DO TAX AUDITS DETER CORPORATE INCOME NONCOMPLIANCE? EVIDENCE FROM RWANDA

Christos Kotsogiannis[1,2] **Luca Salvadori**[1,3] John Karangwa[4]
Theonille Mukamana[4]

[1]TARC, University of Exeter - Business School

[2]CESIfo, [3]IEB, [4] Rwanda Revenue Authority

TARC Workshop on Tax Audits Evaluation, 9th December 2020

Disclaimer

The views expressed in this paper do not necessarily reflect those of the Rwanda Revenue Authority and its Management.

Motivation

- Tax audits: crucial parameters of multidimensional tax system
- Understanding how audits impact on deterring future noncompliance is important for efficiency and fairness of tax system
- Especially true (but not only) in low-income countries where administrative capacity is still generally limited
- These are pressing issues, following development sustainability goals (SDGs, see UN-DESA, 2016) ...
- ... but also post pandemic responses which require significant revenue mobilisation.

Motivation

Recent increased interest among academics and policy makers has led to a proliferation of studies (framed in developed country context):

- **Random audits:** e.g. Kleven et al. (2011); Gemmell and Ratto (2012); Pomeranz (2015); DeBacker et al. (2018*a*,*b*); Advani, Elming and Shaw (2019);
- **Risk-based audits:** e.g. DeBacker et al. (2015); Løyland et al. (2019); Beer et al. (2020); Erard, Kirchler and Olsen (2019)
- . . . But lack of evidence for developing world (notable exceptions being Lediga, Riedel and Strohmaier, 2020; Best, Shah and Waseem, 2021)
  - There is need of enhancing fiscal capacity;
  - Audit assessment might lead to slightly different results.

Our study in a nutshell

- **Our research questions:**
  1. Are tax audits (across tax bases) deterring future noncompliance on CIT?
  2. What is the impact of different types of tax audits?
- **We combine 3 sources of data:**
  - Universe of CIT and VAT anonymised tax declarations for the period 2013-2018
  - Universe of (risk-based) audit data for the 2015 wave (anonymised)
  - Detailed risk rules/criteria and risk weighting scheme $\Rightarrow$ risks scores for audit selection
- **Our approach: matched-DID**
  1. CEM, Kernel-PSM, Kernel-MHD, Nearest Neighbour MHD
  2. IPTW

Our study in a nutshell

Our results:

- **Significant pro-deterrence effect** on CIT reporting one year after audit that corresponds to an increase of 20.7% (12.3%) in corporate income (CIT payable) reported by audited businesses the year after receiving the audit (not significant impact after 2-3 years).
- Noncompliant drive the results.
- **Type of audit matters:**
  - Comprehensive audits drive the results.
  - Narrow-scope audits have counter-deterrence effect (after 2 years).

Outline

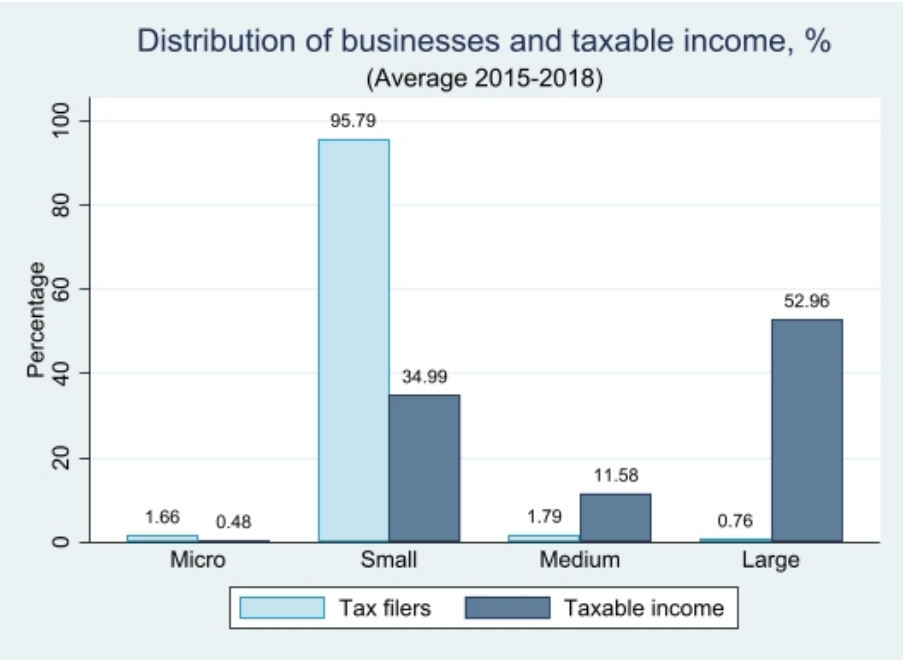
- Institutional setting and data
- Empirical approach and identification strategy
- Main results
- Conclusions

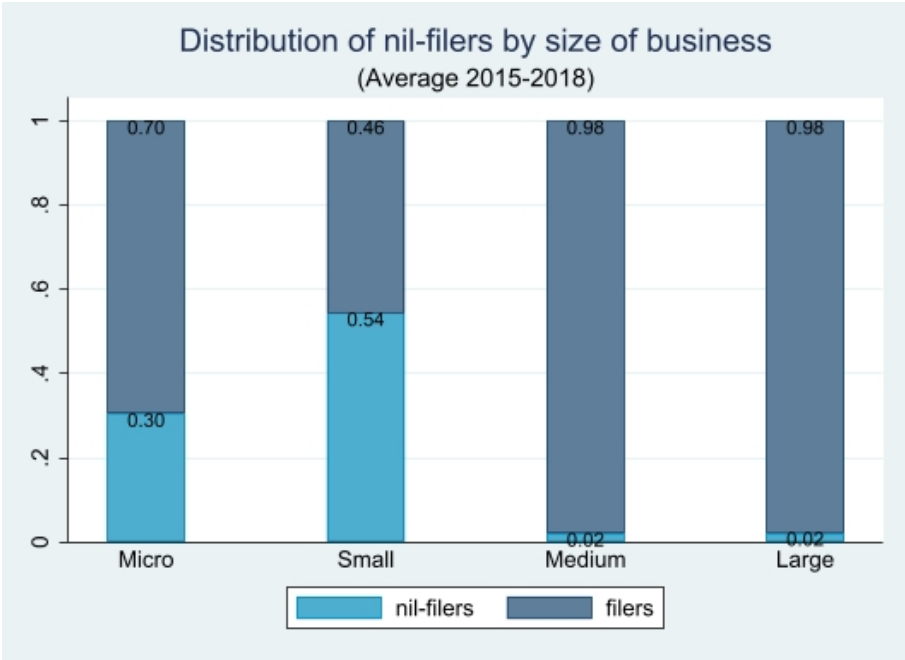
Institutional setting

Why Rwanda?

- **Representative of low-income countries** in terms of fiscal capacity and tax structure...
- **... but embracing reforms and developing capabilities fast** to improve service delivery and enhance tax compliance e.g. through:
  - intensification of tax education and information programs,
  - monitoring of non-filers and non-payers,
  - risk-based audit selection,
  - enhancement and integration of EBM, e-tax system and local government tax management system,
  - further progress towards enhancement of an electronic single window system.

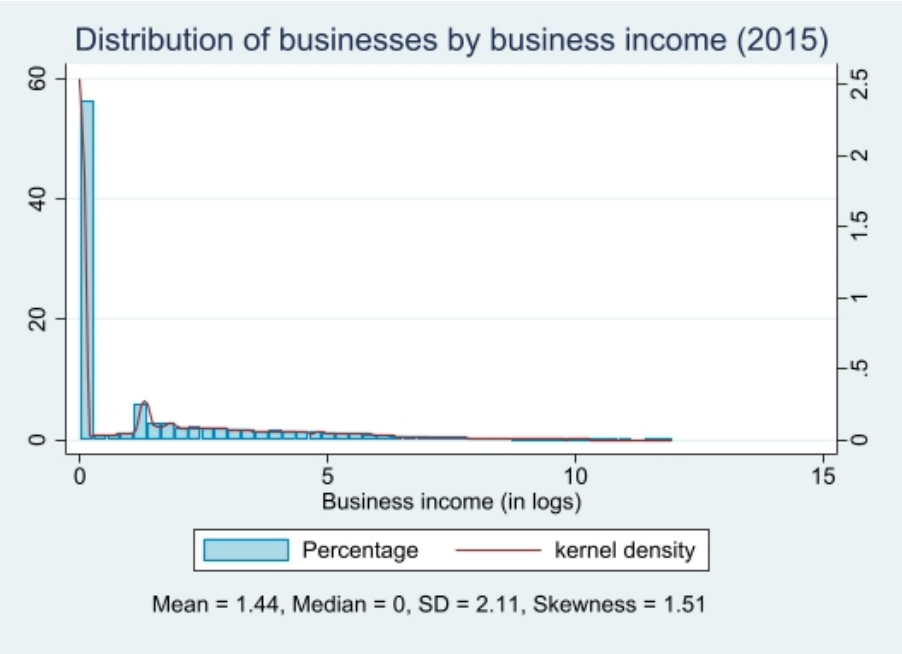
Data: CIT

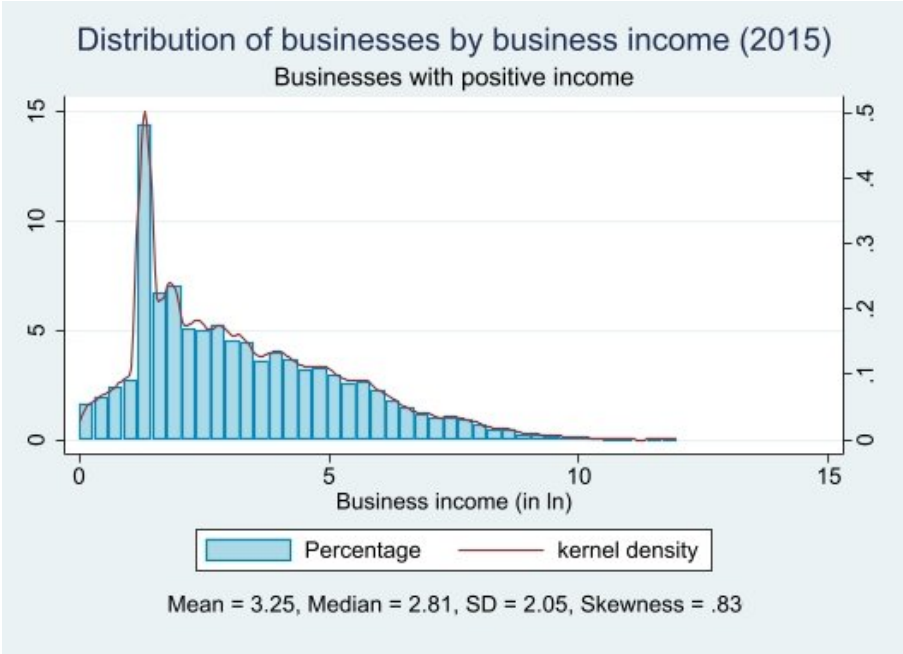




- Universe of CIT tax declarations for the period 2013-2018
- Income generated by incorporated businesses
- Standard CIT rate of 30% of profit with some reductions
- Simplified revenue-based tax regime 3% of turnover (small firms)
- Flat tax regimes (micro-enterprises)

Data: CIT





Number of CIT filers by fiscal year (2013-2018)

Tax period	2013	2014	2015	2016	2017	2018
Total number of CIT declarations	13,778	24,405	29,174	32,572	36,793	40,490

More on CIT data here

Data: Audits

Universe of anonymised completed tax audits for the 2015 wave

- **Desk audits (45%):** conducted by RRA staff using information already submitted to RRA.
- **Issue audits (18%):** usually focused on a single tax type, single aspect or single tax period; usually desk-based
- **Comprehensive audits (37%):** in-person, in-depth and time-intensive across tax bases.

Variable	Obs	Measurement Unit	Mean	Std.Dev	Min	Max
Underreporting detected	435	1000 US $	101.15	969.81	0	19,369.84
Total fines	435	1000 US $	56.36	585.85	0	11,621.90
Interests charges	297	1000 US $	8.26	24.79	0	294.98
Total audit outcome	435	1000 US $	163.14	1554.99	0	30,991.74
Total audit outcome (%)	418	% Potential revenues	66.87	40.21	0	100

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

Data: Risk Scoring

- RRA provided us with information on the criteria for risk-based audit selection (including the risk rules employed to assign risk scores to all tax declarations and correspondent weighting scheme)
- These rules involve the use of a large amount of data including data from other tax bases, in particular VAT used to flag discrepancies
- We employed this data to replicate the aggregate risk scores for any taxpayer
- Data have been complemented by the anonymised VAT declarations of CIT filers

Empirical Strategy: the selection bias

- We want to estimate the ATT of audits on treated CIT filer's reporting behaviour

$$ATT = E\left[\Delta Y^1 | D = 1\right] - E\left[\Delta Y^0 | D = 1\right]. \tag{1}$$

- Risk-based audits target tax returns that are most suspected of noncompliance

$$\Downarrow$$

- Selection bias: $E\left[\Delta Y^0 | D = 1\right] - E\left[\Delta Y^0 | D = 0\right] \neq 0 \Rightarrow$ estimate the counterfactual
- We combine matching methods with a difference-in-difference approach

$$\widehat{ATT} = \frac{1}{N^1} \sum_{i:D_i=1} \left[\Delta Y_i^1 - \sum_{j:D_j=0} W(i,j)\, \Delta Y_j^0\right]. \tag{2}$$

Empirical Strategy: matching methods

- **Exact Matching** matches a treated unit to all control units with the same covariate values
  - **Pros:** perfectly balanced matched data
  - **Cons:** very few matches
- **Approximate matching methods:** specify a metric to find control units that are close to the treated unit (e.g. PSM, MHD).
  - **Pros:** convenient synthetic measures do overcome EM limitations
  - **Cons:** the user has to set the size of the matching solution ex ante, then check for balance ex post
- **Coarsened Exact Matching (CEM):** temporarily coarsens variables into meaningful groups, exact match on these coarsened data (through a "bin signature") and then balance original matched data through weights.
  - **Pros:** coarsening bounds the maximum imbalance through an ex ante choice. CEM tends to perform better in balancing and can improve other matching methods
  - **Cons:** as any other matching method, trade-off balance/size

Empirical Strategy: matching variables

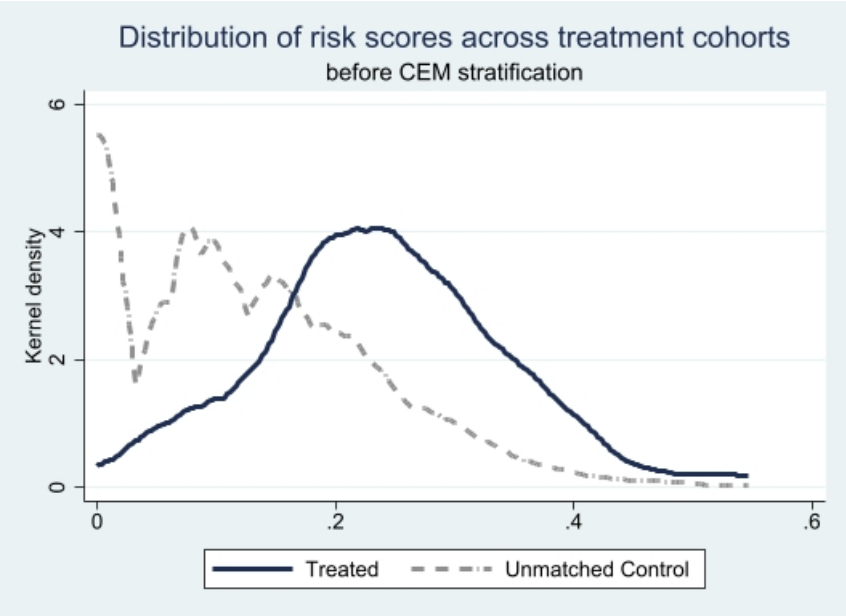
- We are applying CEM-DID as main strategy and CEM-improved Kernel - PSM, Kernel - MHD and Nearest-Neighbours DID estimators as robustness
- Main matching variables: Risk Score, taxable income in year t, $t-1$ and $t-2$
  - Alternative broader sets of variables employed also include VAT paid on inputs, sector of activity (exact matching), index of source of income, tax centre indicator (exact matching) ad lags among others.
  - While corrobrating the results, these alternative sets of covariates lead to inferior matching solutions in terms of the trade-off size/imbalance
- Propensity Score: sequential selection process employed to select the set of pre-treatment covariates to predict the PS based on predictive power.

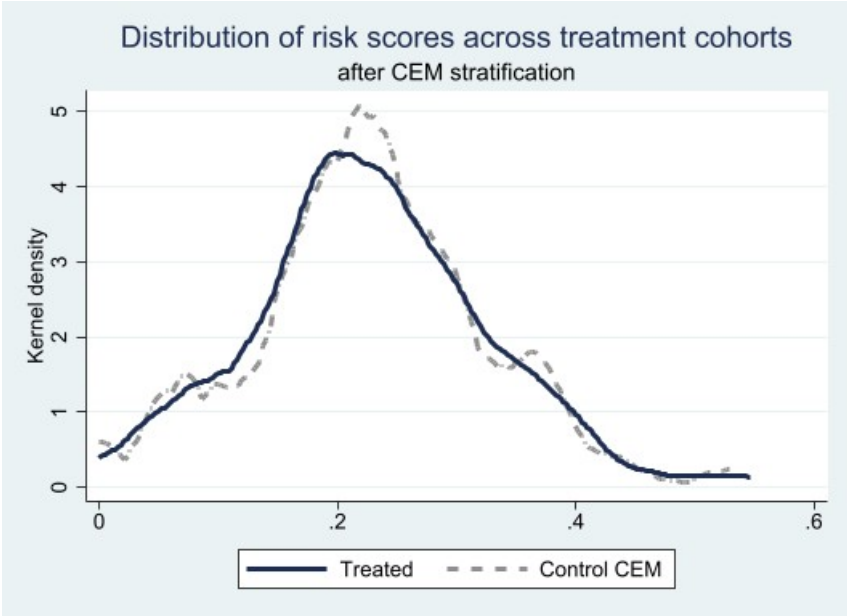
Empirical Strategy: sample selection

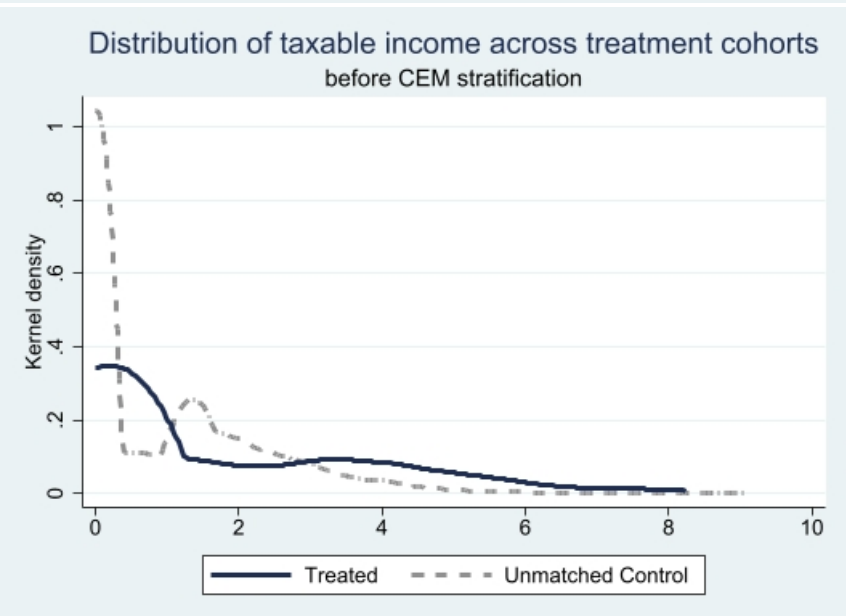
Step	Description	Control Sample	% Δ	Audit Sample	% Δ	Total Sample	% Δ
0	Universe of CIT filers in 2015	28,619	-	435	-	29,174	-
1	Drop outliers with effective tax rate >1	28,610	99.97%	435	100.00%	29,165	99.97%
2	Failure to file timely before treatment	11,203	39.16%	424	97.47%	11,627	39.87%
3	Violation of (pre&post 2015) non-audit restrictions	10,859	96.93%	362	85.38%	11,221	96.51%
4	Final matched sample after CEM	5,577	51.36%	304	83.98%	5,881	52.41%

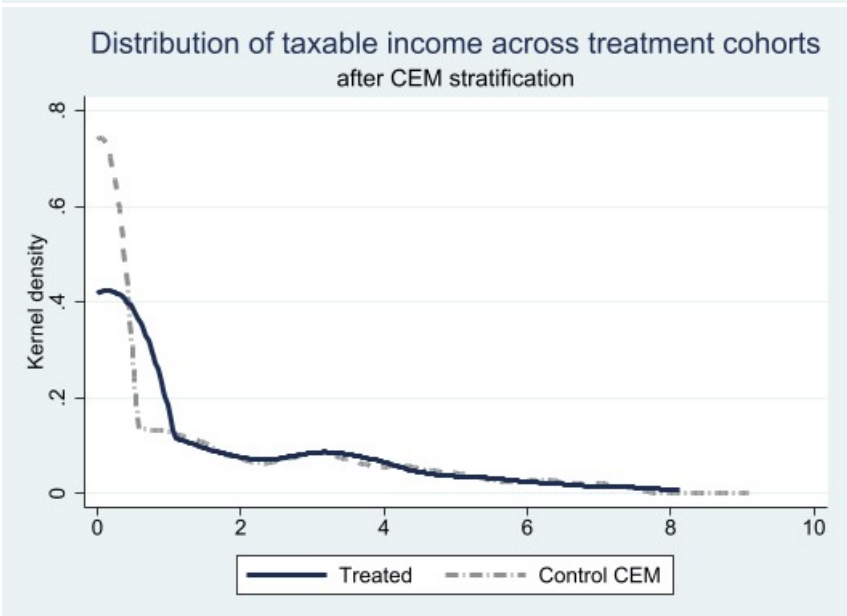
Note: Authors' calculations based on data provided by RRA.

Empirical Strategy: balance performance - graphs









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

More on imbalance here

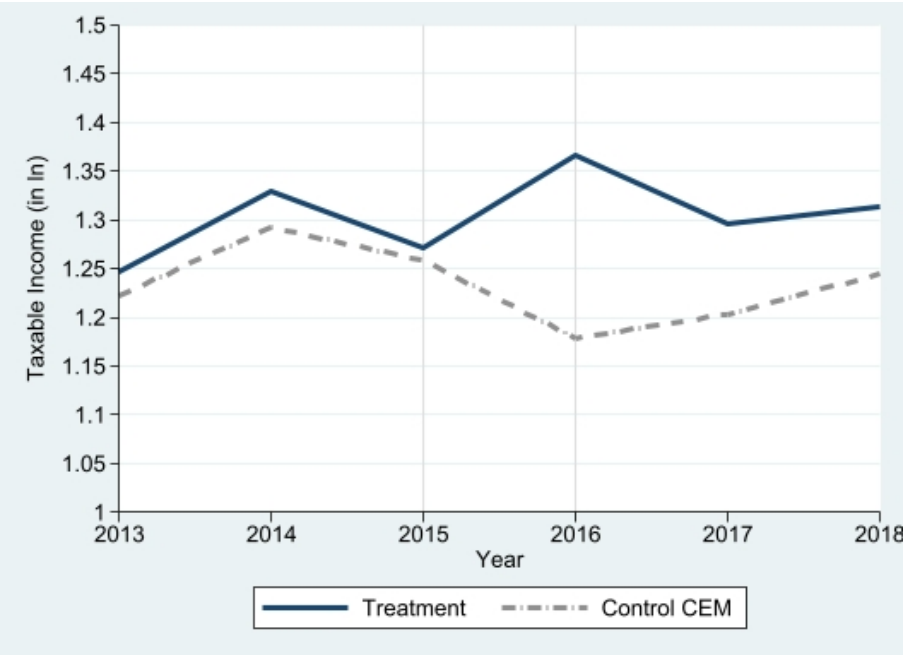
Main Results – Aggregate ATT

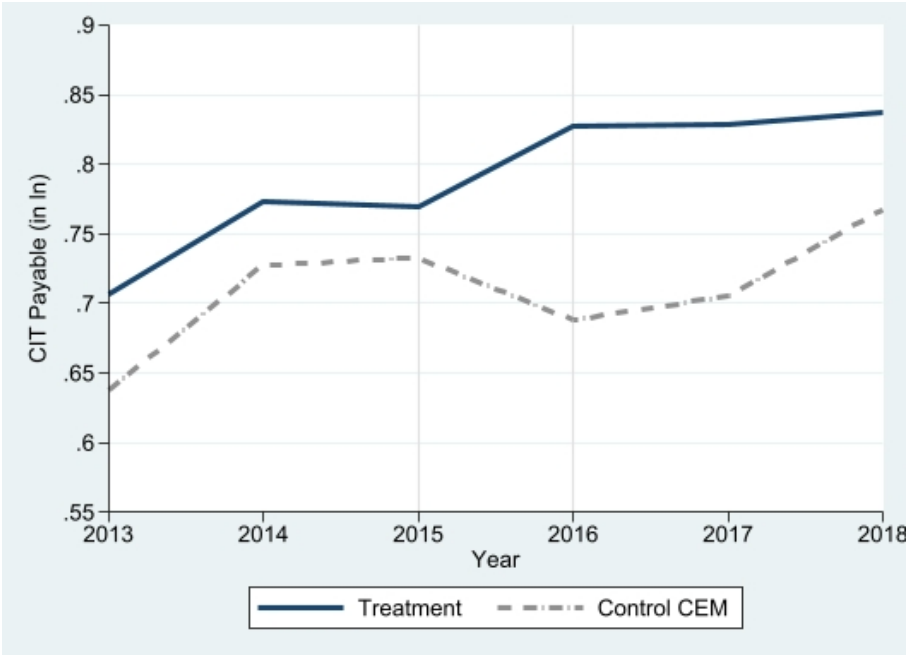
Dependent Variable	Taxable Income			CIT payable		
Years after the audit	I	II	III	I	II	III
Matching estimator	(1)	(2)	(3)	(4)	(5)	(6)
CEM	**0.175**	0.080	0.056	**0.103**	0.087	0.033
	(0.023)***	(0.147)	(0.111)	(0.017)***	(0.107)	(0.081)
	(0.033)***	(0.205)	(0.136)	(0.028)***	(0.147)	(0.098)
	(0.085)**	(0.109)	(0.118)	(0.061)*	(0.079)	(0.083)
Kernel - MHD	**0.208**	0.003	0.025	**0.124**	0.030	0.012
	(0.023)***	(0.147)	(0.111)	(0.017)**	(0.107)	(0.081)
	(0.084)**	(0.100)	(0.097)	(0.057)**	(0.068)	(0.069)
	(0.072)***	(0.091)	(0.088)	(0.047)***	(0.065)	(0.062)
Kernel - PSM	**0.148**	-0.074	-0.145	**0.119**	0.023	-0.059
	(0.081)*	(0.107)	(0.117)	(0.059)**	(0.073)	(0.081)
	(0.085)*	(0.103)	(0.114)	(0.059)**	(0.071)	(0.083)
	(0.080)*	(0.102)	(0.103)	(0.054)**	(0.070)	(0.083)
Nearest Neighbour	**0.297**	0.125	0.195	**0.147**	0.079	0.097
	(0.099)***	(0.120)	(0.143)	(0.072)**	(0.084)	(0.096)
	(0.095)***	(0.218)	(0.198)	(0.072)**	(0.164)	(0.145)
	(0.115)***	(0.143)	(0.162)	(0.080)*	(0.102)	(0.113)

Note: Alternative standard errors are reported in parentheses for any specification. CEM: robust standard errors (clustered by tax center), bootstrapped standard errors (clustered by tax center) based on 500 replications, and stratified bootstrapped standard errors based on 500 replications; Kernel - MHD and Kernel - PSM: bootstrapped standard errors based on 200, 500 replications and stratified bootstrapped standard errors based on 500 replications; Nearest Neighbour: heteroskedasticity-consistent analytical standard errors proposed by Abadie and Imbens (2006), wild bootstrapped standard errors based on 500 replications and stratified bootstrapped standard errors based on 500 replications; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

More on inference here

ATT in graphs





Note: *ATT* of Audits on audited taxpayers (CEM): Taxable Income in ln (left panel); CIT Payable in ln (right panel)

Main Results – ATT by Audit Outcome

	Determined Noncompliant						Determined Compliant					
Dep. Variable	Taxable Income			CIT payable			Taxable Income			CIT payable		
After audit	I	II	III	I	II	III	I	II	III	I	II	III
Estimator	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
CEM	0.166	0.086	0.049	0.097	0.105	0.036	0.248	0.039	0.123	0.152	-0.051	0.006
	(0.034)***	(0.172)	(0.123)	(0.020)***	(0.123)	(0.089)	(0.151)	(0.260)	(0.217)	(0.120)	(0.168)	(0.125)
	(0.040)***	(0.229)	(0.138)	(0.025)***	(0.162)	(0.100)	(0.176)	(0.303)	(0.277)	(0.141)	(0.197)	(0.169)
	(0.097)*	(0.121)	(0.129)	(0.068)	(0.188)	(0.093)	(0.116)**	(0.212)	(0.245)	(0.079)*	(0.137)	(0.148)
Kernel - MHD	0.212	0.022	0.033	0.128	0.058	0.023	0.089	-0.043	-0.072	-0.008	-0.089	-0.015
	(0.080)***	(0.101)	(0.106)	(0.055)**	(0.069)	(0.073)	(0.174)	(0.281)	(0.199)	(0.093)	(0.164)	(0.120)
	(0.090)**	(0.108)	(0.106)	(0.061)**	(0.075)	(0.075)	(0.178)	(0.267)	(0.209)	(0.099)	(0.164)	(0.117)
	(0.085)**	(0.100)	(0.100)	(0.056)**	(0.072)	(0.069)	(0.168)	(0.257)	(0.200)	(0.091)	(0.150)	(0.115)
Kernel - PSM	0.152	-0.042	-0.098	0.124	0.055	-0.023	0.058	-0.351	-0.372	0.029	-0.260	-0.248
	(0.086)*	(0.110)	(0.122)	(0.060)**	(0.078)	(0.087)	(0.155)	(0.283)	(0.336)	(0.109)	(0.179)	(0.216)
	(0.093)	(0.116)	(0.129)	(0.064)*	(0.082)	(0.091)	(0.164)	(0.298)	(0.346)	(0.109)	(0.192)	(0.222)
	(0.086)*	(0.105)	(0.112)	(0.065)*	(0.075)	(0.077)	(0.139)	(0.266)	(0.302)	(0.093)	(0.170)	(0.206)
Nearest	0.320	0.182	0.206	0.184	0.141	0.087	0.351	-0.009	0.14	0.151	-0.083	0.054
Neighbour	(0.143)**	(0.146)	(0.188)	(0.102)*	(0.118)	(0.140)	(0.207)*	(0.309)	(0.315)	(0.141)	(0.186)	(0.196)
	(0.144)**	(0.249)	(0.156)	(0.090)**	(0.184)	(0.083)	(0.260)	(0.216)	(0.251)	(0.151)	(0.119)	(0.173)
	(0.129)**	(0.156)	(0.174)	(0.093)**	(0.110)	(0.118)	(0.212)*	(0.311)	(0.339)	(0.122)	(0.184)	(0.202)

Note: Alternative standard errors are reported in parentheses for any specification. CEM: robust standard errors (clustered by tax center), bootstrapped standard errors (clustered by tax center) based on 500 replications, and stratified bootstrapped standard errors based on 500 replications; Kernel - MHD and Kernel - PSM: bootstrapped standard errors based on 200, 500 replications and stratified bootstrapped standard errors based on 500 replications; Nearest Neighbour: heteroskedasticity-consistent analytical standard errors proposed by Abadie and Imbens (2006), wild bootstrapped standard errors based on 500 replications and stratified bootstrapped standard errors based on 500 replications; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Main Results – Audit Type (IPTW)

Dep. Variable	Taxable Income			CIT payable		
Years after audit	I	II	III	I	II	III
Type of Audit	(1)	(2)	(3)	(4)	(5)	(6)
Comprehensive	**0.285**	0.130	-0.040	**0.246**	0.136	0.030
	(0.162)*	(0.228)	(0.241)	(0.128)*	(0.185)	(0.161)
	(0.173)*	(0.216)	(0.193)	(0.135)*	(0.173)	(0.139)
	(0.173)*	(0.205)	(0.168)	(0.135)*	(0.162)	(0.132)
Desk Issue	0.020	**-0.235**	**-0.170**	0.006	**-0.095**	**-0.078**
	(0.030)	(0.066)***	(0.046)***	(0.026)	(0.047)**	(0.042)*
	(0.074)	(0.103)**	(0.107)	(0.041)	(0.062)	(0.063)
	(0.065)	(0.088)***	(0.094)*	(0.036)	(0.053)*	(0.058)

Note: Alternative standard errors are reported in parentheses. In particular we report robust standard errors (clustered by tax center), bootstrapped standard errors based on 500 replications and stratified bootstrapped standard errors based on 500 replications; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Conclusions

- **Tax audits in Rwanda deliver sizeable pro-deterrence effects on future reporting behaviour**
  - Taxable income declared by audited firms one year after the process increases by **20.7%** (CIT by **12.3%**).
  - This corresponds to approximately **2.8%** of total CIT declared by all CIT businesses in 2016.
  - Noncompliant drive the results.
- **... but there are also margins that can improve the performance of tax enforcement policy:**
  - Comprehensive audits drive the pro-deterrence impact.
  - Narrow-scope audits have counter-deterrence effect after 2 years (**-23.5%** on TI, **-9.5%** on CIT).
  - these results are consistent with those provided by Erard, Kirchler and Olsen (2019) for US.
- **Several** robustness analyses corroborate these results.

Thank you!

@TARC2013
@KotsogiannisC
@LVSalvadori

UNIVERSITY OF EXETER | BUSINESS SCHOOL

TARC
Tax Administration Research Centre

UKRI Economic and Social Research Council

References I

Abadie, Alberto and Guido W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74(1):235–267.

Abadie, Alberto and Guido W. Imbens. 2008. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica* 76(6):1537–1557.

Advani, A., W. Elming and J. Shaw. 2019. "The Dynamic Effects of Tax Audits." Warwick Economics Research Paper, 1198.

Beer, S., M. Kasper, E. Kirchler and B. Erard. 2020. "Do Audits Aeter Future Noncompliance? Evidence On Self-Employed Taxpayers." *CesIfo Economic Studies,* 66(3):248—-264.

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

References II

Bodory, Hugo, Lorenzo Camponovo, Martin Huber and Michael Lechner. 2020. "The Finite Sample Performance of Inference Methods for Propensity Score Matching and Weighting Estimators." *Journal of Business & Economic Statistics* 38(1):183–200.

DeBacker, Jason, Bradley T. Heim, Anh Tran and Alexander Yuskavage. 2015. "Legal Enforcement and Corporate Behavior: An Analysis of Tax Aggressiveness after an Audit." *Journal of Law and Economics,* 58(2):291–324.

DeBacker, Jason, Bradley T. Heim, Anh Tran and Alexander Yuskavage. 2018*a*. "Once Bitten, Twice Shy? The Lasting Impact of Enforcement on Tax Compliance." *Journal of Law and Economics,* 61(1):1–35.

DeBacker, Jason, Bradley T. Heim, Anh Tran and Alexander Yuskavage. 2018*b*. "The Effects of IRS Audits on EITC Claimants." *National Tax Journal,* 71(3):451–484.

References III

Erard, B., Erico Kirchler and Jerome Olsen. 2019. Audit impact study: The specific deterrence implications of increased reliance on correspondence audits. In *Taxpayer Advocate Service: Annual Report to Congress 2019*. pp. 257–268.

Gemmell, Norman and Marisa Ratto. 2012. "Behavioral Responses to Taxpayer Audits: Evidence From Random Taxpayer Inquiries." *National Tax Journal,* 65(1):33–57.

Iacus, Stefano M., Gary King and Giuseppe Porro. 2019. "A Theory of Statistical Inference for Matching Methods in Causal Research." *Political Analysis* 27(1):46–68.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen and Emmanuel Saez. 2011. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." *Econometrica,* 79(3):651–692.

References IV

Lediga, Collen, Nadine Riedel and Kristina Strohmaier. 2020. "Tax Enforcement Spillovers – Evidence from South Africa.". Mimeo.

Løyland, Knut, Oddbjørn Raaum, Gaute Torsvik and Arnstein Øvrum. 2019. "Compliance Effects of Risk-Based Tax Audits.". CESifo Working Paper, 7616.

Pomeranz, Dina. 2015. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax." *American Economic Review,* 105(8):2539–2569.

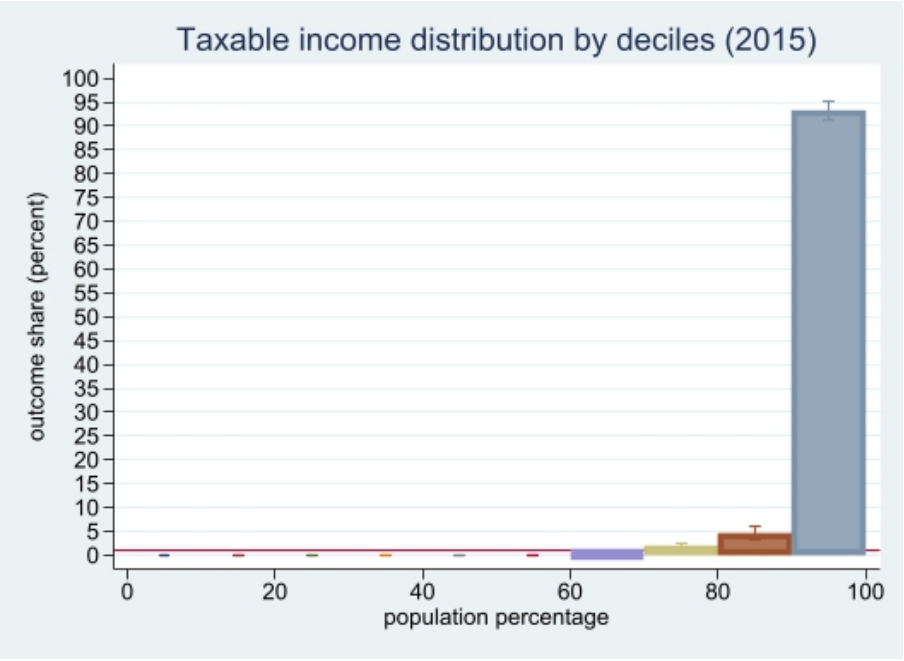
UN-DESA. 2016. "The Millennium Development Goals Report 2015.". https://doi.org/10.18356/6cd11401-en.

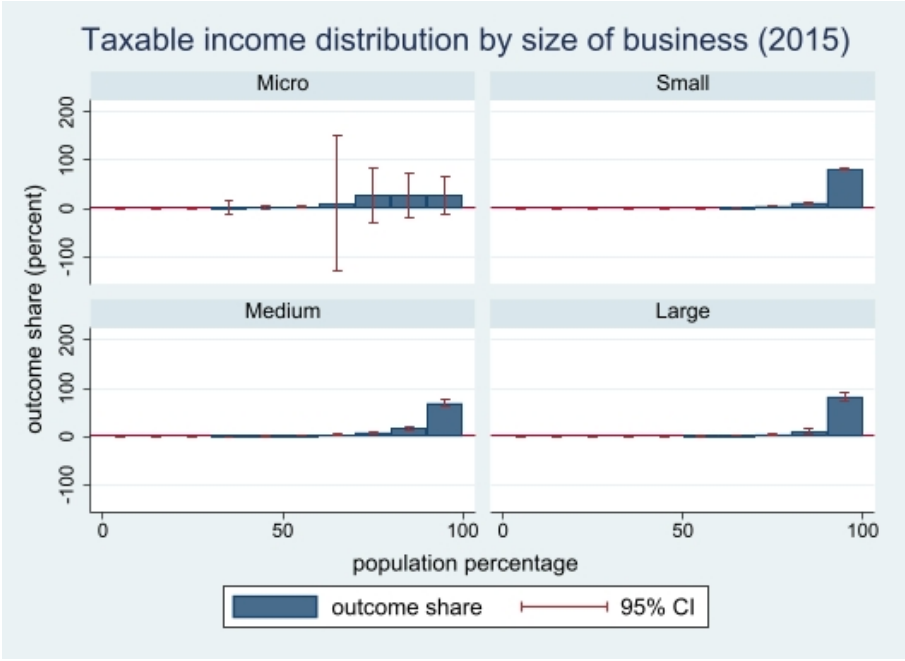
Wooldridge, Jeffrey M. 2002. "Inverse probability weighted M-estimators for sample selection, attrition, and stratification." *Portuguese Economic Journal* 1(2):117–139.

References V

Wooldridge, Jeffrey M. 2007. "Inverse probability weighted estimation for general missing data problems." *Journal of Econometrics* 141(2):1281–1301.

More on CIT data



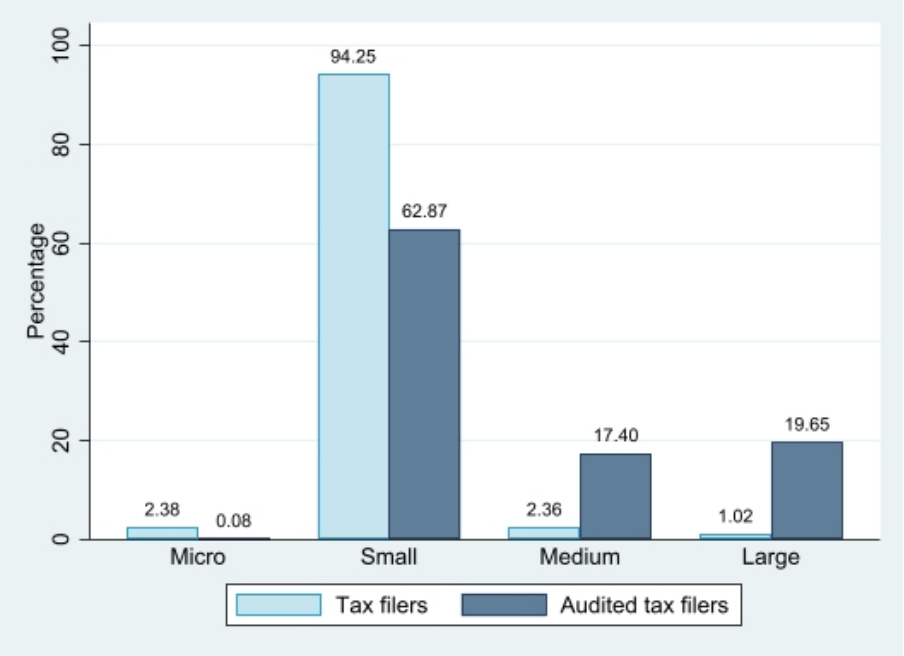


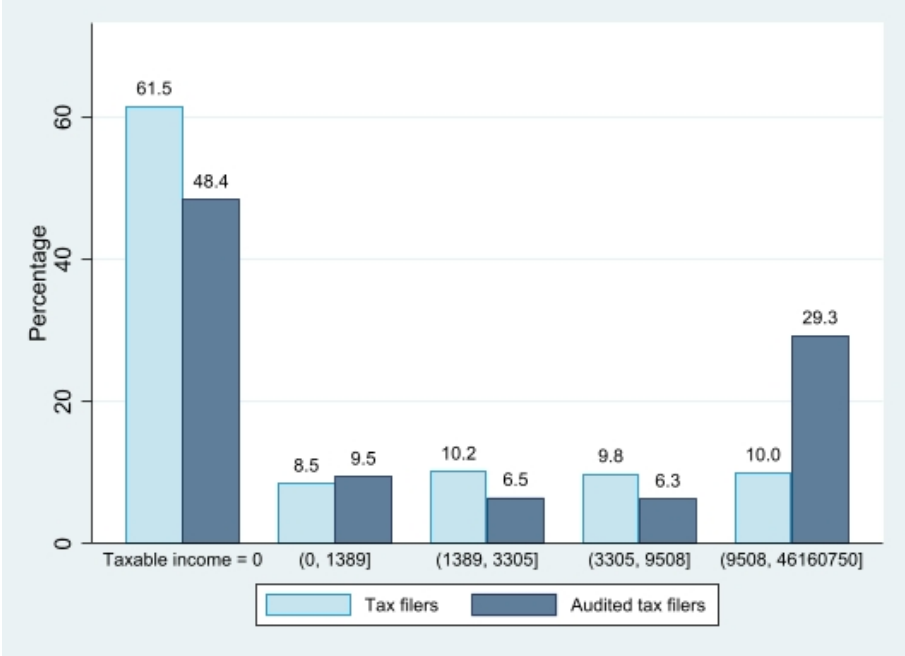
Note: Authors' calculations based on data provided by RRA

- Firms in the tenth decile report more than 90% of taxable income (left-hand-side panel).
- The majority of reported income across firm type is reported by the top deciles of their corresponding distribution (right-hand-side panel).

Back to main CIT data.

More on Audit data





Note: Authors' calculations based on data provided by RRA

- Left-hand-side panel reports the distribution of audits by size and firms by size.
- Right-hand-side panel shows the distribution of firms and audits by deciles of taxable income.

Back to main audit data.

Empirical Strategy: balance performance

Panel A: Overall imbalance, Multivariate L_1	
L_1 statistic pre CEM:	0.61
L_1 statistic post CEM:	0.28

Panel B: Univariate imbalance

	L_1 pre CEM	L_1 post CEM
Aggregate Risk Score	0.48	0.12
Taxable income 2013	0.14	0.08
Taxable income 2014	0.19	0.07
Taxable income 2015	0.18	0.06

Note: The table depicts L_1 statistics for multivariate and univariate imbalance as defined in Iacus et al. (2011). Back to main imbalance.

Inference I

Still substantially debated issue in this context:

- Standard bootstrapping usually applied but not generally justified:
  - Valid for Kernel-based methods (asymptotically linear) (Bodory et al., 2020; Abadie and Imbens, 2008)
  - Not valid for Nearest-Neighbour (Abadie and Imbens, 2008); Abadie and Imbens (2006) provide heteroskedasticity-consistent analytical solution; wild bootstrapping is also justified (Bodory et al., 2020)
- Wooldridge (2007, 2002) has shown that ignoring the first-stage estimation of the selection probabilities when performing inference yields to more conservative standard errors for IPTW.
- Iacus, King and Porro (2019) argue that when ex-ante stratification solutions are employed (as, for example, for CEM) these concerns are misplaced and unaltered regression standard errors are correct.

Inference II

- Given these premises, we provide inference by reporting alternative SEs for any specification:
  - **CEM and IPTW:** robust SEs (clustered by tax center), bootstrapped SEs (clustered by tax center) based on 500 replications
  - **Kernel PSM and Kernel MHD:** bootstrapped SEs (based on 200 and 500 replications)
  - **Nearest-neighbour MHD:** heteroskedasticity-consistent SEs proposed by Abadie and Imbens (2006), wild bootstrapped SEs based on 500 replications
  - **For all specifications:** given CEM preprocessing, we additionally report stratified bootstrapped SEs (based on 500 replications and CEM strata).

Back to main results.

Robustness checks

Several additional sensitivity analyses are performed to test the robustness of the findings. We follow two main avenues:

1. Regression specifications for the outcome variables controlling for residual imbalance:
   - weighted regression models based on the weights calculated with our baseline models;
   - double-robust regression adjustment models (IPW-RA).
2. Stricter selection of the matched sample through the CEM stratification by employing two alternative less parsimonious sets of matching variables for our baseline models.

The results corroborate our main findings. here Back to conclusions.

Sensitivity analysis I

Weighted regression models

Dependent Variable	Taxable Income			CIT payable		
Years after the audit	I	II	III	I	II	III
Matching estimator	(1)	(2)	(3)	(4)	(5)	(6)
CEM	0.315***	0.246	0.253	0.210**	0.232	0.184
	(0.095)	(0.197)	(0.212)	(0.091)	(0.151)	(0.169)
Kernel - MHD	0.279***	0.020	0.051	0.173***	0.056	0.031
	(0.094)	(0.126)	(0.127)	(0.054)	(0.081)	(0.084)
Kernel - PSM	0.191*	-0.029	-0.063	0.137**	0.065	-0.004
	(0.104)	(0.135)	(0.136)	(0.067)	(0.090)	(0.097)
Nearest Neighbour	0.525***	0.353	0.419	0.298**	0.245	0.258
	(0.169)	(0.291)	(0.369)	(0.142)	(0.225)	(0.279)

Note: Standard errors [(1) of main table] are reported in parentheses. Covariates: the risk score assigned to the taxpayer each of the three years before treatment, the taxable income reported in 2014 and 2013, the VAT paid on inputs reported each of the three years before treatment, a set of indicator variables for the tax centre, the sector of activity and the finer classification of the section of activity (according to the ISIC classification), dummies for diverse type of income reported each of the three years before treatment and a dummy for CIT tax return reported after the deadline during the year of the audit process.

Sensitivity analysis II

Double-robust regression adjustment models

Dependent Variable	Taxable Income			CIT payable		
Years after the audit	I	II	III	I	II	III
Matching estimator	(1)	(2)	(3)	(4)	(5)	(6)
IPW-RA (set I)	0.141**	-0.003	-0.032	0.111*	0.092	0.042
	(0.071)	(0.191)	(0.137)	(0.058)	(0.140)	(0.103)
IPW-RA (set II)	0.115*	-0.047	-0.080	0.092*	0.052	-0.000
	(0.066)	(0.170)	(0.139)	(0.055)	(0.122)	(0.103)

Note: Robust standard errors (clustered by tax center) are reported in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Two sets of covariates are employed. Set I includes the risk scores for the latest two pre-treatment years, reported taxable income declared in the year before treatment and a dummy for the sector of activity. Set II also includes dummies for diverse type of income reported each of the three years before treatment, a dummy for CIT tax return reported after the deadline during the year of the audit process and a dummy identifying the three tax centers in Kigali.

Sensitivity analysis III

Double-robust regression adjustment models - Type of audits

Dependent Variable	Taxable Income			CIT payable		
Years after the audit	I	II	III	I	II	III
Type of Audit	(1)	(2)	(3)	(4)	(5)	(6)
Set I						
Comprehensive	0.384*	0.172	0.107	0.317**	0.172	0.155
	(0.167)	(0.226)	(0.274)	(0.133)	(0.194)	(0.211)
Desk Issue	0.019	-0.238***	-0.177***	0.005	-0.099**	-0.086**
	(0.029)	(0.065)	(0.045)	(0.028)	(0.048)	(0.041)
Set II						
Comprehensive	0.297**	0.127	0.115	0.250**	0.137	0.168
	(0.120)	(0.160)	(0.266)	(0.097)	(0.140)	(0.204)
Desk Issue	0.017	-0.231***	-0.170***	0.007	-0.093*	-0.080**
	(0.028)	(0.065)	(0.040)	(0.028)	(0.049)	(0.039)

Note: Robust standard errors (clustered by tax center) are reported in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Sensitivity analysis IV

Main Results – Aggregate *ATT* (using Set II of matching covariates)

Dependent Variable	Taxable Income			CIT payable		
Years after the audit	I	II	III	I	II	III
Matching estimator	(1)	(2)	(3)	(4)	(5)	(6)
CEM	0.296***	0.202	0.229	0.175***	0.160	0.133
	(0.058)	(0.176)	(0.146)	(0.047)	(0.132)	(0.114)
Kernel - MHD	0.279***	0.100	0.115	0.160***	0.088	0.072
	(0.086)	(0.103)	(0.109)	(0.057)	(0.071)	(0.072)
Kernel - PSM	0.198**	-0.131	-0.137	0.138**	-0.025	-0.059
	(0.085)	(0.111)	(0.121)	(0.060)	(0.080)	(0.082)
Nearest Neighbour	0.421***	0.265**	0.336**	0.260***	(0.179)**	0.187
	(0.133)	(0.116)	(0.158)	(0.098)	(0.080)	(0.115)

Note: Standard errors [(1) of main table] are reported in parentheses. Set II of matching covariates includes the initial set of control variables and dummies for the sector of activity (according to ISIC classification). The matched set of observations include 263 treated units (73%) and 4406 untreated units (40.6%). Multivariate imbalance measure before CEM equals 0.62 and after CEM reduces to 0.34 (55% of initial imbalance).

Sensitivity analysis V

Main Results – *ATT* by audit type (using Set II of matching covariates)

Dependent Variable	Taxable Income			CIT payable		
Years after the audit	I	II	III	I	II	III
Type of Audit	(1)	(2)	(3)	(4)	(5)	(6)
Desk Issue	0.094***	-0.183***	-0.132***	0.045**	-0.061	-0.063**
	(0.021)	(0.056)	(0.032)	(0.020)	(0.047)	(0.031)
Comprehensive	0.394***	0.223	0.006	0.329***	0.206	0.067
	(0.149)	(0.222)	(0.207)	(0.121)	(0.180)	(0.133)

Note: Robust standard errors (clustered by tax center) are reported in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Set II of matching covariates includes the initial set of control variables and dummies for the sector of activity (according to ISIC classification). The matched set of observations include 263 treated units (73%) and 4406 untreated units (40.6%). Multivariate imbalance measure before CEM equals 0.62 and after CEM reduces to 0.34 (55% of initial imbalance). Back to main robustness.