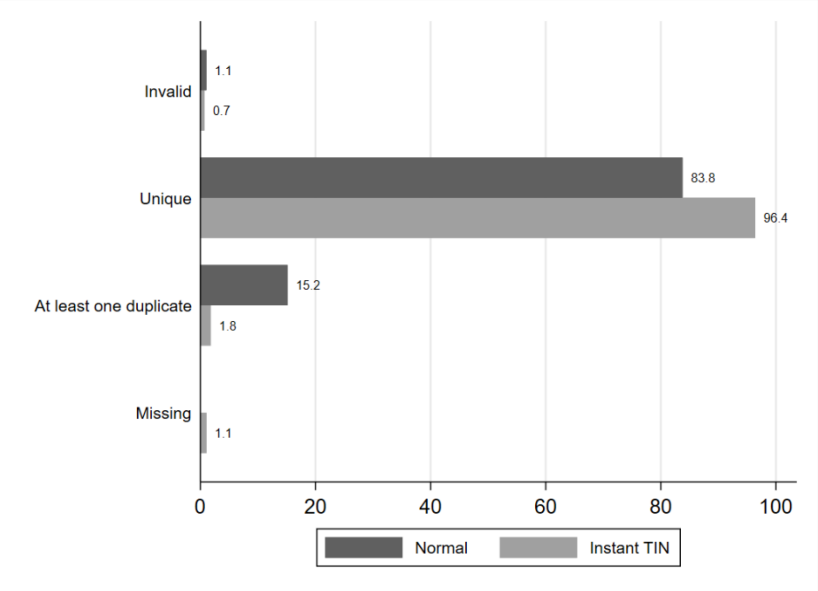
Source: Own calculations on URA administrative data.

Figure A2.6 Rate of email validity by type of registration



Source: Own calculations on URA administrative data.

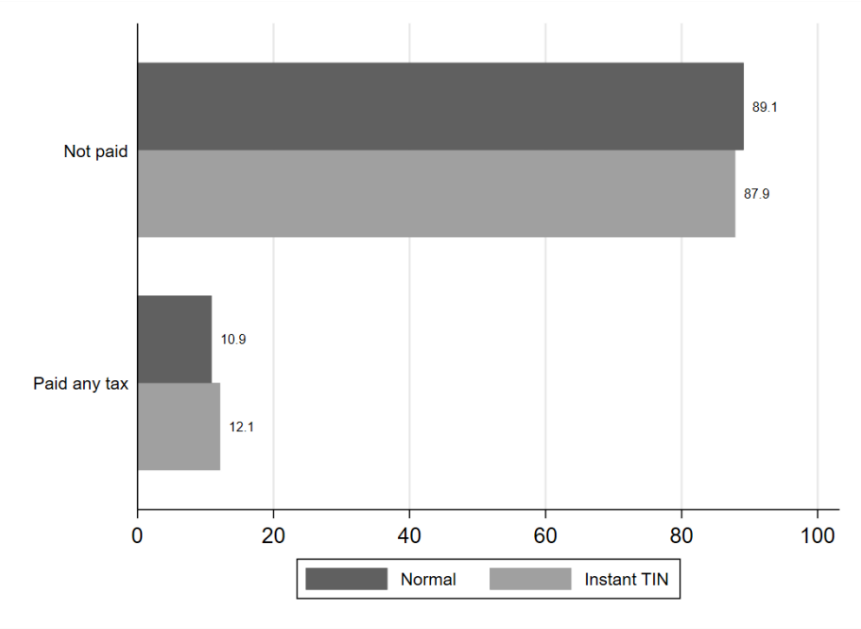
Figure A2.7 Rate of phone number validity by type of registration



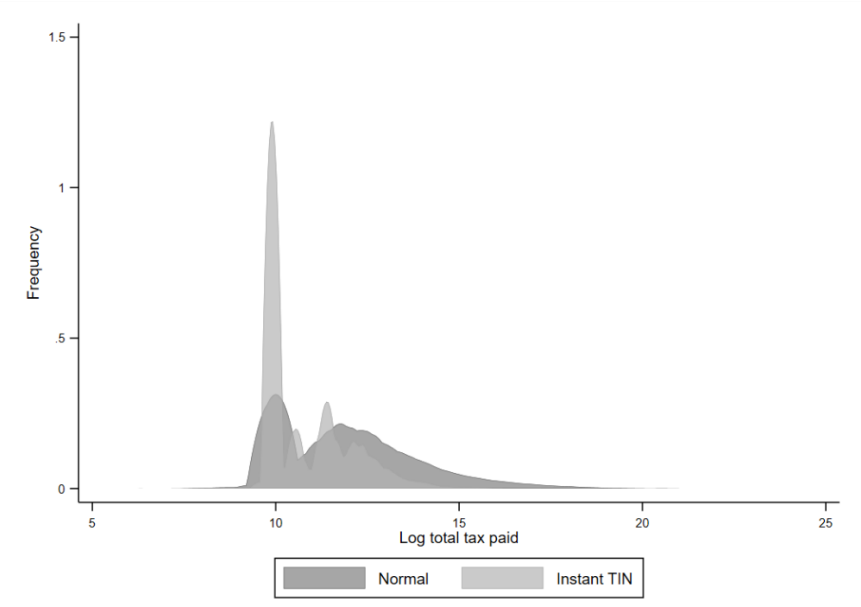
Source: Own calculations on URA administrative data.

Figure A2.8 Payment behaviour by registration type

(a) Share of taxpayers paying any tax at all in 2022



(b) Distribution of log total tax paid in 2022



Source: Own calculations on URA administrative data.
Note: In Figure A2.8a, paying any tax at all indicates whether the taxpayer made at least one tax payment in 2022 for any tax of the ten tax heads available in the data. In Figure A2.8b, the log total tax paid is the log transformation of total tax paid in 2022, in turn built as the sum of all payments made in 2022 for all ten tax heads available in the data.

References

Ali, M.; Shifa, A.B.; Shimeles, A. and Woldeyes, F. (2017) *Building Fiscal Capacity: The Role of ICT*, Working Paper Series No. 290, Abidjan, Côte d'Ivoire: African Development Bank Group

Benhassine, N.; McKenzie, D.; Pouliquen, V. and Santini, M. (2018) 'Does inducing informal firms to formalize make sense? Experimental evidence from Benin', *Journal of Public Economics* 157: 1–14, DOI: 10.1016/j.jpubeco.2017.11.004

Besley, T. and Persson, T. (2014) 'Why Do Developing Countries Tax So Little?', *Journal of Economic Perspectives*, 28(4), 99–120, DOI: 10.1257/jep.28.4.99

Besley, T. and Persson, T. (2013) *Taxation and Development*, *STICERD - Economic Organisation and Public Policy Discussion Papers Series* 041, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE

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

Bruhn, M. (2011) 'License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico', *Review of Economics and Statistics*: 93(1): 382-386

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

Carrillo, P.; Pomeranz, D. and Singhal, M. (2017) 'Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement', *American Economic Journal: Applied Economics* 9.2: 144–64

de Andradi, G.; Bruhn, M. and McKenzie, D. (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

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

de Mel, S.; McKenzie, D.; and Woodruff, C. (2013) 'The Demand for, and Consequences of, Formalization among Informal Firms in Sri Lanka', *American Economic Journal: Applied Economics* 5(2): 122-50

Fan, H.; Liu, Y.; Qian, N.; and Wen, J. (2018) *Computerizing VAT Invoices in China*, NBER Working Paper 24414, National Bureau of Economic Research

Gallien, M.; Moore, M. and van den Boogaard, V. (2021) *Taxing the Informal Economy is not a Silver Bullet for Financing Development – or the Covid-19 Recovery*, ICTD Summary Brief 24, Brighton: Institute of Development Studies

Hakizimana, N. and Santoro, F. (2023) *Technology Evolution and Tax Compliance: Evidence from Rwanda*, ICTD, African Tax Administration Paper 30, Brighton: Institute of Development Studies, DOI: 10.19088/ICTD.2023.033

Jensen, A. (2019) *Employment Structure and the Rise of the Modern Tax System*, NBER Working Paper 25502, National Bureau of Economic Research. DOI: 10.3386/w25502

Jouste, M.; Nalukwago, M.I. and Waiswa, R. (2021) *Do tax administrative interventions targeted at small businesses improve tax compliance and revenue collection? Evidence from Ugandan administrative tax data, WIDER Working Paper 2021/17,* UNU-WIDER, DOI: 10.35188/UNU-WIDER/2021/951-8

Kleven, H.J.; Knudsen, M.B.; Kreiner, C.T. and Saez, E. (2011) 'Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark', *Econometrica* 79(3): 651–692, DOI: 10.3982/ECTA9113

Kleven, H.J.; Kreiner, C.T. and Saez, E. (2016) 'Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries', *Economica*, 83(330): 219–246, DOI: 10.1111/ecca.12182

Lediga, C.; Riedel, N. and Strohmaier, K. (2020) *What You Do (and What You Don't) Get When Expanding the Net—Evidence from Forced Taxpayer Registrations in South Africa*, SSRN Scholarly Paper 3616808, Social Science Research Network, DOI: 10.2139/ssrn.3616808

Mascagni, G.; Santoro, F.; Mukama, D.; Karangwa, J. and Hakizimana, N. (2022) 'Active Ghosts: Nil-filing in Rwanda', *World Development* 152: 105806, DOI: 10.1016/j.worlddev.2021.105806

Mayega, J.; Waiswa, R.; Nabuyondo, J. and Nalukwago Isingoma, M. (2021) *How Clean Are Our Taxpayer Returns? Data Management in Uganda Revenue Authority*, ATAP Working Paper 24, Brighton: Institute of Development Studies, DOI: 10.19088/ICTD.2021.007 (accessed 19 March 2024)

Mayega, J.; Ssuuna, R.; Mubajje, M.; Nalukwago, M.I. and Muwonge, L. (2019) *How Clean is our Taxpayer Register? Data Management in the Uganda Revenue Authority*, ICTD, African Tax Administration Paper 12, Brighton: IDS

Medina, L. and Schneider, F. (2019) *Shedding Light on the Shadow economy: A Global Database and the Interaction with the Official One*, CESifo Working Paper No. 7981, SSRN Scholarly Paper ID 3502028

Mittal, S. and Mahajan, A. (2017) *Vat in Emerging Economies: Does Third Party Verification Matter?*, SSRN Scholarly Paper 3029963, DOI: 10.2139/ssrn.3029963

Moore, M. (2022) 'Tax obsessions: Taxpayer registration and the 'informal sector' in sub-Saharan Africa', *Development Policy Review* 41(1): e12649, DOI: 10.1111/dpr.12649

Naritomi, J. (2019) 'Consumers as Tax Auditors', *American Economic Review 109*(9): 3031–3072, DOI: 10.1257/aer.20160658

Occhiali, G.; Akol, D. and Kargbo, P. M. (2022) *ICT and Tax Administration in Sub-Saharan Africa: Adopting ITAS in Uganda and Sierra Leone*, ICTD Working Paper 148, Brighton: Institute of Development Studies, DOI: 10.19088/ICTD.2022.014 (accessed 19 March 2024)

OECD (2017) *Shining Light on the Shadow Economy: Opportunities and threats*, Organisation for Economic Co-operation and Development

Okunogbe, O. (2021) *Becoming Legible to the State: The Role of Detection and Enforcement Capacity in Tax Compliance*, Policy Research Working Papers, The World Bank, DOI: 10.1596/1813-9450-9852

Okunogbe, O. and Pouliquen, V. (2018) *Technology, Taxation, and Corruption: Evidence from the Introduction of Electronic Tax Filing*, Policy Research Working Papers, Washington DC: World Bank, DOI: 10.1596/1813-9450-8452

Okunogbe, O. and Santoro, F. (2022) 'The Promise and Limitations of Information Technology for Tax Mobilization', *The World Bank Research Observer* 38(2): 295-324, DOI: 10.1093/wbro/lkac008

Pomeranz, D. (2015) 'No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax', *American Economic Review*, 105(8): 2539–2569, DOI: 10.1257/aer.20130393

Rocha, R.; Rachter, L. and Ulyssea, G. (2014) *Do Entry Regulation and Taxes Hinder Firm Creation and Formalization? Evidence from Brazil*, Discussion Paper no. 11, Instituto de Economia, Universidade Federal Do Rio de Janeiro, Brazil

Santoro, F. and Waiswa, R. (2022) *Small Nets for Big Fish? Tax Enforcement on the Richest – Evidence from Uganda*, ICTD Working Paper 143, Brighton: Institute of Development Studies, DOI: 10.19088/ICTD.2022.009

Santoro, F. *et al*. (2023) *Technology and Tax: Adoption and Impacts of E-services in Rwanda*, ICTD Working Paper 153, Brighton: Institute of Development Studies, DOI: 10.19088/ICTD.2023.001 (accessed 19 March 2024)

Santoro, F.; Munoz, L.; Prichard, W. and Mascagi, G. (2022) *Digital Financial Services and Digital IDs: What Potential do They Have for Better Taxation in Africa?* ICTD Working Paper 137, Brighton: Institute of Development Studies, DOI: 10.19088/ICTD.2022.003, accessed 19 March 2024)

Slemrod, J.; Collins, B.; Hoopes, J.L.; Reck, D. and Sebastiani, M. (2017) 'Does credit-card information reporting improve small-business tax compliance?', *Journal of Public Economics* 149: 1–19, DOI: 10.1016/j.jpubeco.2017.02.010

UBOS (2019) *Statistical Abstract 2019*, Uganda Bureau of Statistics

Ulyssea, G. (2020) 'Informality: Causes and Consequences for Development', *Annual Review of Economics* 12(1): 525–546, DOI: 10.1146/annurev-economics-082119-121914

URSB (2018) *Annual Report 2017-2018*, Uganda Registration Services Bureau

Waseem, M. (2022) 'The Role of Withholding in the Self-Enforcement of a Value-Added Tax: Evidence from Pakistan', *The Review of Economics and Statistics* 104(2): 336–354, DOI: doi.org/10.1162/rest_a_00959

www.ictd.ac