How long-lasting are the effects of audits?

Preliminary. Please do not cite without authors' permission

Arun Advani* William Elming† Jonathan Shaw*

December 18, 2015

Abstract

Understanding tax non-compliance and the effectiveness of strategies to tackle it is crucial for a modern tax authority. In this paper we study the indirect benefits of conducting audits, focusing on how the reported tax liability of audited individuals responds after an audit. We exploit data from a random audit program covering income tax self-assessment returns in the UK. We find that audits have a large impact on reported tax liability, between 8 and 17 per cent for those who remain in self-assessment, and this persists for at least 10 years after the audit. This is the same order of magnitude as the initial audit adjustment across all taxpayers. Future work will study how much of this affect might be attributable to differential attrition from self-assessment tax for the audited group.

*The Institute for Fiscal Studies (IFS), the Tax Administration Research Center (TARC) and University College London. This work contains statistical data from HMRC which is Crown Copyright. The research datasets used may not exactly reproduce HMRC aggregates. The use of HMRC statistical data in this work does not imply the endorsement of HMRC in relation to the interpretation or analysis of the information.

†IFS and TARC.

1 Introduction

The ongoing drive for a more efficient public sector increases the importance of understanding what determines tax non-compliance—and of having effective strategies to identify and recover unpaid revenue. Taxpayer audits are one widely-used example of such a strategy. Audits have a direct benefit in terms of additional revenue raised. There is, however, also potential for indirect benefits of audits.

Indirect effects take two forms: dynamic effects and spillover effects. Dynamic effects are changes in the future behaviour of the audited taxpayer. Spillover effects are changes in the behaviour of other taxpayers who know the audited taxpayer. These effects come from updated information about the probability of an audit, the effectiveness of an audit and the cost of an audit.

In this paper we study on the dynamic effects of an audit. We study the tax returns of individuals, and examine how they change over time after an audit. Historically, tax authorities such as the Internal Revenue Service (IRS) have primarily focused on direct revenue maximization in selecting tax returns for their examination program (see Bloomquist, 2013). However, since most taxpayers pay taxes for many years, it is potentially important to understand the long term effects these audits have. This paper therefore seeks to quantify the amount of additional revenue received from an individual taxpayer in the years after audit. Understanding this is crucial in determining the total return from an audit, and hence in determining the optimal extent of enforcement.

We exploit a random audit programme run by the UK tax authority (HMRC) that focuses on income tax self-assessment taxpayers. Not all income taxpayers have to submit a self-assessment return, only those with circumstances likely to mean Pay As You Earn (PAYE) doesn't withhold the right amount of tax. This includes the self-employed, high earners and individuals with capital gains or income from land or property (among others).[1] An average of around 2,800 individuals are selected for random audit each year, corresponding to a probability of 0.03 per cent (three in $10,000$). Taxpayers are selected with equal probability from among the population of self-assessment taxpayers. We have data on audits for fourteen years between 1996/97 and 2009/10 and income tax returns for twelve years between 1998/99 and 2011/12. This means we can track individuals for a substantial amount of time before and after audit.

As a control group, we use individuals who could have been selected for a random audit but weren't. Since some of those assigned to audit are not audited in practice – some are explicitly deselected (e.g. because they no longer meet the self-assessment criteria) and others are not started before the required deadline – the parameter we estimate is the intention to treat (ITT) parameter. It should be interpretted as the effect of being selected for random audit relative to a baseline of not being selected for random audit but facing the normal policy environment, which includes a small chance of a targeted audit.

[1]For mo;re details of who has to complete a self-assessment return, see https://www.gov.uk/self-assessment-tax-returns/who-must-send-a-tax-return.

Our main results are as follows. As expected, there are no substantial differences between treatment and control groups before the audit has taken place, suggesting we have constructed a valid control group. After the audit has taken place, there is a large and persistent impact of being selected for audit. The effect on reported tax liability is between 8 and 17 per cent, and it persists for at least 10 years after the audit. This is a sizable impact that corresponds to between about £500 and £1,600 in cash terms—so roughly the same order of magnitude as the initial audit adjustment across all taxpayers (£700). It is important to note, however, that this could be driven either by an extensive margin effect (e.g. low profitability audited taxpayers dropping out of self-assessment) or an intensive margin effect (audited taxpayers changing their reporting or real behaviour). If the former is responsible, then HMRC does not gain by the equivalent of £700 each year after the audit per taxpayer randomly selected for an audited—indeed HMRC could lose overall. If the latter is responsible, HMRC will gain overall. Distinguishing between extensive and intensive margin responses are therefore the next step in our future work.

The next section presents some of the related literature. Section 3 outlines the policy context, while section 4 describes the data and presents some summery statistics. Section 5 sets out the method used and discusses how we construct our control group. The results are presented in section 6. Section 7 concludes.

2 Related literature

Following the seminal paper by Allingham and Sandmo (1972) there has been a large literature that studies tax compliance, evasion and enforcement regimes (see Andreoni et al. (1998), Slemrod and Yitzhaki (2002) and Slemrod (2007) for more exhaustive reviews of the literature). In general, the literature can be divided into three subgroups: a literature that uses individual taxpayer (micro) data, a literature that relies on aggregate data, and finally a literature that uses laboratory experiments.

2.1 Micro data

Within the literature using micro data, a number of previous papers have investigated the deterrence effect of audits. Kleven et al. (2011) analyze the effect of a tax enforcement field experiment in Denmark. Half of a sample of income tax filers were randomly selected for audit, while the other half were not audited. The following year, letters threatening an audit were randomly assigned to tax filers in both groups. Declarations of self-reported and third-party reported income were then followed up. They find that prior audits and threat-of-audit letters have significant effects on self-reported income, but no effect on third-party reported income.

Gemmell and Ratto (2012) investigate behavioural responses to taxpayer audits using earlier versions of the random audit data we use. They distinguish between taxpayers found to be non-compliant and those found to be compliant, arguing that the former are likely to increase their subsequent compliance while the latter could reduce their compliance.[2] However, this distinction between compliant and non-compliant taxpayers is endogenous, making it hard to interpret the comparison with an unconditionally randomly selected control group as causal.

There are few other comparable studies of the behavioural effects of audits on future income reporting. Long and Schwartz (1987) find that audits in 1969 had no effect on average non-compliance in 1971 and only a small effect on the frequency of non-compliance. Erard (1992) uses IRS data to assess the effect of audits on taxpayers' compliance in the following year. His findings indicate that evaders increase the tax declared following an audit, though the results are not conclusive. The effects he finds cannot be solely attributed to the effect of the audit, and the results are sensitive to the way selection issues are dealt with in the estimation.

Tauchen et al. (1993) estimate the general deterrence effect of audits on the amount of income that taxpayers choose to report on their tax returns. They use a stratified random sample of approximately 50,000 individual tax returns (from the 1979 Taxpayer Compliance Measurement Program) and combine the returns with IRS administrative records for District Offices and 1980 Census data at the five-digit zip code level. They find weak evidence that higher audit rates are associated with greater levels of compliance. They find that a 10 per cent increase in audits lead to a 2.3 per cent increase in reported income for their high-income group, with no significant effects on the other income groups. Further they find that the general deterrence effects of audits are over $2 for every $1 of direct revenue yield.

Bergman and Nevarez (2006) use VAT tax return information and enforcement data to determine the effect of audits on subsequent compliance of taxpayers in Argentina and Chile. They find no evidence that audits increase individual compliance.

While the papers above focus on the effect of actual audits on reporting behaviour, Slemrod et al. (2001) and Agostini and Martnez (2014) studies the effect of tax authority letters containing *threat* of audits. Slemrod et al. (2001) compare the change in reported tax for 1,724 randomly selected Minnesota taxpayers, who received a letter with a threat of audit, relative to a control group that did not receive such a letter. They find that low and middle-income taxpayers in the treatment group on average increased tax payments compared to the previous year. The effect is only significant for those with more opportunities to evade taxes (i.e. those with self-employment or farm income). For the high-income treatment group, however, the conclusions are radically different; upon receiving the threat-of-audit letter they decrease their reported tax liability relative to the control group. Slemrod et al. (2001) argue that this is because higher income taxpayers see the audit threat as the beginning of a negotiation

[2]This could be the case if they are audited and found compliant, but they were in fact non-compliant.

over actual tax liability.

Agostini and Martnez (2014) study the impact of a tax enforcement programme implemented by the Chilean Internal Revenue Service, where letters requiring information about diesel purchases and use and vehicle ownership were sent to around 200 firms in 2003. They find that firms receiving a letter decreased their diesel tax credits by around 10 per cent.

2.2 Aggregate data

In addition to the literature that uses micro data, there is a literature that relies on aggregate data to examine the effects of audits. Dubin et al. (1990) uses the Annual Report of the Commissioner of Internal Revenue and the Statistics of Income for the years 1977-1986 to create a state-level data set. They set up three different linear equations where 'reported tax per return', 'returns filed per capita' and 'assessed tax liability per return' are the dependent variables. As explanatory variables they use the state income tax, the audit rate, per-capita income and other socioeconomic variables. They find that the audit rate had a significant and positive effect on reported tax and assessed liabilities per return. Hence they conclude that audits are an effective deterrent against taxpayer non-compliance. They then use the estimated linear relationships to calculate the counterfactual scenario where audit rates are kept constant at their 1977 level throughout the period. They find that had the audit rate remained constant at the 1977 level throughout the period from 1977-1986 then the total reported taxes would have been greater by 15.6 billion dollars in 1986 (corresponding to roughly 4 per cent of total reported tax).

Plumley (1996) uses a ten-year (1982-91) panel data set aggregated to the state level. He finds that the general deterrence effect of audits for the general population is about 11 times as large as the adjustments proposed by the audits themselves. Ali et al. (2001) use data from the Annual Report of the Commissioner of Internal Revenue Service and the Data Book for 1980 to 1995 to investigate the relationship between taxpayer compliance, audit rates and penalties if detected. They find that both the audit and penalty rate are effective deterrents of non-compliance. Further, they find that the effectiveness of these deterrents is increasing in income. Overall, they find that compliance increases with income, though at a decreasing rate.

Dubin (2007) follows the method of Dubin et al. (1990) and uses state-level data covering the years 1988-2001 to empirically test whether measurable activities of the IRS Criminal Investigation Division (CI) affect taxpayer compliance. He finds that CI activities have a measurable and significant effect on voluntary compliance and that incarceration and probation (rather than fines) have the largest effect.

2.3 Experimental data

Finally, a number of papers use experimental set-ups to test how different circumstances surrounding audits affect reporting behaviour. The papers by Alm and Mckee (2004), Fortin et al. (2007) and Alm et al. (2009) investigate how interactions among taxpayers affects the deterrence effects of audits. Alm and Mckee (2004) studies compliance behaviour when returns are selected for audit based on the deviation of each individuals tax report from the average reported tax of other taxpayers in their cohort. Thus, the optimal strategy of the participants in the experiment would be to cooperate and reach the zero compliance equilibrium, where no one is audited and all have the maximum possible gain.[3] They find that participants struggle to reach this equilibrium, but that pre-game communication facilitates coordination.

In a similar spirit, Fortin et al. (2007) studies the impact of social interactions on tax evasion using experimental data. Their experimental results provide evidence of fairness effects. Specifically, for a given gross income and personal tax rate, the individual will report less if they feel they are being treated unfair by the tax system (e.g. if they pay a higher tax rate than other participants in their group). Alm et al. (2009) investigate the effects of information dissemination concerning enforcement and compliance behaviour of others on the tax reporting behaviour of individual taxpayers in laboratory experiments. They find that taxpayers will respond to wide ranging information sources that report the enforcement effort and estimate the general deterrence effect on compliance to be 4.4 times the direct effect.

While the experimental papers above study the interactive effects of audits the papers of Friedland et al. (1978), Kirchler et al. (2007) and Kastlunger et al. (2009) examine the intertemporal effects of audits. Friedland et al. (1978) investigate the tax evasion behaviour of 15 participants in a game-simulation context. They find that large fines tend to be more effective deterrents than frequent audits. Kirchler et al. (2007) investigate the effectiveness of audit probabilities and sanctions in a dynamic setting focusing on the time lag between audits. They find that compliance decreases immediately after an audit, suggestive of a 'bomb crater effect'.[4] Contrary to Friedland et al. (1978), Kirchler et al. (2007) find that the effect of sanctions on compliance to be relatively less important than higher audit probabilities. They do, however, find that larger sanctions are associated with the tendency of participants to repair their losses following and audit.

Kastlunger et al. (2009) use a laboratory experiment where participants file taxes 60 times (i.e. participants' tax lifecycle). They then subject participants to different patterns of audits in two different studies. The first study focuses on the immediate reaction in reported income following an audit. Similarly to

[3]In some of their experiments the audit rule based on differences in the reported tax from the mean is augmented with a random audit rule such that there is a positive probability of audit regardless of the reporting strategy employed.

[4]The 'bomb crater effect' refers to the idea that individuals might perceive the risk of being audited to fall immediately after an audit. The name originates from preference of WW1 soldiers to hide out in bomb craters, believing that it was unlikely that a bomb would strike the exact same place again (see Mittone, 2006). A competing explanation for the decline in reported tax following an audit is the mechanism of loss repair: experiencing an audit may make taxpayers "want to evade more in the future in an attempt to 'get back' at the tax agency" (p. 844 in Andreoni et al., 1998).

Kirchler et al. (2007) they find a strong decrease in compliance in the period following the audit. Further they find that this is most likely to be the 'bomb crater effect' rather than loss-repair tendencies. The second study investigates the effect of the timing of the audits in participant's tax lifecycle. They find that early, rather than late, audit experiences in participants tax lifecycle lead to increased compliance.

Lastly, Choo et al. (2013) investigate the compliance behaviour of 92 self-assessed taxpayers. They find that increasing the audit rate had no significant effect on the number of people attempting to evade, and no significant effect on the amount of evasion among those who did evade. Further, the main finding of their experiment is that the compliance levels by self-assessed taxpayer participants are extremely high and non-responsive to changes in audit rates. Further, post-experiment questionnaire data shows compliant participants are driven by strong norms of honesty. In contrast, non-compliant participants are driven by profit maximisation. Their results suggest that the deterrence power of random audits is quite limited. Finally, they do not find any evidence that compliance levels drop immediately after an audit (the 'bomb crater effect'), neither for compliant or non-compliant participants. This is in contrast to the experimental evidence from Kirchler et al. (2007) and Kastlunger et al. (2009), which used student participants.

3 Policy context

Income tax is the largest of all UK taxes, contributing 26.2 per cent of total government receipts in 2012-13. Most (but not all) sources of income are subject to income tax, including earnings, retirement pensions, income from property, interest on deposits in bank accounts, dividends, and some benefits. Income tax is levied on an individual basis and operates through a system of allowances and bands. Each individual has a personal allowance, which is deducted from total income. The remainder – taxable income – is then subject to a progressive schedule of tax rates.

A total of around 30 million individuals in the UK pay income tax each year. Out of these, around 8 million are subject to self-assessment and required to submit a tax return. These tend to be individuals with forms of income not subject to withholding or for whom the tax system struggles to calculate and withhold the right amount of tax. It includes self-employed individuals, those with very high incomes, company directors, landlords and many pensioners.

Since incomes covered by self-assessment tend to be harder to verify, there is a significant risk of non-compliance. As a result, HMRC carries out audits each year to deter non-compliance and recover lost revenue. HMRC runs two types of income tax audits. Targeted audits are based on perceived risks of non-compliance. Random audits are used to ensure that all self-assessment taxpayers face a positive probability of being audited, as well as to collect information about the scale of non-compliance and predictors of non-compliance. The Risk department of HMRC performs the selection of taxpayer returns into the two audit categories. The lists of taxpayers to audit are then passed on to local 'compliance teams', where compliance officers perform the audits. The distinction between these two arms of HMRC is important, as not all cases selected for audit by the Risk team are actually audited by the compliance teams, typically due to resource constraints.

The timeline for the audit process is as follows. The tax year ends on 5 April. In April or May, HMRC issues a notice to file to taxpayers who they believe need to submit a tax return.[5] Cases to be subjected to a random audit are provisionally selected from the population of individuals issued with a notice to file. The main filing deadline for taxpayers is 31 January the following calendar year. This is the date by which individuals must submit their tax return. HMRC then deselects some cases from random enquiry.[6] At the same time, targeted audits are selected on the basis of the information provided in self-assessment returns and other intelligence. Individuals selected for a random audit are selected before decisions about targeted audits are made, and as a random audit is worked in exactly the same way as targeted full audits, they cannot be selected for a targeted audit in the same tax year. Audits must be opened within a year of the date when the return was filed (until 2007/08, it was a year from the 31 January filing deadline for returns filed on time). Taxpayers subject to an audit are informed of this when the audit is opened but they are not told whether it is a random or targeted audit.

Table 1 shows the average number of cases HMRC have selected for income tax audits per year over the period 1996/97 to 2009/10. HMRC selected an average of 3,150 cases for random audits and 162,553 cases for targeted audits per year. The vast majority of these are audits of individuals (as apposed to partnerships and trusts). For individuals, the corresponding probabilities of being selected for an audit are .03 per cent (three in 10,000) for random audits and 1.8 per cent for targeted audits.

Table 1: Average number of cases selected for income tax audits

	Random	Random audit probability	Targeted	Targeted audit probability
Individuals	2,827	.0003	152,585	.0176
Partnerships	261		8,172	
Trusts	61		1,797	
Total	3,150		162,553	

Notes: Annual averages for tax years 1996/97 to 2009/10. Data for audits is not available after 2009/10. Source: calculations based on HMRC administrative datasets.

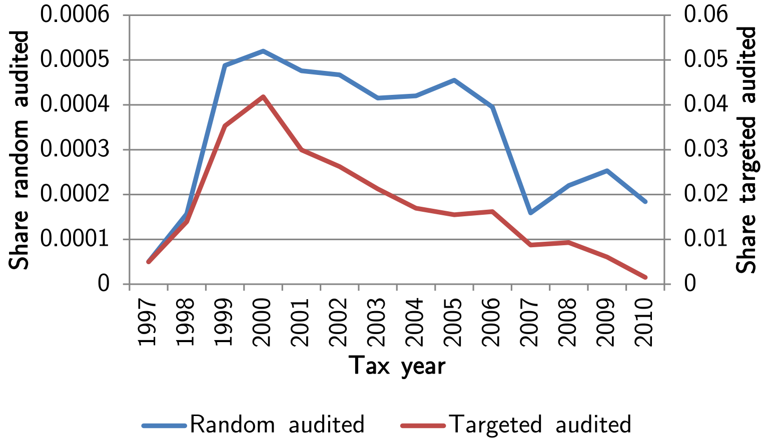
Figure 1 shows that there has been considerable variation in the audit probability for individuals over time. Audit rates peak in 1999/2000 at .05 per cent

[5]This excludes most first-time filers, which HMRC doesn't yet know about.

[6]This deselection implies that we cannot simply compare 'randomly' audited individuals to non-audited individuals.

(five in 10,000) for random audits and 4.2 per cent for targeted audits, and then fall back substantially. The latest publicly-available estimates from the random audits programme described in HMRC (2013) suggest that 27 per cent of self-assessment taxpayers under-declared their tax liabilities in 2009/10.

Figure 1: Number of individuals selected for income tax audits on



Source: calculations based on HMRC administrative datasets.

4 Data

4.1 Data Sources

We exploit data on self-assessment random audits together with information on income tax returns. This combines a number of different HMRC datasets, linked together on the basis of encrypted taxpayer reference numbers and tax year.

Audit records come from CQI (Compliance Quality Initiative), an operational HMRC dataset that records audits made into income tax self-assessment and corporation tax self-assessment returns. It does not include audits by the HMRC's Large Business Service, Special Investigations or Employer Compliance Reviews. It includes operational information about the audits, such as start and end dates, and audit outcome (size of any correction, penalties and interest). There are also limited details about the taxpayer. There are around 50 variables in the data made available to us and, in the main table, there is one observation for each audit. We have CQI data covering audits for tax years 1996/97 to 2009/10.

We track individuals following the audit using information from tax returns. It is important to recognise that we have no way to identify actual compliance behaviour in the follow-up (i.e. subsequent to the initial audit). The number of random audit taxpayers that are re-audited is far too small for it to be possible just to focus on them.

Information from tax returns comes from two sources: SA302 and Valid View. The SA302 dataset contains information that is sent out to taxpayers summarising income and tax liability (SA302 forms). It is derived from self-assessment tax returns, which are put through a tax calculation process. There are a total of around 150 variables but the range of information it contains is relatively limited, e.g. there is no information about turnover and expenses for self-employed individuals, demographics or filing data. For each tax year, it contains one observation per self-assessment taxpayer. We currently have access to it for tax years 2004/05 to 2011/12.

SA302 is supplemented by variables drawn from Valid View, a dataset that provides information taken more directly from individuals' tax returns. Valid View includes variables relating to detailed income sources and tax liabilities, some demographics and filing information, but few income and tax totals of the form found in SA302. The dataset was assembled for internal HMRC analytical purposes and contains about 700 variables. For each year, there is one observation per taxpayer. We have access to it for tax years 1996/97 to 2011/12.

SA302 and, to a lesser extent, Valid View are both updated during the course of the tax year as new tax returns come in. The versions we are using correspond to the October following the end of the tax year to which the extract relates.

4.2 Descriptives

Table 2: Lag in months between tax return filing and audit start

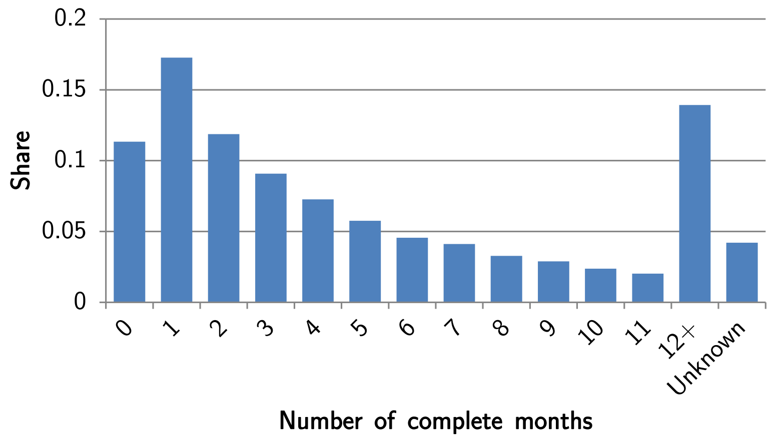
	Mean	Std. dev.
Random audits	8.6	4.0
Targeted audits	10.2	6.1

Notes: Annual averages for tax years 1996/97 to 2009/10. Includes all individuals under income tax self-assessment. Source: calculations based on HMRC administrative datasets.

Currently, if HMRC wants to audit a tax return, it is required begin the audit within twelve months of the return being filed. Table 2 shows the average lag, in months, between the date the tax return was filed and the date at which the audit was started. Random audits begin more quickly, after 8.6 months, compared with targeted audits at 10.2 months. Figure 2 shows the distribution of random audit durations. The majority take less than six months, but there is a long tail of audits lasting a year or more. This substantial lag between when an audit is opened to when it is settled is important to keep in mind when interpreting the result later on.

Table 4.2 summarizes the outcomes of random audits on individuals. More than half of all returns are found to be correct. Of those which are found to be incorrect, one-quarter do not involve any underpayment of tax, for example incorrectly categorising valid expenses, whilst the remaining three-quarters are 'non-compliant', and have a tax underpayment. Among the non-compliant, the

Figure 2: Distribution of audit durations



Source: calculations based on HMRC administrative datasets.

average additional tax owed is £1,950. However, as with income, the distribution is skewed: 65 per cent of non-compliant individuals owe additional tax of less than £1,000, whilst a small fraction (3 per cent) owe more than £10,000.

Table 3: Random audit outcomes

	Mean	Std. dev.
Proportion of audited returns deemed		
Correct	.529	.499
Incorrect (no underpayment)	.112	.316
Non-compliant (underpayment)	.359	.480
Mean additional tax if non-compliant (£)	1,950	8,004
Distribution of additional tax if non-compliant		
Share £1-100	.145	.353
Share £101-1,000	.495	.500
Share £1,001-10,000	.328	.470
Share £10,001+	.031	.172

Notes: Annual averages for tax years 1996/97 to 2009/10. Includes all individuals who faced a random audit. *Source:* calculations based on HMRC administrative datasets.

One approach we outline below exploits the randomisation in the random audit programme to select a suitable control group. For this to be a valid approach, it is crucial that random audits have been selected randomly. Table 4 tests for this by carrying out a set of balancing tests that check whether outcomes in the treatment and control groups are balanced with each other in the run-up to the audit. The outcomes we do this for are income and tax totals, income components and characteristics between five years and one year before the audit. These tests show a very good degree of balancing. Income and tax totals are balanced at all horizons (there are no statistically significant differences, indicated by asterisks on the difference parameters). Among income components, income from share schemes is marginally significantly different four years before audit. All other income components and horizons are not significantly different across treatment and control groups. For characteristics, age is significantly different one year before audit; otherwise all other characteristics are balanced. This extent of differences is roughly what we would expect given the number of tests undertaken and we therefore conclude that our treatment and control groups are comparable and that the randomization has worked.

5 Method

5.1 Parameters of interest

For each audited taxpayer we observe the amount of tax paid in year t, $\mathbb{E}[Y_t]$. If $D(t)$ is the event that a taxpayer was audited in year t, then we can define the tax paid h years after audit as $Y_{t+h}|D(t)$. There are three questions we wish to answer: (i) how does an individual's tax return change in response to receiving an audit; (ii) how does this vary with the probability that the individual is non-compliant ('cheating'); and (iii) how does the effect of an audit vary with whether an individual is *found* to be cheating.

We define $C \in 0, 1$ as the event that a taxpayer is cheating on their taxes.[7] $\Pr(C|\mathbf{X}, Y) \equiv p(C)$ is then the probability that a taxpayer is cheating, given their observed characteristics, $\mathbf{X}$, and their current tax declaration, Y. Characteristics $\mathbf{X}$ include demographics, such as age, gender, and region of residence; the history of income and tax declarations; and potentially summary statistics of the population distribution of income and tax declarations (since deviations from this are likely to be an important flag of something unusual taking place).

We can now define the quantities of interest. We first want to know the average effect of an audit on the expected tax paid h years after audit:

$$\beta_h^{ATE} := \mathbb{E}[Y_{t+h}|D(t) = 1] - \mathbb{E}[Y_{t+h}|D(t) = 0]$$

This parameter tells us how much additional tax HMRC should expect to receive h years after audit if they conduct an audit on a taxpayer uniformly randomly selected from the population of existing taxpayers.

We next want to know how this varies with the estimated probability an individual was cheating. For an individual whose predicted probability of cheating was p this is:

$$\beta_h^{MTE}(p) := \mathbb{E}[Y_{t+h}|D(t) = 1, \hat{p}(C) = p] - \mathbb{E}[Y_{t+h}|D(t) = 0, \hat{p}(C) = p]$$

[7]More accurately, C is the event that, if audited, HMRC would find something in the taxpayer's return that is deemed non-compliant.

Table 4: Balancing test of income totals, income components and characteristics

Income and tax totals						
		-5	-4	-3	-2	-1
Total taxable income	Diff	721.5	348.2	-293.7	660	576.2
and capital gains	p-value	0.651	0.418	0.723	0.124	0.646
Total income tax and	Diff	247.8	78.1	-171.7	196.5	118.9
capital gains tax liability	p-value	0.327	0.36	0.699	0.076	0.837

Income components						
		-5	-4	-3	-2	-1
Employment	Diff	-114.2	-269.5	-582.3	552.1	598.0
	p-value	0.163	0.205	0.572	0.156	0.236
Self-employment	Diff	311.6	270.2	188.1	138.2	172.4
	p-value	0.135	0.364	0.632	0.516	0.465
Pension	Diff	208.7	220.6	195.7	186.2	188.1
	p-value	0.257	0.135	0.138	0.075	0.059
Share schemes	Diff	-189.1	79.3*	-57.5	-25.6	-101.9
	p-value	0.980	0.039	0.995	0.994	0.675
Property income	Diff	-52.1	-62.7	-22.6	-4.3	-3.0
	p-value	0.624	0.344	0.265	0.239	0.771
Trust and estates	Diff	-15.3	4.0	-10.6	-28.5	-30.4
	p-value	0.999	0.401	0.452	0.928	0.991
Intrest	Diff	13.0	37.8	11.1	-22.2	-21.6
	p-value	0.943	0.932	1.000	0.944	0.916
Dividends	Diff	-357.0	-28.6	213.2	136.0	190.1
	p-value	0.946	0.984	0.453	0.994	0.769
Other	Diff	-3.7	-16.6	-5.5	-2.0	7.6
	p-value	0.897	0.951	0.805	0.565	0.901

Characteristics						
		-5	-4	-3	-2	-1
Female	Diff	-0.01	-0.007	-0.005	-0.005	-0.005
	p-value	0.069	0.199	0.515	0.465	0.556
Address abroad	Diff	0.000	0.000	-0.001	0.000	0.001
	p-value	0.531	0.532	0.467	0.299	0.219
Has agent	Diff	-0.004	-0.005	-0.006	-0.002	-0.001
	p-value	0.070	0.082	0.271	0.695	0.91
Age	Diff	0.193	0.298	0.237	0.200	0.24**
	p-value	0.495	0.204	0.253	0.267	0.005
Region	Diff	-0.020	0.003	0.000	0.003	0.011
	p-value	0.565	0.458	0.261	0.606	0.461

Notes: 'Diff' parameters are the coefficient on a treatment dummy in a regression of the outcome of interest on tax year dummies and a treatment dummy. P-values are derive from a test that interactions between treatment and tax year dummies are all zero in a regression of the outcome of interest on tax year dummies and the interaction between treatment and tax year dummies. Standard errors are clustered by taxpayer. * $p < .05$, ** $p < .01$, *** $p < .001$. *Source:* calculations based on HMRC administrative datasets.

This parameter tells us how much additional tax HMRC should expect to receive h years after audit if they select for audit a taxpayer whose predicted probability of cheating it p. Taking a weighted average of these parameters at different values of p also allows us to calculate, for example, the return of extending targeted audits to all taxpayers with predicted probability of cheating in p_0, p_1. Formally, if $f(p)$ denotes the probability density function for the probability of cheating, we can calculate the additional return (h years after audit) of extending the targeted audit program as $\int_{p_0}^{p_1} \beta_h^{MTE}(p) f(p)\,\mathrm{d}p$. With this notation we can also usefully define the average treatment effect parameter as $\int_0^1 \beta_h^{MTE}(p) f(p)\,\mathrm{d}p$.

Finally we would like to calculate these parameters separately depending on whether an individual is or is not found to be cheating:

$$\beta_h^{MTE}(p,c) := \mathbb{E}\left[Y_{t+h}|D(t)=1,\hat{p}\left(C\right)=p,C=c\right] - \mathbb{E}\left[Y_{t+h}|D(t)=0,\hat{p}\left(C\right)=p,C=c\right]$$

This allows us to understand whether the affect of an audit on tax paid varies depending on the result of the audit. For example, one might find that among those with low values of $p(C)$, being audited and not found cheating ($C = 0$) might nevertheless increase future tax payments since these individuals become worried about being audited again, and so choose not to claim as many allowances. Conversely, those with a high value of $p(C)$ who are not found to be cheating ($C = 0$) might in the future reduce their payments since they realise any misstatements were not noticed. Such results are important for determining the scrutiny with which an audit should take place, and whether follow up audits in future years might be helpful.

5.2 The "could have" control group

To answer our first two questions we compare a treatment group of all individuals who are selected for a random self-assessment tax audit in year t ($Z(t) = 1$), to a control group of individuals who *could have* been selected for a random audit, but who weren't ($Z(t) = 0$).

A simple comparison of the average tax paid between the two groups, gives us the average effect of being *assigned* to the audit group (sometimes described as the 'intention to treat' parameter):

$$\beta_h^{ITT} = \mathbb{E}\left[Y_{t+h}|Z(t) = 1\right] - \mathbb{E}\left[Y_{t+h}|Z(t) = 0\right]$$

If compliance were perfect, so all individuals in the treatment group and none in the control group were audited, then this would be equal to the average treatment effect parameter, β_h^{ATE}. However, there are four reasons why these parameters differ: (i) targeted audits; (ii) deselections; (iii) incomplete auditing; and (iv) differential attrition.

Firstly, HMRC's targeted audit program selects for audit individuals with a high probability of cheating. Hence even in our control group there are few people who have $p(C)$ close to 1 who are not audited. To the extent that those with a high probability of cheating have a systematically different response to audit, the intention to treat parameter will differ from the desired average treatement effect. For example, if these individuals are likely to increase their tax payments once audited by more than the average audited individual, the estimated β_h^{ITT} will be a lower bound for β_h^{ATE}.

A second problem is that some taxpayers are deselected for audit after being assigned to the treatment group, because they are no longer required to submit a tax return. This occurs because returns are selected on the basis of the letters sent out to taxpayers asking them to file a return. Where the taxpayer no longer needs to submit a return, for example because they no longer have an income or because they have died, no return will be received. These cases account for around 15% of all selected taxpayers. If a control group of taxpayers were selected at the same time as the treatment group were defined, the parameter estimated from comparing the received returns among the assigned groups might differ from the ATE because the probability of deselection might vary with past individual characteristics. Without such a control group, there is an additional problem that even before the audits the two groups may now differ because the probability of deselection means that the treated group are not a uniform draw from taxpayers who submitted returns (though they are from taxpayers asked to file returns).

A third problem is that not all of the taxpayers who are selected for audit actually receive an audit. After being selected to receive a random audit, taxpayers' details are passed to local compliance offices, where officers are given the random audits (along with the other audits) to do. Whilst they are informed that these cases are considered high priority, they always receive more cases than they can complete and so some of these cases (around 10%) are not audited. If officers are selective in the order in which they work on cases, then it is likely that the cases they leave until last are the ones where they expect the lowest returns from audit, which would suggest that (in the absence of the earlier issues) the estimated β_h^{ITT} will be an upper bound for β_h^{ATE}.

These first three issues are all cases of (static) selection problems, where selection into *receiving* an audit is not uniformly random across the population of self-assessment taxpayers. There are two solutions available to us for this class of problems. The first is to estimate marginal treatment effects, and reweight these. The second is to try to bound the average treatment effect.

Estimating marginal treatment effects, which are of interest in themselves, can be done under two assumptions. First we assume that the unaudited (potential) tax declared, $Y^{(0)}$, is independent of audit status once we account for the probability of cheating, $Y_{t+h}^{(0)} \perp\!\!\!\perp D(t)|p(C)$. Since we observe the same data the HMRC use to estimate this probability, for any given probability of cheating p, which individuals are audited is random. The reason that some individuals are audited and others are not depends on the assignment to receive a random audit and on the particular resource constraints faced by compliance officers at the time. It is reasonable to assume these events are independent from the taxpayers' tax declaration decisions. Second, we require that the probability of receiving an audit is neither zero nor one for taxpayers at all values of $p(C)$. This assumption is testable, and is necessary for us to recover the full distribution of MTEs. If these assumptions hold, we can compare the tax paid h years after audit between audited and non-audited individuals with the same probability p of cheating. This difference will identify the MTE for $p(c) = p$. We can then recover the ATE using the definition of the ATE in the previous subsection, by summing up the MTEs, weighting them by the share of the population who have that probability of cheating.

When there are some values of $p(C)$ for which all individuals are audited (or not audited), we cannot recover the MTE for such individuals. An alternative approach is then to bound the ATE (for each horizon h) by making assumptions about the MTE for those individuals who are not observed. For those who are unlikely to have been cheating, a lower bound estimate is likely to be that they do not make any additional payment, whilst an upper bound would perhaps be to assume that the MTE for these individuals is the same as for the lowest probability individuals who are observed. Formally, if $\underline{\mathrm{p}}$ is the lowest value of $p(C)$ for which the probability of audit is not deterministic, then we might bound the MTEs for all $p < \underline{\mathrm{p}}$ as $\beta_h^{MTE}(p) \in \left[0, \beta_h^{MTE}(\underline{\mathrm{p}})\right] \quad \forall p < \underline{\mathrm{p}}$. Conversely, in the upper tail of $p(C)$ we might assume the MTE is at least the same as for the highest estimable MTE, and no greater than twice this.

The fourth problem we face is that over time individuals drop out of self-assessment taxation. To the extent that this represents a loss of tax revenue for HMRC, we assume zero returns for these individuals in all years after they stop submitting.[8] A problem arises if we have differential attrition, where the probability of survival differs between the treatment and control groups. Since the treatment and control groups are selected from slightly different populations – treatment from taxpayers asked to file returns, control from taxpayers who submit returns – the differences in observables between these groups might lead to differential attrition. Let $S(t, \tau) \in 0, 1$ denote the event that an individual 'survives' in the sample to year t given that they were first observed in year τ. If $\Pr(S|\mathbf{X}, t - \tau) \neq \Pr(S)$ then the probability of survival depends on either observed characteristics, time since entry into the sample, or both. If characteristics $\mathbf{X}$ and $t - \tau$ differ systematically between treated and control, then we get differential attrition between these groups. If these characteristics are also correlated with tax paid, then this will bias our estimates.

To solve the dynamic selection problem, there are two possibilities. Firstly, if we could construct a control group on the same basis as the treated group, then the characteristics at baseline would (in expectation) be balanced, so the

[8]Since we cannot link self-assessment tax records to other tax records, we cannot see whether these dropouts leave the tax system altogether, or merely switch to paying tax only through witholding of various forms of income. We therefore include in our results some robustness analysis, comparing how the treatment effect estimated would change if we instead assumed that these individuals continued to pay their last observed amount of tax in all years after they stop being observed (and similarly pay their first amount in all years before they are observed).

probability a treatment group taxpayer would drop out is on average the same as for a control group taxpayer, in the absence of audit. The control group is then a valid counterfactual, although care must again be taken to interpret the results since the population from which selection is done is not the same as the population that is most of interest (we are interested in those who submit returns, but we select from those asked to submit). A second approach would be to estimate the probability that an individual survives $t - \tau$ years after audit, $\widehat{\Pr}\left(S|\mathbf{X}, t-\tau, p(C)\right) \equiv \hat{q}(S)$ as a function of his characteristics, $\mathbf{X}$, and his probability of cheating, $p(C)$. We can then estimate the expected tax paid, allowing it to depend on audit status, probability of cheating, and additionally the probability of survival: $\mathbb{E}\left[Y_{t+h}|D(t)=1, \hat{p}\left(C\right)=p, \hat{q}(S)\right]$. Then the MTE for probability of audit p can be calculated as:

$$\beta_h^{MTE}(p) = \int_0^1 \left[\mathbb{E}\left[Y_{t+h}|1,p,q\right] - \mathbb{E}\left[Y_{t+h}|0,p,q\right]\right] f_{Q|P}(q|p)\,\mathrm{d}q$$

where $f_{Q|P}$ is the conditional density of survival probability given the probability of cheating.

Whilst the uniformly randomly selected control group allows us to identify, under some conditions, the ATE and the distribution of MTEs for each time horizon since audit, it does not provide a route to understanding how these treatment effects vary with whether an individual was found cheating. In the next subsection we discuss an alternative form of control group that can help us to answer this question.

5.3 The "future-audited" control group

For the second control group approach, we use the pre-audit tax returns of individuals who are randomly audited in the future. For example, to identify the effects of audits one year after the audit, control individuals consist of 2005 returns for taxpayers audited in 2006 and 2007 returns for individuals audited in 2008 (i.e. the reported tax liability one year prior to their random tax audit), all pooled together and taking out year effects. The advantage of using this "future-audited" control group approach is that we can use it to separately identify the treatment effect based on the outcome of the audit. For example, to identify the effect of random audits on individuals found to be non-compliant we use the pre-audit tax returns of individuals we know (ex post) will be found to be non-compliant in the future and compare them with randomly-audited individuals who have been found to be non-compliant.

We will be able to rely on this method of constructing a control group if the selection mechanism for audits does not change over time and if the audit process is stable over time. The first of these conditions is trivially met as we are considering a random audit programme. With regard to the second condition, from talking to HMRC representatives we have had no indications of formal shifts in the audit intensity over time, but we cannot determine to what extent changes in the audit process affect the results using this approach.

Even if these conditions are satisfied, there is a risk that we end up comparing individuals who are not comparable with each other. The reason is that (without careful conditioning) individuals need to survive as self-assessment taxpayers for different numbers of years to be in the treatment group compared to the control group. For example, for outcomes two years after the audit, treated individuals have to appear for three years (the audit year and two years later), while control individuals only have to appear one year before audit and in the audit year. This problem becomes worse the further into the future we consider. We can think of two different properties of the data that might cause problems for using the future-audited approach where there are differences in average number of years we require treatment and control groups to be in the sample.

First, there could be learning in the self-assessment population: i.e. individuals change their reported tax liability over time over and above general productivity improvements and inflation. This could be the case if, say, they become better at filling in their tax return (e.g. by gaining knowledge about what types of expenses are deductible) or if they become better at the activity that made them part for the self-assessment population (e.g. plumbers might become better at plumbing the longer they have been in business). If there is evidence of such learning effects in our data (we have not yet tested for this) then we will need to ensure that we compare individuals who have been observed equally long in the self-assessment population at the time of comparison.

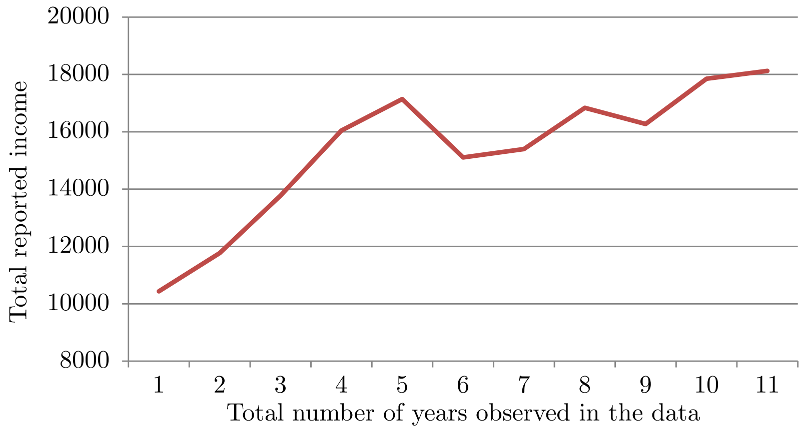
Second, it could be that individuals who survive longer in the data (ex post) are, on average, of better quality than individuals who survive shorter in the sample. Suppose that there are different "types" of individuals with different levels of productivity in our sample. Further suppose that there is a random process of temporary productivity shocks. In this case, the better the individual's "type", the more negative a productivity shock will have to be to put him out of business. That is, individuals who are in the sample longer will, on average, be of better quality. From looking at Figure 3 we see that there is evidence that the average reported tax liability increases with the total number of years an individual is observed (ex post) in the sample. That is, when we use the future-audited control approach we need to ensure that the treatment and control group and treatment group are (ex post) observed equally long in the sample.

6 Results

In this section we present results for the "could have" control group (constructed from individuals who could have been selected for random audit but who weren't). The outcome we focus on is reported tax liability from the self-assessment return.

We begin by showing the patterns in the raw data that drive our results. Figure 4 plots the reported tax liability in treatment and control groups in the run-up to the audit and after the audit. Before the audit (years -10 to -1), reported tax liability is the same in treatment and control groups at around

Figure 3: Average reported income by total number of years observed in the data

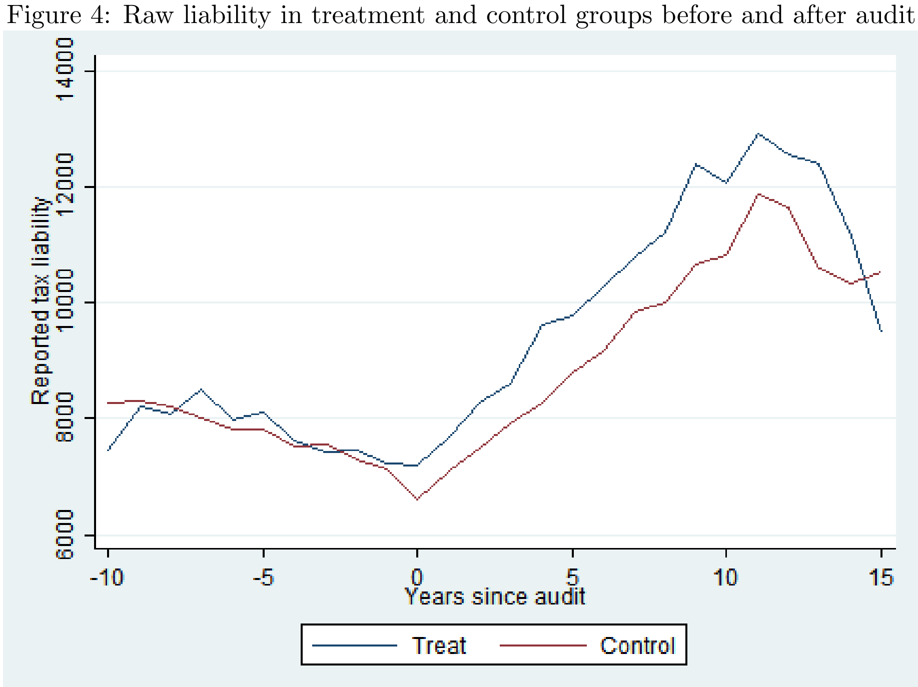


Source: calculations based on HMRC administrative datasets. Note: The average reported income has been calculated using only non-treated individuals, and year effects have been taken out.

£8,000. This is what we hope to see if treatment and control groups are balanced (i.e. the control group is a valid counterfactual for the treatment group). In years after the audit, the treatment and control lines diverge, with reported tax liability in the treatment group exceeding that in the control group by £500-1,000 for at least the first 10 years after the audit. The difference may disappear by around 15 years after the audit. This suggests that the audit may have a positive and fairly substantial impact on reported tax liability over a prolonged period.

One interesting pattern in Figure 4 is the distinct V-shape in reported tax liability for both the treatment and control groups. The explanation for this is that individuals observed in the data for longer have a higher reported tax liability on average. This is what Figure 3 shows. The reason this leads to a V-shape in reported tax liability is that individuals are, on average, observed in the data longer the further they are from year zero (the year of the audit) in Figure 4 (because, for example, for 10 years after the audit the individual must have been observed in the audit year and 10 years after that). The V-shape is not symmetric because of a trend in reported tax liability over time (a higher number of years since audit will tend to be composed of a later tax years).

Figure 4: Raw liability in treatment and control groups before and after audit



Source: calculations based on HMRC administrative datasets.

Table 5: Estimated differences between treatment and control groups in the years before audit

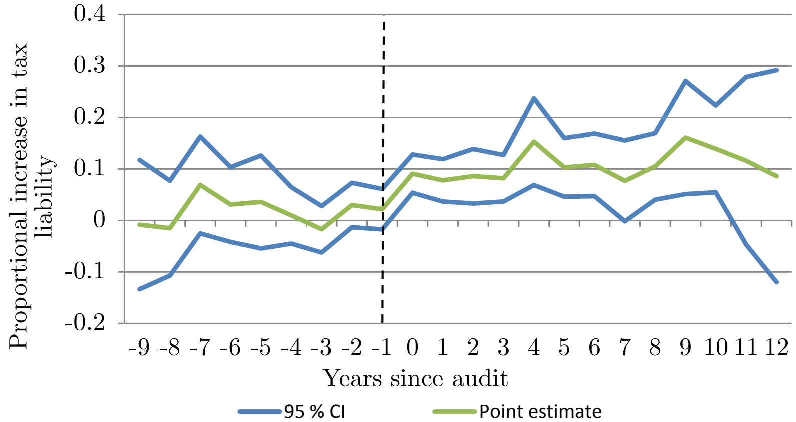
Years since audit	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
Proportional treatment effect	-0.117	-0.101	-0.008	-0.015	0.069	0.031	0.036	0.01	-0.017	0.03	0.022
Std. Error	(0.088)	(0.064)	(0.064)	(0.047)	(0.048)	(0.037)	(0.046)	(0.028)	(0.023)	(0.022)	(0.02)
Avg. log reported tax liability by control group	8.981***	8.994***	8.966***	8.945***	8.919***	8.901***	8.908***	8.867***	8.86***	8.812***	8.79***
Std. Error	(0.036)	(0.033)	(0.022)	(0.018)	(0.012)	(0.012)	(0.013)	(0.011)	(0.014)	(0.01)	(0.013)

Table 6: Estimated differences between treatment and control groups in the year of - and in years after - the audit

Years since audit	0	1	2	3	4	5	6	7	8	9	10
Proportional treatment effect	0.091***	0.078***	0.086**	0.082***	0.153***	0.103***	0.108**	0.077	0.105**	0.161**	0.139**
Std. Error	(0.019)	(0.021)	(0.027)	(0.023)	(0.043)	(0.029)	(0.031)	(0.04)	(0.033)	(0.056)	(0.043)
Avg. log tax liability reported by control group	8.717***	8.79***	8.85***	8.896***	8.939***	9.007***	9.048***	9.123***	9.13***	9.18***	9.194***
Std. Error	(0.011)	(0.011)	(0.014)	(0.014)	(0.016)	(0.021)	(0.02)	(0.025)	(0.023)	(0.032)	(0.03)

Notes: Based on a Poisson regression of reported tax liability on a an indicator for number of years since audit, number of years since audit interacted with an indicator for being in the treatment group, year fixed effects, and a constant as shown by equation (1). The proportional treatment effect is only an approximation. To get the precise proportional difference between the treatment and control group we apply the formula: $\exp(\beta) - 1$. The output tables does not show the coefficients of the estimated year dummies. However all of the estimates above should be interpreted as net of year effects. * $p < .05$, ** $p < .01$, *** $p < .001$.
Source: calculations based on HMRC administrative datasets.

Figure 5: Impact of audit on reported tax liability



Source: calculations based on HMRC administrative datasets.

We now investigate formally what the impact of being audited is on subsequent reported tax liability by running our main regression specification. Tables 5 and 6 and Figure 5 present the results from running a Poisson regression of reported tax liability on years since audit dummies, year since audit dummies interacted with a treatment dummy, year dummies and a constant. We use Poisson regression because reported tax liability is highly skewed and as we want to allow for zero's in the data. Furthermore, we are interested in identifying the proportional increase.[9] We therefore fit a model of the form:

$$y_{it} = \exp\left(\alpha + \sum_{\tau} \beta_\tau D_i T_{it}(\tau) + \sum_{\tau} \theta_\tau T_{it}(\tau) + \gamma_t + u_{it}\right) \tag{1}$$

where y_{it} is reported tax liability for individual i, at time t, , $T_{it}(\tau)$ is a dummy equal to one if time since audit (in years)=τ, D_i is the audit group dummy and γ_t is the vector of year dummies. Our parameter(s) of interest is β_τ, which is the additional reported tax liability of treated individuals τ years since the tax year to which the audit relates.

In order to be confident that treatment and control groups are balanced, we hope to see that there is no effect of the audit before the audit has happened. This is exactly what we observe: none of the treated dummies are significant in Table 5 and the confidence bands straddle zero before the audit in Figure 5.

After the audit, Table 6 shows that the effect of the audit is strongly statistically significant in almost all years, with the coefficient ranging between about 0.08 and 0.16. Given the regression specification, this corresponds to an effect of between 8 and 17 per cent ($e^{.161} - 1$). Recall that this is an ITT parameter

[9] We calculate robust standard errors to relax the usual Poisson assumption that the mean is equal to the variance.

that is the effect of being selected for random audit relative to a baseline of not being selected for random audit but facing the normal policy environment, which includes a small chance of a targeted audit. Thus it is likely to be a lower bound on the impact of being selected for audit relative to not being selected. There is no noticeable decline in the size of the impact as the horizon is extended, suggesting that the effect of audits on reported tax liability is strongly persistent. This is a sizable impact: it corresponds to between about £500 and £1,600 in cash terms—so roughly the same order of magnitude as the initial audit adjustment across all taxpayers (£700) (the product of the non-compliant share and mean additional tax if noncompliant in Table 4.2).

Note, however, that the 8-17 per cent impact of the audit that we find may be due to an extensive effect, an intensive effect or some combination of the two. By an extensive effect, we mean that the audit leads some taxpayers to drop out of self-assessment and therefore disappear from our data. Given the large positive impact of audits, it would have to be the lowest-earning/least-profitable self-assessment individuals who drop out. By an intensive effect, we mean that the audit leads some taxpayers to change the amount tax that they report as due. This could either be as a result of a change in reporting behaviour or as a result of a change in underlying real behaviour.

Which of these channels is responsible matters a lot. Indeed, if the extensive margin alone is responsible, then HMRC could lose tax revenue. If the intensive margin alone is responsible, then HMRC gains by around £700 each year per individual selected for a random audited. If some combination of extensive and intensive effects is responsible, then the figure will be between the two. In principle it would be possible for us to disentangle the extensive and intensive channels. We will explore this in future work.

7 Conclusion

This paper investigates how long-lasting is the effect of audits on reported tax liability. This is an important question from the perspective of quantifying the returns to an audit, as well as understanding the mechanisms by which audits might influence taxpayer behaviour.

To answer this question we exploit a random audit program run by the UK tax authority (HMRC), under which an average of around 2,800 individuals are selected for random audits each year. We use data on audits and individual income tax returns for the tax years 1998/99 to 2011/12.

Our results suggest that there is a large and persistent impact of audits on reported tax liability of between 8 and 17 per cent that persists for at least 10 years after the audit. This is roughly the same order of magnitude as the initial audit adjustment across all taxpayers. This is a sizable effect and emphasises the importance of taking the indirect revenue effects into account when deciding on the optimal enforcement strategy. Note, however, that this could be due to an extensive or an intensive effect (or a combination of the two).

Future work will seek to disentangle extensive and intensive treatment effects, investigate how responses vary depending on the audit outcome (e.g. do taxpayers found to be non-compliant increase their reported tax liabilities by more?) and compare impacts for across different income types subject to different third-party reporting requirements.

References

Agostini, C. A. and C. Martnez, "Response of Tax Credit Claims to Tax Enforcement: Evidence from a Quasi-Experiment in Chile," *Fiscal Studies*, 2014, *35* (1), 4165.

Ali, M. M., H. W. Cecil, and J. A. Knoblett, "The effects of tax rates and enforcement policies on taxpayer compliance: A study of self-employed taxpayers," *Atlantic Economic Journal*, June 2001, *29* (2).

Allingham, M. G. and A. Sandmo, "Income tax evasion: A theoretical analysis," *Journal of Public Economics*, 1972, *1*, 323–338.

Alm, J. and M. Mckee, "Tax compliance as a coordination game," *Journal of Economic Behavior & Organization*, 2004, *54*, 297312.

_ , B. R. Jackson, and M. Mckee, "Getting the word out: Enforcement information dissemination and compliance behaviour," *Journal of public economics*, 2009, *93*, 392402.

Andreoni, J., B Erard, and J Feinstein, "Tax compliance," *Journal of Economic Literature*, June 1998, *36* (2), 818–860.

Bergman, M. and A. Nevarez, "Do Audits Enhance Compliance? An Empirical Assessment of VAT Enforcement," *National Tax Journal*, 2006.

Bloomquist, K. M., "Incorporating Indirect Effects in Audit Case Selection: An Agent-Based Approach," *Forthcoming: IRS Research Bulletin*, 2013, *1500.*

Choo, L., M. A. Fonseca, and G. Myles, "Lab Experiment to Investigate Tax Compliance: Audit Strategies and Messaging," *HM Revenue and Customs Research*, 2013, *Report 308.*

Dubin, J. A., "Criminal Investigation Enforcement Activities and Taxpayer Noncompliance," *Public Finance Review*, 2007, *35* (3), 500–529.

_ , M. J. Greatz, and L. L. Wilde, "The effect of audit rates on the federal individual income tax, 1977-1986," *National Tax Journal*, December 1990, *43* (4), 395–409.

Erard, B, "The influence of Tax audits on reporting behaviour," *Why People Pay Taxes: Tax Compliance and Enforcement*, 1992.

Fortin, B., G. Lacroix, and M. Villeval, "Tax evasion and social interactions," *Journal of Public Economics*, 2007, *91*, 20892112.

Friedland, N., S. Maital, and A. Rutenberg, "A simulation study of income tax evasion," *Journal of Public Economics*, 1978, *10*, 107–116.

Gemmell, N and M Ratto, "Behavioral responses to taxpayer audits: Evidence from random taxpayer enquiries," *National Tax Journal*, 2012.

HMRC, "Measuring tax gaps: 2013 edition," 2013.

Kastlunger, B., E. Kirchler, L. Mittone, and J. Pitters, "Sequences of audits, tax compliance, and taxpaying strategies," *Journal of Economic Psychology*, 2009, *30*, 405418.

Kirchler, E., B. Macijovsky, and H Schwarzenberger, "Misperception of chance and loss repair: On the dynamics of tax compliance," *Journal of Economic Psychology*, 2007, *28*, 678691.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez, "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica*, 2011.

Long, S and R. Schwartz, "The Impact of IRS Audits on Taxpayer Compliance: A Field Experiment in Specific Deterrence," *Annual Law and Society Association Meeting 1987*, 1987.

Mittone, L., "Dynamic behaviour in tax evasion: An experimental approach," *The Journal of Socio-Economics*, October 2006, *35* (5), 813835.

Plumley, A. H., "The Determinants of Individual Income Tax Compliance Estimating the Impacts of Tax Policy, Enforcement, and IRS Responsiveness," *Internal Revenue Service Publication*, 1996, *1916*, Rev. 11–96.

Slemrod, J., "Cheating Ourselves: The Economics of Tax Evasion," *Journal of Economic Perspectives*, Winter 2007, *21* (1), 25–48.

_ and S. Yitzhaki, "Tax avoidance, evasion and administration," *Handbook of Public Economics*, 2002, *3.*

_ , M. Blumenthal, and C. Christian, "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota," *Journal of Public Economics*, 2001.

Tauchen, H. V., A. D. Witte, and K. J. Beron, "Tax compliance: An investigation using individual taxpayer compliance measurement program (TCMP) data," *Journal of Quantitative Criminology*, 1993, *9* (2).