E-invoicing, Tax Audits and VAT Compliance

BSE Working Paper 1454| July 2024

Christos Kotsogiannis, Luca Salvadori, John Karangwa, Innocente Murasi

bse.eu/research

E-invoicing, Tax Audits and VAT Compliance

Christos Kotsogiannis[a], Luca Salvadori[b,†], John Karangwa[c], Innocente Murasi[c]

This version: 24th July 2024

Abstract: Difficult to find another policy shift that has promised as much for tax compliance in developing countries as digitalization. Yet the evidence on its impact is scant. Using the universe of tax filings in Rwanda over the period 2012-2019, this paper investigates the extent to which digitalization (in the form of e-invoicing) has impacted on VAT compliance and, in particular, the effectiveness of tax audits. The evidence suggests that while e-invoicing adoption per se has increased firms' net VAT payments, this impact is quantitatively limited, as firms seem to re-adjust their expenses so to keep VAT payments low. Interestingly, e-invoicing had a sizable compliance impact on net VAT liabilities reported by audited firms, with this impact being attributed to tax audits becoming more efficient, rather than to VAT registered firms becoming more cautious following their participation in the e-invoicing mechanism.

Keywords: Tax Audit Evaluation; Technological Change; Digitalisation Initiatives; Tax Administration; Tax Evasion; Tax Compliance.

JEL classification: H25, H26, H32, O17, O33, D02, D22.

[a]Tax Administration Research Centre (TARC), Department of Economics, University of Exeter Business School, United Kingdom; and CESIfo, Germany.

[b]Autonomous University of Barcelona (UAB), Barcelona School of Economics (BSE), Spain; Tax Administration Research Centre - TARC (University of Exeter), United Kingdom; and Barcelona Institute of Economics (IEB), Spain.

[c]Planning and Research Department, Rwanda Revenue Authority, Rwanda.

[†]Corresponding author. Email: luca.salvadori@uab.cat.

1 Introduction

To strengthen the capacity of tax collection,[1] enhance the overall performance and quality of services provided to individuals and firms, and ultimately make taxation more growth-friendly, African tax administrations have begun embracing digitalization (ATAF, 2021a). Deploying modern technology enabling more accurate (and formal) linkages between sellers and purchasers —typically taking the form of electronic invoicing (e-invoicing)[2] through which every sales transaction, both to other businesses and to final consumers, is recorded—can help businesses improve record keeping, enhancing revenue protection as well as protecting honest taxpayers from unfair competition. It can also contribute to tax compliance through the digital audit trail that these transactions enable, creating information that can be readily available to verify the accuracy of tax declarations at both the pre-audit and post-audit stage.[3] Perhaps not surprisingly—given the importance in total revenues of the Value Added Tax (VAT)—the integration of digital innovations, and e-invoicing in particular, has been more prominent in VAT.[4]

Despite, however, the importance of this issue for revenue mobilization in developing countries, the evidence of the impact of such policy innovation on firms' compliance (reviewed shortly below) is relatively unexplored and mixed. The use of technology, such as the e-invoicing, makes it possible for tax administrators to systematically verify the accuracy of taxpayers' records and perform cross-checks between trading partners. In this context, audits can enhance the potential compliance impact of technological innovations. It is this issue that the paper deals with. More specifically, the aim of this paper is to evaluate the compliance impact – through audit enhancement – of the introduction of e-invoicing in Rwanda, which allowed for the automated transfer of billing information between firms and the tax authority and the creation of a digital trail. Undoubtedly, the challenges for such a reform are daunting and profound (let alone for a developing country), for both taxpayers and the tax administration.

[1] At an average of around 15 percent, tax ratios in developing countries remain low. In about half of those countries, the tax ratio is below the 15 percent widely considered as the minimal level for which there is some prospect of them meeting their urgent needs. Achieving the Sustainable Development Goals has been estimated to require, on average, additional revenue of around 15 percent of GDP, with this calculation being made before the revenue setback imposed by COVID-19. Since the onset of the pandemic general government debt in developing countries has risen by about 5 points of GDP.

[2] More than fifty countries around the world have already adopted e-invoicing (Barreix et al., 2018), and the number is growing (OECD, 2022).

[3] For example, for VAT refunds (pre-audit) and for identifying fake invoices issued on purchases of inputs (post-audit).

[4] VAT contributes a substantial portion of tax revenues in Africa, accounting for, on average, over 35 percent of the total tax revenue collected, ATAF (2021b). VAT adoption has been one of the most significant recent development in tax policy. While only 47 countries adopted VAT in 1990, 170 countries embraced this tax in 2020, OECD et al. (2020).

Making use of the universe of administrative tax data on VAT filings and tax audit records in Rwanda, as well as information on the adoption of the e-invoicing by firms over the period spanning the years 2012-2019, this paper shows that e-invoicing adoption per-se yields a limited increase in firms' net VAT payments. The bulk of the positive impact on compliance stems from tax audits becoming more effective in deterring tax evasion in the post-audit period. The results also show that audits involving firms adopting the e-invoicing system are the only ones that yield a significant compliance improvement in net VAT payments, highlighting the importance of e-invoicing for VAT audit enhancement. Finally, the results also suggest that taxpayers react to tax enforcement or to the adoption of e-invoicing, by over-reporting their input VAT: the purpose of doing so is to reduce their final liability. This pattern is nuanced and reverted when tax audits are strengthened by the use of e-invoicing (see more details in Section 5).

The reminder of the paper is organized as follows. Section 2 briefly reviews the literature. Section 3 provides some institutional background, discusses the mechanisms originating the synergies between tax audits and e-invoicing, and describes the data the analysis is based on. Section 4 describes the methodological approach followed, and Section 5 presents the results, while Section 6 provides some concluding remarks.

2 Literature review

The evidence of the introduction of e-invoicing on compliance is mixed. Using cross-country data, Casey and Castro (2011) find that the introduction of e-invoicing alone is generally not associated with significant increases in VAT revenue collection or permanent compliance improvements. The reason for this is that effective implementation of e-invoicing requires sufficient interoperability with systems that support tax audits and, more broadly, the tax compliance function. Indeed, while the implementation of an e-invoicing technology is potentially enabling tax administrations to strengthen the integrity of their processes and facilitate efficient audits, this is not necessarily achieved in the absence of tax capacity. More recently, Mascagni et al. (2021) show that the introduction of e-invoicing in Ethiopia increased both VAT revenues and income tax revenues. But, they also find that e-invoicing users tend to increase their expenses reported in order to try to mitigate their tax liability. In two recent contributions—focusing on Rwanda and Uganda, respectively—Mascagni et al. (2019) and Almunia et al. (2022)[5] show that electronic submission of information to tax authorities does not necessarily lead to more compliance. The reason for this is that, despite the fact that transactions are digitalised,

[5]See also the discussion in Section 3.2.

there are discrepancies in transactions recorded, driven by the lack of effectiveness in cross-checking sales and purchases.[6] Earlier work on the initial adoption of e-invoicing, which used data from Ethiopia and Rwanda, suggests that the pilot implementation of this innovation led to a significant increase in VAT revenues after the first year of adoption in these countries (Ali et al., 2021; Eissa and Zeitlin, 2014). More recently, Bellon et al. (2022) examine the impact of switching from paper to electronic invoicing on firm tax compliance and performance using quasi-experimental variation in the roll-out of VAT e-invoicing in Peru. They find that e-invoicing increases reported firm sales, purchases and VAT liabilities by over 5 percent in the first year after adoption.[7] This paper relates to all these contributions, but it is also distinctively different in the emphasis, as it discusses the channel of impact of electronic invoicing on future tax compliance through audits.

In this light, this paper also contributes to the literature investigating the compliance impact of operational audits. Despite the extensive evidence on the deterrence impact of audits in the developed world (see, for example, among others, Kleven et al., 2011; Gemmell and Ratto, 2012; DeBacker et al., 2015, 2018; Advani et al., 2021) the issue of evaluating tax audits in developing countries is somewhat neglected. Notable exceptions are the recent contributions by Best et al. (2021) and Kotsogiannis et al. (2024). Best et al. (2021) focus on Pakistan and find that although VAT audits uncover a substantial amount of evasion, they do not deter future noncompliance. These authors suggest that, given that these inspections tend to focus on checking mechanical violations, they are unlikely to move firm priors on the detection probability upwards. Kotsogiannis et al. (2024) analyse Corporate Income Tax (CIT) audits in Rwanda and provide evidence of an aggregate positive impact in terms of future compliance, an effect that is completely driven by comprehensive audits, with audits that are narrower in their scope (desk and issue audits) delivering a net counter-deterrence effect.

The evidence across these two strands of the literature briefly discussed in the preceding paragraphs, suggests that, depending on the context, technological innovations and tax audits may not be fully exploited as tax enforcement instruments. Despite being recognized as a theoretical possibility (in the discussion of the paper trail perspective of digitization), these issues have attracted, to the best of our knowledge, limited empirical

[6] One reason for the poor revenue performance of many developing countries is that their tax administrations lack effectiveness in administering their core functions which in turn reflects weaknesses not only in processes and procedures, including in the use of available technologies, but also in governance arrangements and in quality and quantity of the resources available. A point emphasized in, among others, Acemoglu et al. (2001), Besley and Persson (2009), and Besley et al. (2013).

[7] There are also some other papers on this issue and on the impact of digital technologies on tax administration, including not only e-invoicing (Ramírez and Oliva, 2018; Bérgolo et al., 2018; Templado and Artana, 2018; Castro et al., 2016; Lee, 2016), but also the electronic submission of tax returns or e-filing (Yilmaz and Coolidge, 2013; Kochanova et al., 2016; Okunogbe and Pouliquen, 2022).

attention in the literature.

3 Timing of the E-invoicing Reform, Mechanisms and Data

Rwanda is a representative low-income country both in terms of fiscal capacity and tax structure (ATAF, 2021b).[8] To improve efficiency in VAT collection and strengthen compliance, Rwanda made e-invoicing mandatory for VAT registered taxpayers in 2013. E-invoicing meant that a firm could issue certified VAT electronic invoices recording every sales transaction, both to other businesses and to final consumers, and submitting them electronically directly to the Rwanda Revenue Authority (RRA).

3.1 Timing of the reform

It was recognized that switching to e-invoicing would create significant costs for firms and the tax administration,[9] including updating IT capacity and staff training. For this, e-invoicing was introduced in a gradual and staggered manner, with large businesses and firms in specific sectors (the largest issuers of invoices) required to adopt e-invoicing first. During the first period of adoption e-invoicing expanded rapidly and, by September 2014, there were over 3,943 taxpaying firms actively using e-invoices, which corresponds to 77.8 percent of all VAT-registered firms at that time.[10]

In January 2014, RRA announced the deadline for the adoption of e-invoicing by all VAT registered taxpayers (by end of March of the same year). During this phase there was increased enforcement effort by RRA and the number of e-invoicing adoptions increased significantly in the same year, one year after the launch. Despite announcing the deadline for the adoption of e-invoicing, some taxpayers did not join the system, while others started adopting it in later years. There were also issues with the quality of receipts

[8] VAT is the main source of revenue: in 2017/2018 the VAT contributed 33 percent of total domestic revenue, followed by the employment income tax under the Pay as You Earn (PAYE) scheme (23 percent), business income taxes (CIT and PIT, 19 percent), excise tax (12 percent), and import duty (8 percent) (RRA, 2019). Further details on the Rwandan tax system see, for example, Tourek (2022), Mascagni et al. (2022), Mascagni et al. (2023), and Kotsogiannis et al. (2024).

[9] During the transition phase e-invoicing presented a number of challenges, mainly consisting in practical barriers, and inconveniences that taxpayers experienced, often with negative repercussions on their tax morale, perceptions and attitudes (Mascagni et al., 2023). It was also significantly costly, especially for small firms, which had to cover the purchase and maintenance costs over time. The upfront cost in installation required four full years of VAT payments to be recovered (Eissa and Zeitlin, 2014). It was also not possible, for example, for RRA to track taxpayers' inventory and specific details of the items sold. In addition the RRA could not provide remote online support to taxpayers and monitor the status of the machines.

[10] E-invoicing was also expected to enhance the RRA capacity to monitor firm transactions. The evidence for the first stage of the introduction of e-invoicing is that is increased VAT payments by 6.5 percent between March 2013 and September 2014 (Eissa and Zeitlin, 2014).

that used to easily deteriorate, which posed challenges to RRA with record-keeping and verification in audits, as well as limitations in the information that the electronic billing machines could store. These challenges prompted the need for an upgrade to the system, which resulted in the introduction of an upgraded e-invoicing system in March 2017, which aimed at increasing the adoption and use of e-invoices by making it more accessible to all eligible taxpayers. Indeed, unlike the previous requirement for firms to purchase a device, this system is provided as software, which the RRA installs free of charge. The implementation of the second phase in e-invoicing followed a similar staggered approach to that of the initial transition, in which a sample of large businesses were first requested to switch from the initial phase to the second phase. Later in 2020 all VAT registered businesses were requested to switch to the latest e-invoicing system.

3.2 E-invoicing and Tax Audits

As with almost all VAT systems, Rwanda follows the credit-invoice system which allows for the deduction of VAT already paid at each stage of production.[11] This sequence of transactions creates a paper trail of records that show the flow of goods and services together with the associated cost of VAT transactions. This auditable paper trail along the production chain serves as a self-enforcement mechanism (Pomeranz, 2015; Naritomi, 2019).

As alluded to in the introductory section, an e-invoicing system strengthens compliance through several channels.[12] First, through providing revenue agencies with real-time transaction information from firms, e-invoicing can play a crucial role in better defining tax liabilities. Through collecting and processing large amounts of data provided through electronic invoicing and tax declarations, revenue authorities can also automatically detect inconsistencies, such as mismatches between self-reported and third party–reported tax liability. Second, e-invoicing could also play a major role in terms of reducing firms' compliance costs. Finally, by providing reliable information flows on firms' transactions to the revenue authority, an e-invoicing system is expected to reinforce its VAT enforcement capacity, enhancing VAT audit effectiveness.

The success of e-invoicing depends on its use. While firms are expected to mechanically increase the amount of total sales reported after the adoption of e-invoicing, in the absence of automatic pre-filling of taxpayer information, they also have an incentive to over-report other margins so to minimize the VAT paid. In this regard, the introduction

[11]The idea of the credit-invoice system is that firms issue invoices for any sale. When business to business transactions occur, firms can claim the VAT charged on input purchases as taxes already paid and are only required to remit the "value added" on goods sold (sales - business to business purchases).

[12]For a detailed discussion on this see Okunogbe and Santoro (2023).

of e-invoicing in Rwanda had limitations. Indeed, while it allows for the information to be transmitted in real time to the revenue authority, the system is not designed to be used for pre-filling of VAT returns, allowing firms to manipulate their purchases by inflating the VAT paid on the inputs. This is a behaviour that has been documented for Ethiopia (see Mascagni et al., 2021). Of course, there are other margins that can be subject to manipulation, including the misclassification of goods to exempted and zero-rated. One contribution of this paper is to shed some light on these margins too highlighting the role of audits in mitigating this behaviour (see Sections 4 and 5). What this all points to is that the operational significance of e-invoicing might diminish if not tackled through appropriate additional enforcement policies. The aim of this paper is to evaluate the role played by e-invoicing as a mean of a digital trail in indirectly enhancing compliance by improving the quality and performance of risk-based tax audits.

3.3 Data: VAT and Tax Audits

All data employed in this paper is at the taxpayer (business) level. They include mostly financial variables used to calculate taxes (for example, total sales, exempt items, VAT paid on inputs), as well as some taxpayer characteristics, such as size as defined by RRA,[13] geographical location (at tax centre level), and the information on the date in which firms adopted e-invoicing. VAT declarations have been annualised and merged with the detailed records of audits undertaken by the RRA during the years 2013 through 2017.

The RRA tends to audit two tax periods but taxpayers are required to keep their records for a longer period. Tax enforcement examinations involve three types of audits:[14] desk audits, issue audits and comprehensive audits. Comprehensive audits are in-depth and time-intensive examinations and usually are conducted through RRA staff visiting the taxpayer's business premises in order to review all relevant documents. Desk and issue audits are narrower in their scope, generally focusing on a single aspect and single tax period and are conducted by RRA staff using information already submitted to RRA through various sources including from the tax declarations.[15] We have also been given access to the detailed confidential information on the criteria for audit selection which includes the risk rules employed to assign risk scores to the world of tax declarations. The

[13]RRA classifies businesses as follows: *Micro*-businesses declare a turnover of less than 12 million Rwf (USD 13,380 as of February 2019 exchange rate) in a tax period; *Small*-businesses have a turnover between Rwf 12 million and Rwf 50 million (USD 55,750) in a tax period; *Medium* and *Large*-businesses have a turnover above that threshold.

[14]Following an administrative procedure RRA may also amend submitted tax liability which is initiated when the tax administration discovers a miscalculation or omission, an understatement or any other error in which case the tax administration rectifies the submitted tax liability. These amendments are not considered audits and therefore they do not appear in the analysis.

[15]For more details on the audit process, see Kotsogiannis et al. (2024).

risk criteria utilise information that spans across tax bases.[16] The administrative data is retrieved from RRA systems which collect and store tax data from tax procedures followed by taxpayers. The resulting dataset consists of a panel of firms over the period 2012-2019. Before performing our analysis we select a sample of relatively more homogeneous firms. Concretely, we select firms filing a VAT declaration for all years during the observed period and, in doing so, we end up with a perfectly balanced panel of 4,897 firms.

For this sample, Table 1 presents the summary statistics for our outcome variables, which are the annual aggregations of the correspondent fields of VAT declarations. More precisely, *VAT payable* represents the total output VAT that is, the sum of VAT charged on Taxable Sales and VAT Reverse charge. *VAT paid on inputs* is the sum of all input components of VAT (VAT paid on imports and VAT paid on local purchases); while *Total non-taxable sales* aggregates exempted sales, zero-rated sales and exports. Our main margin of interest is the net VAT liability annually paid by firms, that is *VAT due*, which is obtained by subtracting *VAT paid on inputs* from *VAT payable*. All variables are expressed in thousands of US$.

Table 1: Summary statistics - outcome variables (2012-2019)

Variable	Measurement Unit	Observations	Mean	Std. Dev.
VAT Due	1,000 US $	39,176	31.71	447.11
VAT Payable	1,000 US $	39,176	81.33	646.55
VAT Paid on Inputs	1,000 US $	39,176	70.94	405.36
Total Non-Taxable Sales	1,000 US $	39,176	359.5	3,850.74

Note: Authors' calculations based on data provided by RRA.

For our selected sample, Panel A of Table 2 reports some descriptive statistics for the detection performance of the tax audits conducted by RRA for all tax periods audited and available audit waves. In particular, *audit outcome* represents the amount of tax base underreported uncovered with a mean of just over US$ 40,000 and the standard deviation of about US$ 503,000. *Audit outcome* is also reported as a share of the *potential tax base* (defined as the sum of tax base declared by the taxpayer and the *audit outcome*). *Total fines*, which gives the sum of all fines and penalties applied to those businesses found underreporting tax bases, has a mean of almost US$ 23,000 and standard deviation of approximately US$ 312,000. *Total audit outcome* gives the sum of *audit outcome* and *total fines*. Finally, *total audit outcome (%)* is calculated as the percentage of *total audit outcome* over the *potential tax base including fines* (defined as the sum of taxable income

[16] After each return is filed, audit flags are deterministically generated based on the characteristics of the returns. Tax auditors conduct audits by following the procedures outlined in the audit manual, ensuring a systematic and consistent approach to the tax audit process. The integrity of the tax and audit data is assured by the RRA.

declared by the taxpayer and *total audit outcome* as to include tax fines). Thus, Panel A of Table 2 reveals that audits contribute a substantial amount of tax revenues in terms of uncovered tax bases underreported which amounts to about 28 percent of the potential tax base audited (about 31 percent including fines).

Table 2: Audits descriptive statistics (2013-2017)

Variable	Measurement Unit	Observations	Mean	Std. Dev.
Panel A: Total				
Audit outcome	1,000 US $	1,788	40.59	503.66
Audit Outcome (%)	% Potential Tax Base	1,788	28.48	41.61
Total Fines	1,000 US $	1,788	22.64	311.97
Total Audit Outcome	1,000 US $	1,788	63.32	811.95
Total Audit Outcome (%)	% Potential tax base (Including Fines)	1,788	31.56	43.01
Panel B: E-invoicing adopters				
Audit outcome	1,000 US $	960	48.45	665.91
Audit Outcome (%)	% Potential Tax Base	960	28.15	42.2
Total Fines	1000 US $	960	28.10	411.85
Total Audit Outcome	1,000 US $	960	76.56	1,074.18
Total Audit Outcome (%)	% Potential Tax Base (Including Fines)	960	30.99	43.56
Panel C: E-invoicing non adopters				
Audit outcome	1,000 US $	828	31.47	183.78
Audit Outcome (%)	% Potential Tax Base	828	28.87	40.93
Total fines	1,000 US $	828	16.31	116.37
Total Audit Outcome	1,000 US $	828	47.96	293.38
Total Audit Outcome (%)	% Potential Tax Base (Including Fines)	828	32.21	42.37

Note: Authors' calculations based on data provided by RRA.

Panel B and C of Table 2 present the same information by splitting the sample of audited firms between adopters and non-adopters of e-invoicing. In terms of their performance evaluated in levels, audits involving firms e-invoicing uncover a higher amount of underreporting and levy a larger amount of fines, leading to a greater total audit outcome when compared to the others. Nevertheless, when evaluated in relative terms, as a percentage of the potential tax base, audits involving non-adopters of e-invoicing tend to slightly outperform audits targeting users of e-invoicing. This means that the use of e-invoicing leads to an increase of the declared tax bases—the main component of the denominator of this ratio—that is more than proportional to the increase in underreporting detected plus fines levied. This is also true when comparing firms of the same size (see Apprendix A).[17] Read in this way, adoption of e-invoicing seems to improve compliance

[17]It would be interesting to look at firms' industry as an additional source of heterogeneity, but as discussed in Kotsogiannis et al. (2024), in our case there are substantial limitations in these dimensions due to missing data and data reliability.

by increasing the declared tax base and by enhancing the detection power of audits.

The next section presents the methodology employed to estimate the joint effect of e-invoicing and audits on the future reporting behaviour of taxpayers, that is the deterrence impact of these two enforcement instruments which is at the heart of the contribution of this paper.

4 Estimation Strategy

As discussed in Section 3.2, the primary aim of this paper is to estimate the impact of e-invoicing on promoting VAT compliance through tax audits enhancement. While our focus is on the net VAT liability reported by taxpayers as outcome variable, we also analyse different margins of VAT registered firms' reporting behaviour with the aim of decomposing their estimated net effect on the annual *VAT due* and disentangle the treatment effect on the underlying parts of the VAT return. As discussed in Sections 3.2 and 3.3, these are the annual aggregations of *VAT payable*, or in other words total output VAT; *total non-taxable sales*, and *VAT paid on inputs* reported by firms in their VAT returns.

Considering that we are leveraging data from multiple audit waves and that the adoption of e-invoicing has been staggered, our identification strategy employs a Stacked Difference-in-Differences (DID) approach. This means that for any wave of treatment, we compare taxpayers who were treated—that is, audited or adopted e-invoicing—to those who were never treated during the entire sample period (see Cengiz et al., 2019; Cunningham, 2021). Furthermore, we combine this approach with Coarsened Exact Matching (CEM, see Iacus et al., 2011, 2012) matching treated and untreated firms based on their aggregate likelihood of being noncompliant as synthesized by their total risk score, the index used by RRA for audit selection.[18] In this way, we account for both RRA risk-based audit selection and for potential endogeneity issues in e-invoicing adoption.[19]

More precisely, our CEM-improved Stacked DID design (CEM-Stacked DID) is performed in three steps. First, individual stacks are created based on the treatment waves, clustering taxpayers treated in each wave with units never treated. Next, all these samples are appended together in the same dataset. Finally, we estimate a fixed effect regression with taxpayer-by-stack fixed effects, time-by-stack fixed effects, standard errors cluster-

[18] See Appendix B that deepens this method. More details on this can also be found in Kotsogiannis et al., 2024.

[19] All firms were given a common deadline by which to adopt the e-invoicing system. Differences in adoption timings might therefore be associated to some extent to some unobserved firm characteristics. Most obviously, less compliant firms are likely to take longer to adopt a technology that makes noncompliance more difficult. Therefore, also for e-invoicing adoption, matching firms based on their risk score measuring their likelihood of being noncompliant is addressing potential selection bias issues.

ing on taxpayer-by-stack, employing weights from the CEM stratification. This design allows us to identify the average treatment effect on the treated (ATT) within each stack by comparing an individual cohort of treated units to never treated taxpayers, avoiding comparisons between late to early treated taxpayers that might bias a standard Two-Way Fixed Effects (TWFE) estimates if effects vary across treated cohorts (Goodman-Bacon, 2021). Formally, we estimate CEM-weighted versions of equation (1) on the stacked dataset. The dependent variable Y_{it} represents, alternatively, each of the margins described above expressed in levels (thousands of US$), $Audit_{it}$ and $E\text{–}INV_{it}$ are dummy indicators that switch on when the taxpayer is first audited and on the year of adoption of e-invoicing, respectively while τ_t and θ_i account respectively for time-by-stack fixed effects and taxpayer-by-stack fixed effects and ε_{it} is the error term. Since the equation is estimated through a fixed effect panel data model, the data is first differentiated in the estimation and as a result the coefficients in equation (1) estimates the overall DID effect across treatment cohorts. More precisely, the parameters of interest are β_1, β_2—that collect respectively the treatment effects of tax audits and e-invoicing adoption employed in isolation—and in particular β_3 estimating their joint effect on compliance that is, the impact of e-invoicing on deterrence through enhanced audit effectiveness.

$$Y_{it} = \alpha + \beta_1 Audit_{it} + \beta_2 E\text{–}INV_{it} + \beta_3 Audit_{it} \times E\text{–}INV_{it} + \theta_i + \tau_t + \varepsilon_{it}. \tag{1}$$

The next section presents the results of the empirical analysis.

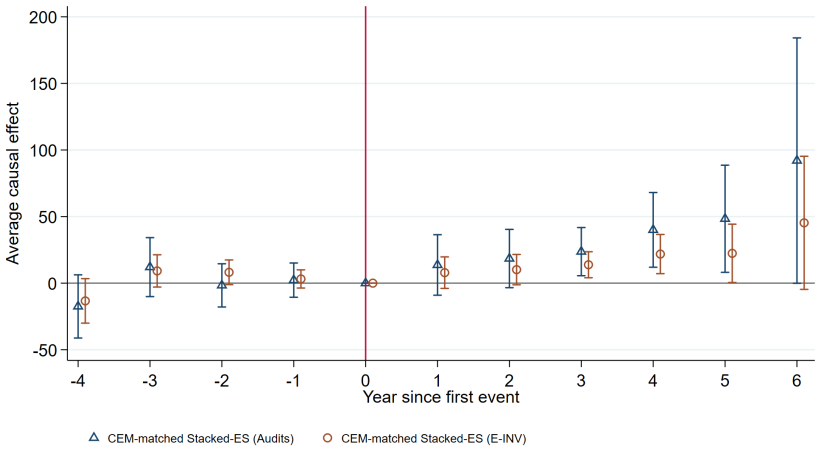
5 Results

5.1 Parallel Trend

As a first step, we test the parallel trend assumption through a CEM-Weighted Stacked Event Study (ES) model applied to our main outcome variable, firms' net VAT liability. This entails estimating dynamic fixed-effects models applied to the audit treatment and e-invoicing adoption separately using the CEM-matched stacked dataset. The absence of significance in pre-treatment differences between treated taxpayers and control units tends to confirm that the pre-treatment parallel trend is achieved and significance in post-treatment estimated coefficients provides a first evidence of the pro-deterrence impact of audits and e-invoicing adoption on firms' net VAT liability reported (see Figure 1).

The analysis turns now to the estimation of equation (1) for the net VAT liability.

Figure 1: Event Study: Aggregate Dynamic response to *VAT Audits* and *E-invoicing*.



Note: This figure reports the estimates of the period-specific treatment effects on net VAT liability for audits and e-invoicing separately. Estimates are obtained through CEM-Weighted Stacked-ES models. Taxpayer-by-stack and year-by-stack fixed effects are controlled for; 95 percent confidence intervals (based on robust standard errors clustered at taxpayer-by-stack level) are shown.

5.2 Net VAT liability

Table 3 focuses on the net VAT liability. Columns (1) and (2) report the aggregate post-treatment impact of audits and e-invoicing adoption respectively when estimated in isolation. The results provide evidence of an aggregate post-treatment increase in the net VAT liability declared of about US$27,000 for audits and about US$11,000 for e-invoicing adopters. These effects tend to persist when included jointly (Column 3). Then we focus on the joint deterrence effect of VAT audits and e-invoicing adoption. The results of the estimation of the fully fledged model described in equation (1) are reported in Column (4) and indicate that the impact on compliance is mostly driven by tax audits performed on firms adopting the e-invoicing system, as the interaction term between *Audit* and $E\text{–}INV$ confirms. Indeed, the only audits leading to a significant improvement in VAT compliance are those involving firms employing e-invoicing, which yield a net combined effect on future VAT payments of about US$33,400. Regarding the impact of the adoption of e-invoicing on compliance, the analysis suggests that it is mostly achieved when combined with audits. Indeed, the estimated coefficient of the adoption of invoicing per se is relatively small in absolute value, while the effect of e-invoicing when paired with audits corresponds to an increase of about US$32,000. All this suggests that the information flow provided to RRA through the digital trail improves the audit performance in terms of deterrence power, conveying to the taxpayer the message that RRA has relevant information that has been used in the audit process which is internalized in taxpayer's future decision-making

process. Additionally, the analysis also indicates that the adoption of e-invoicing per se provides a limited improvement in compliance, and it needs to be paired with further enforcement measures to be fully effective.

Table 3: Impact of Audits and EMB adoption on VAT net liability

	(1)	(2)	(3)	(4)
Audit	26.843***		24.889***	6.677
	(9.404)		(9.248)	(9.746)
E–INV		10.913***	8.501***	5.678*
		(3.475)	(3.205)	(2.933)
Audit*E–INV				26.695*
				(13.630)
Linear combinations				
Audit+Audit*E–INV				33.372***
				(11.853)
EBM_adopted+Audit*E–INV				32.372**
				(13.784)
Observations – Unstacked	39,176	39,176	39,176	39,176
Observations – Stacked	112,472	112,472	112,472	112,472

Note: CEM-Stacked DID models. Taxpayer-by-stack and year-by-stack fixed effects are controlled for. Robust standard errors (clustered at taxpayer-by-stack level) are reported in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3 Other Margins of Non-compliance

To understand what drives these results, we explore next the impact of tax audits and e-invoicing adoption on several additional margins of firms' VAT reporting behaviour. Table 4 reports the results of the estimation of equation (1) on these margins. First, as explained in Sections 3.2 and 4, we decompose the effect on the net VAT liability—reported again in Column (1)—into its output and input components. More precisely, in Column (2) we present the results on VAT payable, the output component of VAT, while in Column (3) we report the results on VAT paid on inputs. Concerning the output component of VAT, the results provide evidence of a significant and positive direct impact of both tax audits and e-invoicing adoption on the aggregation of annual VAT on taxable sales and VAT reverse charge. In this case, when these two instruments are employed in synergy, they do not provide a significant additional impact on compliance. This means that even when used in isolation, audits and e-invoicing enhance compliance on the gross VAT liability.

A consequence of these results, taken together with the progressive expansion of the e-invoicing system, is an expected mechanical increase in adopters' reported VAT

Table 4: Impact of EMB adoption and Audits on the components of net VAT liability

	(1)	(2)	(3)	(4)
Outcome	VAT due	VAT payable	VAT paid on inputs	Non-taxable sales
Audit	6.677	43.953***	29.031	58.754
	(9.746)	(15.383)	(26.772)	(201.189)
E–INV	5.678*	17.738***	6.143	20.975
	(2.933)	(3.741)	(4.211)	(22.957)
Audit*E–INV	26.695*	-6.968	-62.920***	-99.559
	(13.630)	(20.531)	(23.522)	(192.013)
Linear combinations				
Audit+Audit*E–INV	33.372***	36.985**	-33.890*	-40.805
	(11.853)	(17.988)	(18.644)	(151.051)
EBM_adopted+Audit*E–INV	32.372**	10.770	-56.777**	-78.584
	(13.784)	(21.425)	(22.649)	(195.774)
Observations	39,176	39,176	39,176	39,176
Observations − Stacked	112,472	112,472	112,472	112,472

Note: CEM-Stacked DID models. Taxpayer-by-stack and year-by-stack fixed effects are controlled for. Robust standard errors (clustered at taxpayer-by-stack level) are reported in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

paid on inputs. Indeed, as a result of the increase in e-invoicing adoption along the VAT chain, providers are expected to increase the VAT levied (Column 2) in business-to-business transactions (see also Section 3.2). But on top of this effect, there might also be a behavioural one affecting this margin in the same direction. Indeed, firms might decide (in the absence of automatic pre-filling) to over-report VAT paid on inputs as a strategic measure to partially reduce their final VAT liability (VAT due). This behaviour is possible (and not uncommon), but relates in particular to audited firms. Indeed, when it comes to tax audits, there might be two opposite effects on this margin. On the one hand, audits may deliver a pro-deterrence effect through lower VAT paid on inputs following the audit.[20] On the other, audited taxpayers may also decide to increase their gross VAT liability (as documented in Column 2) but try to partially counter this effect over-reporting VAT paid on inputs and ending up with a milder increase in their net VAT liability. Thus, the net effect of audits would depend on which of these two components tends to prevail on the other in the average estimated effect. Column (3) reports the results on this margin providing evidence for both these effects. Indeed, while the combined net impact of audits and adoption of e-invoicing on this margin is significantly promoting compliance, we find that firms who are non-adopters and who are audited increase their input VAT reported by US$29,000—an impact that is not precisely identified—while audited adopters reduce

[20]This is likely in particular if this was a detected source of noncompliance, something we cannot identify due to lack of this information.

them by more than the double leading to a significant net pro-deterrence effect.

A similar strategic margin of action VAT payers may utilize to contain their net VAT payments, and one that has been unexplored by the literature, regards non-taxable sales. Indeed, after experiencing an audit or adopting e-invoicing, firms may compensate the increase in gross VAT liability reported (Column 2), by inflating their non-taxable sales. Column (4) analyses this margin, showing that on average this hypothesis seems to apply for audited taxpayers who have not adopted e-invoicing and unaudited firms who have adopted e-invoicing; the opposite pro-deterrence effect seems to be present for audited e-invoicing adopters. There is so some remarkable similarity in behaviour, in so far as misreporting of sales goes, between audited firms non-adopting e-invoicing and unaudited firms adopting e-invoicing. While this similarity is striking with the current sample the magnitude or sign of these coefficients cannot be measured with sufficient precision.

6 Concluding Remarks

Improving VAT compliance is undoubtedly a major challenge for tax administrations across the world, and in particular so for developing countries (IMF, 2015), which tend to rely on this tax base as a source of revenues (ATAF, 2021b). One reason for the poor revenue performance of many developing countries is that their tax administrations commonly lack effectiveness in their core functions— registration of taxpayers, assessment of their liability, and collection of taxes—and so compliance. Digitalization offers increasingly realistic—and potentially transformative— opportunities for improving tax compliance and revenue mobilisation. The limited ability of administration in low income countries to implement basic reforms suggest that digitilisation is not panacea. For tax administrations operating under weak capacity the implementation of technological solutions should be carefully assessed. This is an important issue and one that is directly related to the design and effectiveness of tax auditing and capacity building in tax administrations.

By using available data on the universe of anonymised VAT records as well as information on firms' e-invoicing adoption and tax audits performed by the RRA, this paper has investigated the role played by the joint use of e-invoicing and tax audits in promoting VAT compliance in Rwanda, a representative developing country. The analysis provided evidence that the introduction of the e-invoicing system in Rwanda delivered a pro-deterrence impact on VAT liabilities reported mainly through tax audit enhancement. E-invoicing adoption provides per se only a limited increase in firms' net VAT payments, while its impact is significantly higher when paired with audits thanks to the (indirect) enhancement of the enforcement policy provided by e-invoicing. Furthermore,

and perhaps more importantly, the results have shown that the only tax audits leading to a significant pro-deterrence effect on VAT compliance are those involving firms employing e-invoicing yielding a net impact of about US$33,400 in the post-treatment period. The paper has also provided some evidence that audited taxpayers and adopters of e-invoicing may simultaneously react by over-reporting the VAT paid on inputs with the aim of reducing their final liability, a pattern that is nuanced and reverted when tax audits are combined with e-invoicing.

The results do not of course suggest that digitalization is not a good strategy for enhancing compliance. To the contrary, it is an enabling factor, but one that needs to be designed and implemented effectively within an ecosystem that records all information through the business-to-business and business-to-consumers transactions. This journey for developing countries is a long, and expensive, one: for this, every step taken needs to be thoroughly evaluated.

We hope to have shown that the results obtained are instructive and the issues regarding the evaluation of digital solutions for tax compliance merit further investigation.

Acknowledgements: We are grateful to the Management of the Rwanda Revenue Authority (RRA) for supporting this research and to the many officials who have provided feedback through extensive discussions. We also thank the participants at the 2nd Annual RRA Research Day, the ATAF & TARC joint workshop on VAT Compliance and Digitalisation in Africa, the 10th Annual Mannheim Taxation Conference, and at the BSE Summer Forum in Public Economics for their helpful comments and suggestions. Kotsogiannis and Salvadori gratefully acknowledge research funding from the Economic and Social Research Council (ESRC) (Grants ES/S00713X/1 and ES/X003973/1). Salvadori also gratefully acknowledges support from the Serra Húnter Programme, the Generalitat de Catalunya [2021SGR00571] and the Spanish Agencia Estatal de Investigación (AEI), through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S) and the Knowledge Generation Projects Programme (Grants PID2022-137707NB-I00 and CNS2022-135749). The views expressed in this paper do not necessarily reflect those of the Rwanda Revenue Authority and its Management. The usual caveat applies.

References

Acemoglu, D., Johnson, S., Robinson, J.A., 2001. The colonial origins of comparative development: An empirical investigation. American Economic Review 91, 1369–1401.

Advani, A., Elming, W., Shaw, J., 2021. The Dynamic Effects of Tax Audits. The Review of Economics and Statistics 0, 1–45.

Ali, M., Shifa, A., Shimeles, A., Woldeyes, F., 2021. Building Fiscal Capacity in Developing Countries: Evidence on the Role of Information Technology. National Tax Journal 74, 591–620.

Almunia, M., Hjort, J., Knebelmann, J., Tian, L., 2022. Strategic or Confused Firms? Evidence from "Missing" Transactions in Uganda. The Review of Economics and Statistics , 1–35.

ATAF, 2021a. Efficient Implementation and Maintenance of ICT Tax Systems in Africa: Compilation of Good Practices, Success Stories, and Lessons Learnt. ATAF Research Publications.

ATAF, 2021b. The African Tax Outlook (ATO) 2021. ATAF's African Tax Outlook Series.

Barreix, A.D., Zambrano, R., Costa, M.A., da Silva Bahia, Á.A., de Jesus, E.A., de Freitas, V.P., Barraza, F., Oliva, N., Andino, M., Rasteletti, A., Drago, C., Cuentas, G., Paredes, M., Pazos, J., Canales, L.A., Campo, R., Castiñeira, L., González, G., Redondo, J.F.M., de Almeida, A.F.F., Hernández, K., Robalino, J., Ramirez, J., Andrade, J.P.A., Goncalves, N., 2018. Electronic invoicing in latin america. URL: https://api.semanticscholar.org/CorpusID:169986831.

Bellon, M., Dabla-Norris, E., Khalid, S., Lima, F., 2022. Digitalization to improve tax compliance: Evidence from vat e-invoicing in peru. Journal of Public Economics 210, 104661. URL: https://www.sciencedirect.com/science/article/pii/S0047272722000639, doi:https://doi.org/10.1016/j.jpubeco.2022.104661.

Besley, T., Ilzetzki, E., Persson, T., 2013. Weak States and Steady States: The Dynamics of Fiscal Capacity. American Economic Journal: Macroeconomics, 5, 205–235.

Besley, T., Persson, T., 2009. The origins of state capacity: Property rights, taxation, and politics. American Economic Review 99, 1218–44.

Best, M., Shah, J., Waseem, M., 2021. The Deterrence Value of Tax Audit: Estimates from a Randomised Audit Program. Mimeo.

Bérgolo, M., Ceni, R., Sauval, M., 2018. Factura electrónica y cumplimiento tributario: Evidencia a partir de un enfoque cuasi-experimental. IDB-DP-561, Banco Interamericano de Desarrollo.

Casey, P., Castro, P., 2011. Electronic Fiscal Devices (EFDs): An Empirical Study of their Impact on Taxpayer Compliance and Administrative Efficiency. IMF Working Paper, WP/15/73, Fiscal Affairs Department.

Castro, H.J., Carrillo, A.Z., Cortés, S.B., Aragón, G.A., Diez, M.E., Salgado, M.M.S., 2016. Impacto en la evasión por la introducción de la factura electrónica. Instituto Tecnológico y de Estudios Superiores de Monterrey. No. IA-006E00002-E37-2016 .

Cengiz, D., Dube, A., Lindner, A., Zipperer, B., 2019. The Effect of Minimum Wages on Low-Wage Jobs. The Quarterly Journal of Economics 134, 1405–1454.

Cunningham, S., 2021. Causal Inference: The Mixtape. Yale University Press.

DeBacker, J., Heim, B.T., Tran, A., Yuskavage, A., 2015. Legal Enforcement and Corporate Behavior: An Analysis of Tax Aggressiveness after an Audit. Journal of Law and Economics, 58, 291–324.

DeBacker, J., Heim, B.T., Tran, A., Yuskavage, A., 2018. Once Bitten, Twice Shy? The Lasting Impact of Enforcement on Tax Compliance. Journal of Law and Economics, 61, 1–35.

Eissa, N., Zeitlin, A., 2014. Using Mobile Technologies to Increase VAT Compliance in Rwanda. Mimeo, McCourt School of Public Policy, Georgetown University.

Gemmell, N., Ratto, M., 2012. Behavioral Responses to Taxpayer Audits: Evidence From Random Taxpayer Inquiries. National Tax Journal, 65, 33–57.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. Journal of Econometrics 225, 254–277. Themed Issue: Treatment Effect 1.

Iacus, S.M., King, G., Porro, G., 2011. Multivariate Matching Methods That Are Monotonic Imbalance Bounding. Journal of the American Statistical Association, 106, 345–361.

Iacus, S.M., King, G., Porro, G., 2012. Causal Inference without Balance Checking: Coarsened Exact Matching. Political Analysis, 20, 1–24.

IMF, 2015. Current Challenges in Revenue Mobilization - Improving Tax Compliance. IMF Policy Paper, 5.

Kleven, H.J., Knudsen, M.B., Kreiner, C.T., Pedersen, S., Saez, E., 2011. Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. Econometrica, 79, 651–692.

Kochanova, A., Hasnain, Z., Larson, B., 2016. Does e-government improve government capacity? evidence from tax administration and public procurement. Evidence from Tax Administration and Public Procurement (April 28, 2016). World Bank Policy Research Working Paper .

Kotsogiannis, C., Salvadori, L., Karangwa, J., Theonille, M., 2024. Do tax audits have a dynamic impact? evidence from corporate income tax administrative data. Journal of Development Economics 170, 103292.

Lee, H.C., 2016. Can electronic tax invoicing improve tax compliance? a case study of the republic of korea's electronic tax invoicing for value-added tax. A Case Study of the Republic of Korea's Electronic Tax Invoicing for Value-Added Tax (March 7, 2016). World Bank Policy Research Working Paper .

Mascagni, G., Dom, R., Santoro, F., Mukama, D., 2023. The VAT in practice: equity, enforcement, and complexity. International Tax and Public Finance 30, 525–563.

Mascagni, G., Mengistu, A.T., Woldeyes, F.B., 2021. Can icts increase tax compliance? evidence on taxpayer responses to technological innovation in ethiopia. Journal of Economic Behavior & Organization 189, 172–193.

Mascagni, G., Mukama, D., Santoro, F., 2019. An Analysis of Discrepancies in Taxpayers' VAT Declarations in Rwanda. ICTD Working Paper 92.

Mascagni, G., Santoro, F., Mukama, D., Karangwa, J., Hakizimana, N., 2022. Active Ghosts: Nil-filing in Rwanda. World Development 152.

Naritomi, J., 2019. Consumers as tax auditors. American Economic Review 109, 3031–72.

OECD, 2022. Tax Administration 3.0 and Electronic Invoicing. URL: https://www.oecd-ilibrary.org/content/publication/2ffc88ed-en, doi:https://doi.org/https://doi.org/10.1787/2ffc88ed-en.

OECD, AUC, ATAF, 2020. Revenue Statistics in Africa 2020. OECD Report.

Okunogbe, O., Pouliquen, V., 2022. Technology, taxation, and corruption: evidence from the introduction of electronic tax filing. American Economic Journal: Economic Policy 14, 341–372.

Okunogbe, O., Santoro, F., 2023. Increasing Tax Collection in African Countries: The Role of Information Technology. Journal of African Economies 32, i57–i83.

Pomeranz, D., 2015. No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax. American Economic Review, 105, 2539–2569.

Ramírez, J., Oliva, Nicolás; Andino, M., 2018. Facturación electrónica en Ecuador: Evaluación de impacto en el cumplimiento tributario. IDB-DP-563, Banco Interamericano de Desarrollo.

RRA, 2019. Tax Statistics in Rwanda, Fiscal Year 2017/18. Second Edition, June 2019, Kigali.

Templado, I., Artana, D., 2018. Análisis del impacto de la factura electrónica en Argentina. IDB-DP-562, Banco Interamericano de Desarrollo.

Tourek, G., 2022. Targeting in tax behavior: Evidence from Rwandan firms. Journal of Development Economics 158, 102911.

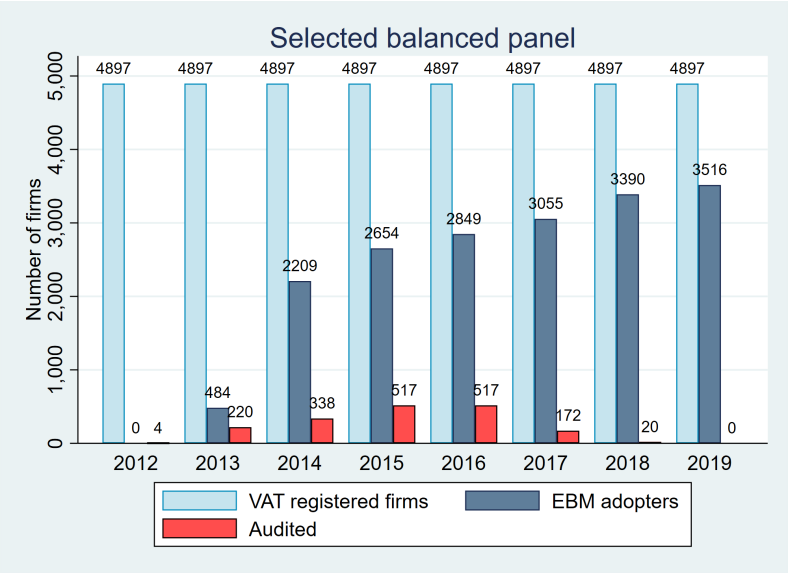
Yilmaz, F., Coolidge, J., 2013. Can e-filing reduce tax compliance costs in developing countries? World Bank Policy Research Working Paper .

Online Appendices

Appendix A: Additional Descriptive Statistics

Figure A.1 presents a graphical representation of the composition of our selected sample in which we include also the firms audited in the waves we have been granted access to.[1]

Figure A.1: Composition of selected sample



Note: The graph presents the composition of our balanced sample.

Tables A.1-A.4 present the descriptive statistics of the main audit outcome variables, clustering our sample of firms by size as defined by the RRA (see also footnote 13) and by e-invoicing adoption status. The aim of this exercise is to compare the average performance of audits involving e-invoicing adopters with those implicating non-adopters but maintaining a comparable size across these cohorts. With the aim of maintaining the subsamples as comparable as possible in terms of their sample size across e-invoicing adoption cohorts, we group together medium and large firms in the first cluster, and small and micro firms in the second cluster. For medium-large firms, both the average underreporting uncovered through the inspection, the audit outcome, the total fines charged, and the total audit outcome are about the double for e-invoicing adopters compared to non-adopters (Tables A.1- A.2). Regarding micro-small firms, we record a similar ratio in terms of audit outcome, but fines tend to be similar across cohorts, with those applied to e-invoicing adopters being just above those charged to non-adopters. The resulting total audit outcome is about 1.5 times higher for e-invoicing adopters. All this suggests that by providing accurate information on firms' transactions, the e-invoicing system can boost the monitoring ability of the revenue authority, enhancing its capacity to detect tax evasion through audits. Interestingly, when we compare audits across these cohorts but

[1] As noted above in the text, we have been granted access to 5 waves of audits, 2013-2017, but they are recorded into a specific tax period depending on the month in which the taxpayers have been notified, this is why few of them are coded as starting in 2012 and 2018.

in terms of their relative performance—that is considering the outcomes as share of the potential tax base—audits involving e-invoicing non-adopters tend to slightly outperform audits targeting e-invoicing users. Given that the potential tax base is defined as the sum of the tax base declared by the taxpayer and the audit outcome, this means that within the same size category, firms using e-invoicing declare higher tax bases than non-adopters and do so more than proportionally compared to the difference in the outcome of the audits.

All this, read together with the main results of the paper on deterrence effects, seems to confirm that the improved compliance channelled through e-invoicing is obtained by increasing the gross declared tax base and by enhancing the detection and deterrence power of audits.

Table A.1: Audits descriptive statisics e-invoicing adopters – Medium & Large firms (2013-2017)

Variable	Measurement Unit	Observations	Mean	Std. Dev.
Audit outcome	1,000 US $	349	118.99	1100.04
Audit outcome (%)	% Potential tax base	349	19	34.77
Total fines	1,000 US $	349	69.7	680.39
Total audit outcome	1,000 US $	349	188.69	1774.53
Total audit outcome (%)	% Potential tax base (including fines)	349	21.36	36.39

Note: Authors' calculations based on data provided by RRA.

Table A.2: Audits descriptive statisics e-invoicing non-adopters – Medium & Large firms (2013-2017)

Variable	Measurement Unit	Observations	Mean	Std. Dev.
Audit outcome	1,000 US $	434	54.46	251.18
Audit outcome (%)	% Potential tax base	434	26.02	37.43
Total fines	1,000 US $	434	28.03	159.54
Total audit outcome	1,000 US $	434	82.76	401.34
Total audit outcome (%)	% Potential tax base (including fines)	434	29.97	39.46

Note: Authors' calculations based on data provided by RRA.

Table A.3: Audits descriptive statisics e-invoicing adopters – Small & Micro firms (2013-2017)

Variable	Measurement Unit	Observations	Mean	Std. Dev.
Audit outcome	1,000 US $	611	8.16	48.06
Audit outcome (%)	% Potential tax base	611	33.37	45.11
Total fines	1,000 US $	611	4.34	31.84
Total audit outcome	1,000 US $	611	12.51	79.06
Total audit outcome (%)	% Potential tax base (including fines)	611	36.5	46.31

Note: Authors' calculations based on data provided by RRA.

Table A.4: Audits descriptive statisics e-invoicing non-adopters – Small & Micro firms (2013-2017)

Variable	Measurement Unit	Observations	Mean	Std. Dev.
Audit outcome	1,000 US $	394	6.15	18.31
Audit outcome (%)	% Potential tax base	394	32.01	44.31
Total fines	1,000 US $	394	3.41	11.49
Total audit outcome	1,000 US $	394	9.63	29.13
Total audit outcome (%)	% Potential tax base (including fines)	394	34.68	45.29

Note: Authors' calculations based on data provided by RRA.

Appendix B: Outcomes of the CEM Procedure

Several matching techniques are discussed in the literature (see, among others, Stuart, 2010; King et al., 2011; Imbens and Rubin, 2015; Guo and Fraser, 2015). In general, all matching methods for causal inference seek a trade-off between maximizing balance on the relevant pre-treatment covariates between the treated and control units while keeping a reasonable matched sample size (King et al., 2011). Among the available methods we employ CEM, which has been proven to possess a set of powerful statistical properties. In particular, CEM has been shown to perform better than commonly used matching methods (like e.g. Propensity Score Matching and Mahalanobis Distance Matching) in reducing the initial imbalance across treatment cohorts.[2] Moreover, CEM algorithm is extremely intuitive. First, CEM temporarily coarsens each relevant pre-treatment variable into meaningful groups through a threshold assigned by the user based on intuitive substantive information, where it is possible, or through alternative standard binning algorithms.[3] Subsequently, units with the same 'bin signature' (that is, with the same values) for all the coarsened variables are placed in a single stratum. And, finally, the control units within each stratum are weighted to equal the number of treated units in that stratum. Strata without at least one treated and one control unit are pruned from the data set. Each treated unit is weighted with 1 while the weights for each control unit equals the number of treated units in its stratum divided by the number of control units in the same stratum, normalized so that the sum of the weights equals the total matched sample size. By employing these weights we analyse the unpruned units through a DID approach to finally estimate equation (1) on the stacked dataset.

Specifically, we employ CEM to stratify the sample based on the sole risk score, which is the relevant index used by RRA to prioritize audits and a measure of firms' likelihood of being noncompliant. In the interest of mantaining a reasonable sample size, we decided to focus on a single variable, also given that in our models taxpayer-by-stack and year-by-stack fixed effects are already controlling for most of potential heterogeneity across cohorts. Thus, focusing on the risk scores allows us to reduce the remaining imbalance across treatment cohorts that is due to potential differences in the likelihood of noncompliance—the main source of sample selection—not already accounted for in our model.

Table A.5, Panel A reports the summary of the matching procedure in terms of Matched and Unmatched observations for the audit treatment. By focusing on the risk score we can match all the observations we have in our selected sample. Table A.5, Panel B provides a measure of imbalance reduction through L_1 statistics introduced by Iacus et al. (2011). Specifically, in our case this imbalance measure is based on the L_1 difference between the histogram of risk scores across treatment cohorts (see Iacus et al., 2011 for a formal definition). In short, L_1 is bounded between 0 and 1—with higher values indicating higher imbalance—and it is an index that should be evaluated in relative rather than absolute terms by comparing the values before and after the stratification process. After CEM imbalance in the risk scores reduces to half of the initial imbalance, indicating that

[2]CEM also reduces model dependence, estimation error, bias, variance, mean square error, and other criteria while seeking a trade-off between sample size and balance (see Iacus et al., 2011, 2012; Blackwell et al., 2009; King et al., 2011; King and Nielsen, 2019 for more details and formal proofs, and Iacus et al., 2019 for a discussion on the inference theory).

[3]In this case, we use the Sturge's rule

homogeneity in the likelihood of being noncompliant across treatment cohorts increases significantly as a result of the CEM process while keeping all observations matched-in. This is visually confirmed in Figure A.2 that plots the distribution of the risk scores before and after the CEM procedure.

Table A.5: Summary of the CEM matching procedure - Audit treatment

Panel A: Matching summary

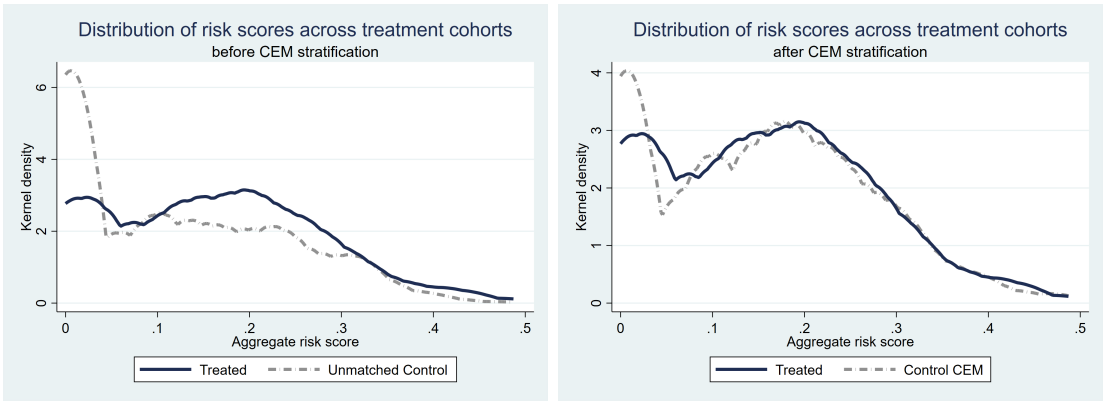
	Never-audited	Audited (first time)
Unique taxpayers	4,026	871
Matched	4,026	871
Unmatched	0	0

Panel B: Difference across treatment cohorts

Before CEM							
	L1	mean	min	25%	50%	75%	max
Risk score	0.209	0.041	0	0.064	0.077	0.038	0
After CEM							
	L1	mean	min	25%	50%	75%	max
Risk score	0.112	0.003	0	0	0.013	0.013	0

Note: The table depicts the matching summary of the CEM procedure, L_1 statistics for imbalance as defined in Iacus et al. (2011), and differences across treatment cohorts in the distribution of the risk scores before and after CEM.

Figure A.2: Risk score imbalance reduction (CEM)



(a) (b)

Note: Authors' calculations based on data provided by RRA.

Similarly, we use the same stratification for the treatment involving e-invoicing adoption. Results are reported in Table A.6 and Figure A.3. In this case, we are able to match all observations and reduce imbalance to less than half of the initial level.

Table A.6: Summary of the CEM matching procedure - E-invoicing treatment

Panel A: Matching summary

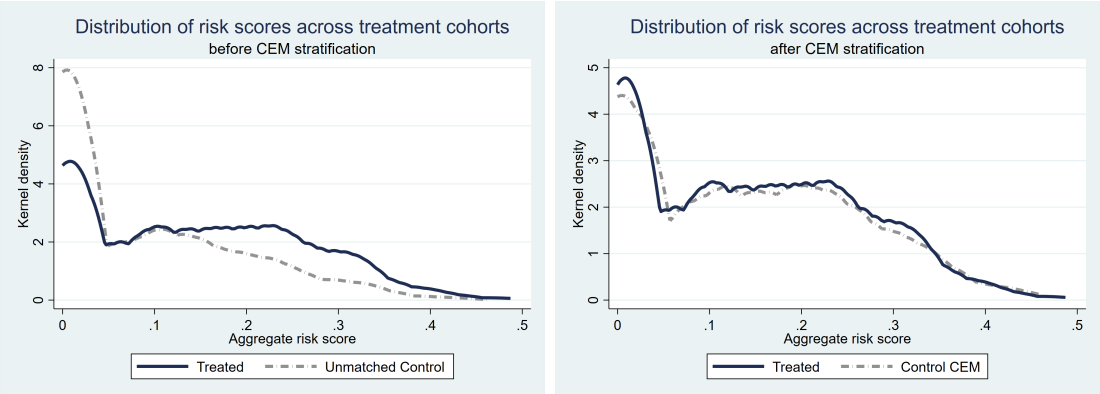
	Non-adopters	E-invoicing adopters
Unique taxpayers	1,381	3,516
Matched	1,381	3,516
Unmatched	0	0

Panel B: Difference across treatment cohorts

Before CEM							
	L1	mean	min	25%	50%	75%	max
Risk score	0.241	0.058	0	0.026	0.103	0.077	0.026
After CEM							
	L1	mean	min	25%	50%	75%	max
Risk score	0.103	0.003	0	0.026	0	0	0.026

Note: The table depicts the matching summary of the CEM procedure, L_1 statistics for imbalance as defined in Iacus et al. (2011), and differences across treatment cohorts in the distribution of the risk scores before and after CEM.

Figure A.3: Risk score imbalance reduction (CEM)



(a) (b)

Note: Authors' calculations based on data provided by RRA.

References

Blackwell, M., Iacus, S., King, G., Porro, G., 2009. Cem: Coarsened Exact Matching in Stata. The Stata Journal, 9, 524–546.

Guo, S., Fraser, M., 2015. Propensity Score Analysis: Statistical Methods and Applications. Advanced Quantitative Techniques in the Social Sciences, SAGE Publications. Second Edition.

Iacus, S.M., King, G., Porro, G., 2011. Multivariate Matching Methods That Are Monotonic Imbalance Bounding. Journal of the American Statistical Association, 106, 345–361.

Iacus, S.M., King, G., Porro, G., 2012. Causal Inference without Balance Checking: Coarsened Exact Matching. Political Analysis, 20, 1–24.

Iacus, S.M., King, G., Porro, G., 2019. A theory of statistical inference for matching methods in causal research. Political Analysis 27, 46–68.

Imbens, G.W., Rubin, D.B., 2015. Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge University Press, USA.

King, G., Nielsen, R., 2019. Why propensity scores should not be used for matching. Political Analysis, 27, 435–454.

King, G., Nielsen, R., Coberley, C., Pope, J.E., Wells, A., 2011. Comparative Effectiveness of Matching Methods for Causal Inference. Mimeo.

Stuart, E.A., 2010. Matching Methods for Causal Inference: A Review and a Look Forward. Statistical Science, 25, 1–21.