The role of data in simplifying tax systems in Commonwealth countries

Fras Laura, Mittra Saugata and Siwale Twivwe

- There has been a huge drive to digitalise and simplify tax systems in Commonwealth countries. Simplified tax systems can reduce the compliance burden for taxpayers and improve efficiency for tax authorities.
- A key benefit of digitalisation is that it produces colossal volumes of data. Yet, for many tax authorities, this data remains underutilised.
- This brief argues that data utilisation should lie at the heart of revenue collection strategies.
- Effective data usage can enhance enforcement, detect non-compliance, generate evidence for policy decisions, and unleash modern tools such as machine learning.
- Establishing specialised data analytics units within tax authorities—such as Senegal's tax data lab—can improve data integration, risk assessment, and compliance monitoring.

POLICY BRIEF XXX-25049

FEBRUARY 2025

DIRECTED BY LSE UNIVERSITY OF OXFORD

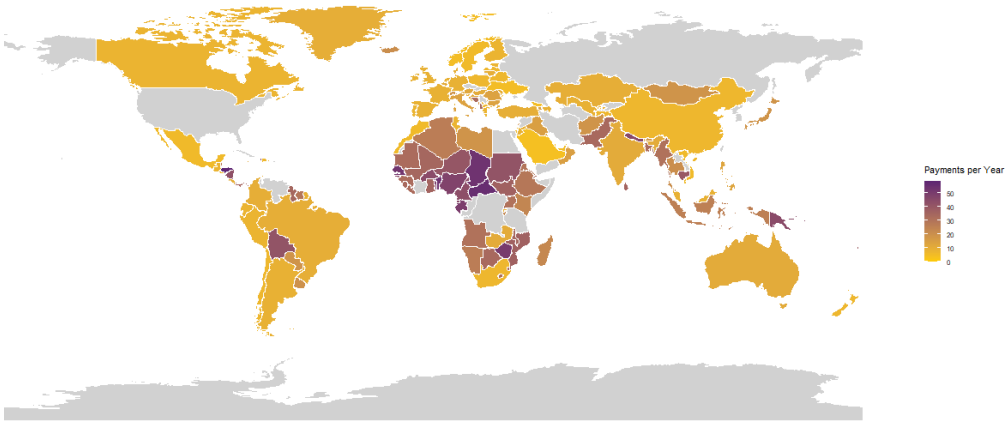
FUNDED BY UKaid from the British people

Introduction

Tax authorities in the Commonwealth have embraced digitalisation. Large investments have been made to overhaul entire tax systems. It is widely agreed that these digital reforms should benefit both the taxpayer and tax authority. One of the primary channels this should occur is through simplification. The goal of simplifying a tax system is to increase the ease of compliance, which is reflected in the costs and time saved.[1]

The process of paying taxes should be as easy as possible for the taxpayer. Regrettably, this is often not the case. In sub-Saharan Africa, 36.6 tax payments are made by businesses each year, compared to only 10 in OECD countries.[2] Beyond the number of payments, the time spent doing taxes is a further burden. On average, in South Asia, businesses spend 273.5 hours filing tax returns, which is the equivalent of 34 working days.[3] Figure 1 shows the distribution of tax payments made by businesses in 2019. There is still a lot of work to be done.

FIGURE 1: Number of tax payments that businesses make in a year.



Source: Created by IGC using World Bank data

For the tax authority, simplification means that enforcing and collecting taxes should be faster and more accurate. This can include using technology to make it easier to spot non-compliers, ideally those who owe large sums of taxes. As well as enabling more efficient and effective auditing methods.

The goal of simplifying a tax system is to increase the ease of compliance, which is reflected in the costs and time saved.

[1] International Finance Corporation, Multilateral Investment Guarantee Agency and World Bank (2009) *A Handbook for Tax Simplification*. Washington, DC: World Bank. Available at: http://hdl.handle.net/10986/28206 (Accessed: [05/02/2025]).

[2] World Bank (2019) Ease of Doing Business Rankings, The World Bank. Available at: https://archive.doingbusiness.org/en/rankings.

[3] World Bank (2019)

Many reports focus on the benefits of simplification for taxpayers. In this brief, we focus on the tax authority and how **data generated by technological adoption** can be used to simplify revenue collection for the tax authority. This is an area which is often overlooked as discussions are dominated by a focus on streamlining policies, rationalising processes and rolling out new technology. Given the underutilisation of data in many tax authorities, a data centric approach could be a game changer

A word about data

Data is the new oil. Like oil, it is a valuable resource that, if unlocked, can produce significant value. For tax authorities, this represents a unique opportunity. The process of digitalisation and technology adoption **produces large volumes of data**. Most tax authorities in the Commonwealth and beyond have bought into the idea that technology use is the future of tax administration. They have invested heavily in processes such as online registration, e-filing of returns, use of Electronic Billing Machines (EBMs) and, increasingly, the use of digital IDs. The data generated by these systems can lead to further simplification of tax systems, if utilised well.

The virtuous cycle of data

High-quality data is both a key output and input for higher tax capacity.[4] It is a key output in that it enables the tax authority to achieve better processes at scale that would be too costly to do manually. It is also an input as it feeds into other tax administration processes. Data quality is an important element in this discussion. **The higher the quality of data, the better the outcomes**. Data quality depends on several factors, including the level of investment in technology infrastructure. How accessible is the technology? How easy is it to use? It also relies heavily on skilled human capital. How well attuned staff are to using digital systems. **The staff at tax authorities need to think of themselves as not just users of data but producers.**

[4] Bachas, P. and Jensen, A. (2023) 'Administrative data and methodologies for tax policy', *International Growth Centre*, 3 April. Available at: https://www.theigc.org/publications/administrative-data-and-methodologies-tax-policy (Accessed: 4 February 2025).

Three ways data and automation can simplify tax processes

1. Creating data inputs for other processes

Automation leads to simplification by creating data that serves as input for other processes. A great example of this is the use of **Electronic Billing Machines (**EBMs) and **smart invoicing**. EBMs are machines that transmit sales transactions directly to the revenue authority in real time. For taxes like VAT, a return detailing all the transactions for that period is typically due at the end of the tax period. If the EBM system is linked to the e-filing system, the return can be prefilled for the taxpayer with the data collected by the EBM. This saves the taxpayer time and improves the accuracy of the information in the returns. There are many other applications where this can be extended.

2. Creating links across disparate data systems, allowing for cross-checking and the creation of data dashboards

Tax authorities typically have data systems for different tax types and associated processes. The typical tax authority will have an e-filing system, a customs processing system and a separate tax registry. These systems are often not integrated. One way of unlocking the power of the data contained in these systems is by creating linkages and cross-checking them for patterns and discrepancies, which algorithms can do in minutes. This can be extremely effective in catching non-compliance and uncollected revenue and may be impossible for a human to do manually. In Uganda, researchers did a cross-checking exercise comparing VAT input and output schedules. They found out that by not carrying out these cross-checks, the Uganda Revenue Authority potentially **lost at least US$ 383 million in uncollected revenue between 2013 and 2016.**[5] A systematised approach to this is the creation of dashboards that display key information in a simple and accessible way. Dashboards can be customised for different functions, such as sector dashboards to understand industry compliance behaviour, or specific to data from a single system, such as an EBM dashboard. Dashboards simplify the analysis and display of large amounts of data.

[5] Almunia, M, J Hjort, J and L Tian (2021), Strategic or Confused Firms? Evidence from 'missing' transactions in Uganda, Review of Economics and Statistics

Tanzania Electronic Fiscal Device (EFD) dashboard

The Tanzanian Revenue Authority (TRA), working with the IGC, used Electronic Fiscal Device (EFD) data to create a dashboard that processed and displayed EFD data in an accessible way for tax authority staff.

The dashboard includes information allowing for regional, EFD machine and taxpayer analysis of data. The dashboard gives the TRA the opportunity to investigate total sales and receipts issued and analyse the data through a range of filters, including time windows and frequencies.

The TRA can use the dashboard to answer highly relevant questions such as which taxpayers' receipt information is outside the normal range and whether there are specific patterns for issuing invoices.

The TRA is working on integrating data systems so that they can create dashboards that include data from personal income tax, customs tax and taxpayer registration information.

An application of cross-checking with huge potential is data collected by other organisations, known as third-party information. Third-party reporting has historically played an important role in improving tax capacity in developed countries.[6] The ability of a tax authority to verify the information provided by other economic agents on taxpayers and the perception of that ability is extremely important for compliance purposes and is a form of third-party reporting. For example, in South Africa the Revenue Service accessed the country's commercial registry which is managed by the Companies and Intellectual Property Commission. They discovered approximately **300,000 firms that were required to register for tax purposes who were not registered.**[7] This type of linkage can be extended to banks, property registries. Digital IDs may present a unique opportunity for this type of cross-checking to happen at scale across all government agencies. The process of leveraging Digital Public Infrastructure for tax purposes is an important one that should be explored.

[6] Pomeranz, D., 2015. No taxation without information: Deterrence and self-enforcement in the value added tax. American Economic Review, 105(8), pp.2539-2569.
[7] Lediga, C, N Riedel and K Strohmaier (2020), What you do (and what you don't) get when expanding the net: Evidence from forced taxpayer registrations in South Africa, 113th Annual Conference on Taxation, National Tax Association.

3. Large tax data systems can simplify auditing by enabling the use of Artificial Intelligence (AI)

Large high-quality datasets enable the use of new digital tools such as machine learning. Machine learning is a form of AI that allows machines to learn and perform tasks automatically.[8] This is an application of cross-checking, as highlighted in point two above. These are tasks that would be impossible for a human to perform manually. This simplifies a process that was previously out of reach or that was previously done in a limited way. A great application of this is in **auditing and risk scoring**. Machine learning can help tax authorities select firms for enforcement activities like audits in a fairer way by making the process more data driven.

In Delhi, a team of researchers used data from previous audits to train an algorithm to identify ghost firms.[9] Ghost firms commit tax fraud by issuing fake receipts and claiming VAT on goods and services that they have not provided. The algorithm was able to sort through large quantities of data and come up with a set of characteristics that are typical of ghost firms. These firms were then selected for audit, which resulted in the collection of US$30 million more in revenue. The audit results were fed back into the system to improve future predictions. The system was more **accurate than manual scoring techniques**.

The use of machine learning and related technology is one way to simplify a process that humans would struggle to do manually and use a data-driven approach that makes audit selection fairer.

Policy Recommendation: Prioritise the creation of data analytics units

The creation of a data analytics unit is an effective way to operationalise the various ways of using data that have been highlighted. The Research, Applied Analytics, and Statistics (RAAS) division at the US Internal Revenue Service and the Knowledge, Analysis and Intelligence group at the UK Her Majesty's Revenue and Customs are examples of these units in High-Income Countries.

The role of the data analytics unit would be to produce user-friendly final datasets which are harmonised, contain only the needed information for each use case, and have been de-identified. These resulting datasets can then be used for budget projections, tax policy analysis, and economic predictions.[10]

[8] Russell, S.J. and Norvig, P., 2016. Artificial intelligence: a modern approach. Pearson.
[9] Shekhar Mittal, Ofir Reich, and Aprajit Mahajan. 2018. Who is Bogus? Using One-Sided Labels to Identify Fraudulent Firms from Tax Returns. https://doi.org/10.1145/3209811.3209824
[10] Bachas, P. and Jensen, A. (2023)

Many tax authorities in the Commonwealth have already established data analytics units. However, the approach in many countries must incorporate more data use cases. In some countries, there needs to be a complete step change. Many tax authorities still think about tax collection in an analogue way and then overlay a digital process onto it. There needs to be a seismic shift to view the entire process of revenue collection through a data lens. This will change many things in fundamental ways including the skills requirements of staff.

There needs to be a seismic shift to view the entire revenue collection process through a data lens.

A promising data analytics unit is the Senegal data lab highlighted below.

Senegal's data lab

In 2015, the Senegalese Tax Administration (DGID) collaborated with researchers on a series of projects, which led to the creation of a tax data lab.[11]

Aim of the lab: A dedicated unit of six data scientists in the lab work to digitalise data, improve audit selection, detect tax evasion by firms, and enforce property tax. A further aim was to create standardised practises to digitalise manually filed information.

How has the data lab improved outcomes?

- It produced a comprehensive guide to navigate the complex information environment and suggested improvements, which were relayed to the IT departments and local tax offices.

- It revealed data anomalies for which data-checking solutions will progressively be implemented.

It increased the administration's capacity to analyse and cross-check tax declarations and consequently detected non-compliant behaviours on a large scale.

One of the key inputs into the data lab's work is talented data scientists who can maximise the use of data.

[11] Bachas, P. and Jensen, A. (2023)

Challenges remain

- In some countries, some parts of the data ecosystem are still **non-**digitalised. Digitalising manually collected data and harmonising with already digital data can be cumbersome. If manual reporting is incomplete, inconsistent, or unclear digitalising data cannot solve the problem of poor quality in tax returns.

- Many tax offices and departments may be collecting similar data but may not be aware of each other's existence. Even if all the possible sources of data are identified, it may still be challenging to combine data sets if they use different taxpayer IDs.

- Tax authorities need to use data and evidence teams to consider what strategies are most appropriate to resolve the specific data challenges they have.

- Tax authorities need to train and recruit staff who are data natives.

Conclusion

Digitalisation can transform tax compliance and enforcement by enabling a data-driven approach. This can lead to simplifying processes that make tax compliance easier and fairer for taxpayers, such as prefilling tax returns. It can also allow for cross-checking and verification at a scale and level of accuracy that is unachievable with traditional systems. Large data systems also facilitate the use of modern tools, such as machine learning, that can improve the detection of non-compliance. Tax authorities can facilitate these gains by creating data analytics units and putting them at the centre of their enforcement strategies, being mindful of the challenges of building the expertise and systems needed to ensure data quality.

References

Almunia, M., Hjort, J. and Tian, L. (2021) Strategic or Confused Firms? Evidence from 'missing' transactions in Uganda. Review of Economics and Statistics.

Bachas, P. and Jensen, A. (2023) 'Administrative data and methodologies for tax policy,' International Growth Centre, 3 April. Available at: https://www.theigc.org/publications/administrative-data-and-methodologies-tax-policy (Accessed: 4 February 2025).

International Finance Corporation, Multilateral Investment Guarantee Agency and World Bank (2009) A Handbook for Tax Simplification. Washington, DC: World Bank. Available at: http://hdl.handle.net/10986/28206 (Accessed: [02/02/2025]).

Lediga, C., Riedel, N. and Strohmaier, K. (2020) 'What you do (and don't) get when expanding the net: Evidence from forced taxpayer registrations in South Africa', 113th Annual Conference on Taxation, National Tax Association.

Mittal, S., Reich, O. and Mahajan, A. (2018) 'Who is Bogus? Using One-Sided Labels to Identify Fraudulent Firms from Tax Returns'. Available at: https://doi.org/10.1145/3209811.3209824 (Accessed: [insert date]).

Pomeranz, D. (2015) 'No taxation without information: Deterrence and self-enforcement in the value added tax', American Economic Review, 105(8), pp. 2539-2569.

Russell, S.J. and Norvig, P. (2016) Artificial intelligence: a modern approach. Pearson.

World Bank (2019) 'Ease of Doing Business Rankings', The World Bank. Available at: https://archive.doingbusiness.org/en/rankings (Accessed: [05/02/2025]).