



May 2019

No.414

**The Dynamic Effects of Tax Audits**

Arun Advani, William Elming and Jonathan Shaw

**WORKING PAPER SERIES**

Centre for Competitive Advantage in the Global Economy

Department of Economics

# The Dynamic Effects of Tax Audits

Arun Advani\*

William Elming<sup>†</sup>

Jonathan Shaw<sup>‡</sup>

May 20, 2019

## Abstract

Understanding causes of and solutions to non-compliance is important for a tax authority. In this paper we study how and why audits affect reported tax in the years after audit – the dynamic effect – for individual income taxpayers. We exploit data from a random audit program covering more than 53,000 income tax self assessment returns in the UK, combined with data on the population of tax filers between 1999 and 2012. We first document that there is substantial non-compliance in this population. One in three filers underreports the tax owed. Third party information on an income source does not predict whether a taxpayer is non-compliant on that income source, though it does predict the extent of underreporting. Using the random nature of the audits, we provide evidence of dynamic effects. Audits raise reported tax liabilities for at least five years after audit, implying an additional yield 1.5 times the direct revenue raised from the audit. The magnitude of the impact falls over time, and this decline is faster for less autocorrelated income sources. Taking an event study approach, we further show that the change in reporting behaviour comes only from those found to have made errors in their tax report. Finally, using an extension of the Allingham-Sandmo (1972) model, we show that these results are best explained by audits providing the tax authority with information, which then constrains taxpayers’ ability to misreport.

JEL codes: D04, H26, H83

Keywords: tax audits, tax revenue, tax reporting decisions, income tax, self assessment, HMRC

---

\*University of Warwick, the Institute for Fiscal Studies (IFS), the Tax Administration Research Center (TARC), and the Centre for Competitive Advantage in the Global Economy (CAGE). The authors thank Michael Best, Richard Blundell, Tracey Bowler, Monica Costa Dias, Dave Donaldson, Clive Fraser, Costas Meghir, Áureo de Paula, Imran Rasul, Yee Wan Yau, and seminar participants at the Tax Systems Conference, Royal Economic Society, Louis-André Gérard-Varet, European Economic Association, Warwick Applied, OFS Empirical Analysis of Tax Compliance, International Institute of Public Finance, Econometric Society European Meetings, and National Tax Administration Conferences, for helpful comments. 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. Correspondence: Dept of Economics, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL. Email: economics@arunadvani.com.

<sup>†</sup>IFS and TARC at the time of involvement in this work.

<sup>‡</sup>Financial Conduct Authority.

# 1 Introduction

Central to the efficient functioning of a modern tax authority is the ability to assess and collect tax revenue owed by taxpayers in a timely and cost-effective manner. While the traditional economic analysis of taxation has focused on discussion of optimal *tax rates* across various *tax bases*, an important third arm of policy is *tax administration*: the assessment and collection of taxes owed. In the UK, around 6.5 per cent of tax liabilities go uncollected (HM Revenue and Customs, 2016); in cash terms this is equivalent to the value of corporation tax receipts. In the US this number is 16.3 per cent (Internal Revenue Service, 2016). Taxpayer audits are one important tool used by many tax authorities to improve compliance. Audits have a mechanical benefit in terms of the unpaid revenue they identify and recover. Historically, this is primarily what tax authorities such as the Internal Revenue Service (IRS) have focused on when selecting tax returns for further scrutiny (see Bloomquist, 2013). However, since most taxpayers pay taxes over many years, it is important to understand the extent of any ‘dynamic effects’ that may occur. Dynamic effects are changes in the behaviour of taxpayers in the years after they have been audited. Measuring these wider impacts of audits, and understanding why they occur, is crucial for determining audit strategy including the optimal extent and targeting of enforcement.

In this paper we study *how* and *why* audits affect reporting behaviour in subsequent years. Using administrative data from the UK, we show that audits have important dynamic effects: the aggregate additional tax paid in the five years after an audit is 1.5 times the direct revenue raised from an audit. These effects are driven by individuals who were found to be underreporting. The reason for this change in reporting is that audits provide the tax authority with information about a taxpayer’s income at the time of audit, making later misreporting more difficult.

To estimate the dynamic effect we exploit a random audit programme run by the UK tax authority (HM Revenue and Customs, HMRC). Over 53,000 individual tax filers were (unconditionally) randomly selected for audit by the programme between 1998/99 and 2008/09. This random selection allows us to address the common concern that audits are typically targeted towards taxpayers believed to be non-compliant i.e. those who are underreporting their tax liability. Using data on the population of UK self assessment taxpayers from 1998/99 to 2011/12, in our first identification strategy we construct a control group for each year of the programme from individuals who could have been selected for a random audit that year but were not. We then study the difference in reporting behaviour over time.

To distinguish the impact by audit outcome we use an alternative identification strategy. We conduct an event study, focusing on sets of individuals who were audited at some point in our sample and all who had the same audit outcome. This approach allows us to control for individual and calendar time fixed effects. Identification of the audit impact comes from comparing individuals audited and found to have a particular outcome with individuals who we know will be audited and found to have the same outcome.

We demonstrate five main empirical results. First, we show that there is significant non-compliance among individual self assessment taxpayers, both in the share of taxpayers who are found non-compliant and the share of tax that is misreported. More than one third of self assessment taxpayers are found to be non-compliant, equal to 12 per cent of all income taxpayers. Second, we demonstrate that third party reporting is important in influencing levels of compliance, but only on the *intensive* margin of how much tax is misreported. Third party information does not predict whether a taxpayer is non-compliant, but conditional on non-compliance it is associated with lower magnitudes of misreporting. Third, we provide evidence of important *dynamic effects*, with the additional tax revenue over the five years post-audit equalling 1.5 times the direct revenue raised by audit. Fourth, we show that for income sources which are relatively stable – have a high autocorrelation in the absence of audit – the impact is very long-lived, while for more volatile sources it goes to zero quickly. Finally, we show that being audited only changes the behaviour of those who are found to have misreported, and this is true whether or not they received a penalty.

To understand the mechanism driving the additional reported tax liabilities we extend the canonical model of tax evasion (Allingham and Sandmo, 1972; Kleven et al., 2011) to incorporate (simple) dynamics in the response to audit. This allows us to study the distinct predictions of different mechanisms that might drive changes in reporting: (1) changes in the perceived penalty for evasion; (2) changes in the perceived probability that evasion is detected; and (3) updates to the information held by the tax authority. Under the first mechanism, any impact should be permanent; while under the second and third it should decline over time. The second and third mechanisms are distinguished by their predictions across income sources: under the second any decline should be common to all income sources, whilst under the third less autocorrelated income sources should see faster declines. Also, under the third mechanism only those who are found to have errors should respond i.e. those for whom the tax authority learns new information.

Intuitively, an audit provides a snapshot measure of the current true level of income. For income sources which are relatively stable, this will also be a good predictor of revenue in the future, making it easier for the tax authority to detect later misreporting. For those which are less autocorrelated,

the predictive power of this snapshot will decline quickly over time, so it quickly becomes less useful in trying to detect misreports.<sup>1</sup> Both the dynamics by income source and the results by audit outcome point towards this third mechanism as explaining *why* we see these impacts by audit outcome.

This paper contributes to a small but growing literature on the dynamic effect of audits. Kleven et al. (2011) analyse the effect of a tax enforcement field experiment in Denmark. They show that audits increase self-reported, but not third party reported, income in the year after audit. We find similar effects of third party information on the amount of underreporting conditional on non-compliance, but in our context the choice of whether to be non-compliant at all appears to have additional determinants. Our analysis also extends up to eight years after the audits take place, allowing us to see that the increased reporting immediately after audit is only a transitory, albeit long-lived, effect.

Gemmell and Ratto (2012) investigate behavioural responses to taxpayer audits in the UK in the year 2000 using a limited version of the random audit data we use. They find no impact on overall tax declared.<sup>2</sup> They also 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. We tackle this same question, but using an event study strategy which can overcome the potential endogeneity present in their study from comparing only compliant or non-compliant individuals to a random control group.

In work conducted concurrently with this study, DeBacker et al. (2018) investigate the impact of random audits conducted by the US Internal Revenue Service (IRS) between 2006 and 2009 on subsequent taxpaying behaviour. One limitation of their study is that taxpayers selected for random audit in the US are explicitly told that they are receiving a random audit, whilst in the UK taxpayers are simply told they are being audited, using the same language as in usual ‘operational’ audits. Despite this, they find remarkably similar quantitative results to us, with the cumulative impact in the five years subsequent to audit also around 1.5 times the direct effect.<sup>3</sup> In addition to providing complementary results to DeBacker et al. (2018), we are able to split results by audit outcome. We also develop a simple framework that allows us to capture different possible mechanisms that might

---

<sup>1</sup>In fact an audit may not uncover all underreporting. Instead these predictions will then hold relative to the revealed level of income, which may be lower than the true income.

<sup>2</sup>The pattern of results they find is consistent with our results, but their limited sample size means that have less power. Similar results are also found in the US by Beer and Erard (2015).

<sup>3</sup>Earlier studies for the US found some limited evidence for such dynamic effects, but only studied one or two years after audit (Long and Schwartz, 1987; Erard, 1992; Tauchen et al., 1993).

explain the observed dynamics, as well as to pin down which mechanism can be supported by the data.

Finally, Maciejovsky et al. (2007) and Kastlunger et al. (2009) test compliance behaviour using laboratory experiments on students. They find compliance decreases immediately after an audit, suggestive of a ‘bomb-crater effect’.<sup>4</sup> Choo et al. (2013) do lab experiments with taxpayers rather than students, and do not find any evidence of a bomb-crater effect. We find that for taxpayers in a ‘real-world’ setting, a bomb-crater effect does not exist. This remains true even when looking by audit outcome.

The remainder of the paper is organised as follows. Section 2 outlines the policy context and data sources. Section 3 provides evidence on who is non-compliant, including the relationship between non-compliance and third party reporting. Section 4 shows how audits affect reporting behaviour in overall tax, and by different income sources. Section 5 uses an alternative identification strategy to estimate the impact by audit outcome. Section 6 outlines a model of tax evasion with dynamics in the response to audits, to show which mechanisms might rationalised the observed behaviour. Section 7 concludes.

## 2 Context and Data

### 2.1 The UK self assessment tax collection and enforcement system

In this paper, we focus on individuals who file an income tax self assessment return in the UK. Over our sample period (1999-2012) this comprised around 9 million individuals, over one-third of all individual taxpayers in the UK. Income tax is the largest of all UK taxes, consistently contributing a quarter of total government receipts over this period. Most sources of income are subject to income tax, including earnings, retirement pensions, income from property, interest on deposits in bank accounts, dividends, and some welfare 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. Table 1 shows the share of individuals in our sample reporting non-zero values

---

<sup>4</sup>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 World War I soldiers to hide out in bomb craters, believing that it was unlikely that a bomb would strike exactly the 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” (see Andreoni et al., 1998, p. 844).

for each component of income. When we later study income components separately, we focus on those components where at least five per cent of the population report non-zero values.

Unlike the US, not all taxpayers have to file a tax return in the UK. The one-third of taxpayers required to submit a tax return 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 incomes over £100,000, 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, HM Revenue and Customs (HMRC, the UK tax authority) carries out audits each year to deter non-compliance and recover lost revenue. HMRC runs two types of audit, ‘targeted’ (also called ‘operational’) and ‘random’. Targeted audits are based on perceived risks of non-compliance. Random audits are unconditionally random from the population, and are used to ensure that all self assessment taxpayers face a positive probability of being audited, as well as to collect statistical information about the scale of non-compliance and predictors of non-compliance that can be used to implement targeting.

The timeline for the audit process is as follows. The tax year runs from 6th April to 5th April. Shortly after the end of the tax year, HMRC issues a ‘notice to file’ to taxpayers who they believe need to submit a tax return. This is based on information that HMRC held shortly before the end of the tax year. Random audit cases are provisionally selected from the population of individuals issued with a notice to file. The deadline by which taxpayers must submit their tax return is 31 January the following calendar year (e.g. 31 January 2008 for the 2006/07 tax year). Once returns have been submitted, HMRC deselects some random audit cases (e.g. due to severe illness or death of the taxpayer). At the same time, targeted audits are selected on the basis of the information provided in self assessment returns and other intelligence. Random audits are selected before targeted audits, and individuals cannot be selected for a targeted audit in the same tax year as a random audit. The list of taxpayers to be audited is passed on to local compliance teams who carry out the audits. Up to and including 2006/07, audits had to be opened within a year of the 31 January filing deadline, or a year from the actual date of filing for returns filed late. For tax returns relating to 2007/08 or later, audits had to be opened within a year of the date when the return was filed. Taxpayers subject to an audit are informed when it is opened but they are not told whether it is a random or targeted audit. Approximately one third of taxpayers on the list passed on to local compliance teams end up not being audited, largely due to resource constraints.

Those who are audited initially receive a letter requesting information to verify what they have reported. If this does not provide all the required information, the taxpayer receives a follow up phonecall, and ultimately in-person visits until the auditor is satisfied.

## 2.2 Data Sources

We exploit data on income tax 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 number and tax year.

Audit records for tax years 1998/99 to 2008/09 come from CQI (Compliance Quality Initiative), an operational database that records audits of income tax self assessment returns. It includes operational information about the audits, such as start and end dates, and audit outcomes: whether non-compliance was found, and the size of any correction, penalties and interest.

We track individuals before and after the audit using information from tax returns for the years 1998/99 to 2011/12. This comes from two datasets: SA302 and Valid View. The SA302 dataset contains information that is sent out to taxpayers summarising their income and tax liability (the SA302 tax calculation form). It is derived from self assessment returns, which have been put through a tax calculation process. It contains information about total income and tax liability as well as a breakdown into different income sources: employment earnings, self employment profits, pensions, and so on. For all of these variables, we uprate to 2012 using the Consumer Prices Index (CPI) to account for inflation, and trim the top 1 per cent to avoid outliers having an undue impact on the results. We supplement these variables with information from Valid View, which provides demographics and filing information (e.g. filing date). Note that we cannot identify actual compliance behaviour subsequent to the audit: the number of random audit taxpayers that are re-audited is far too small for it to be possible to focus just on them.

An explicit control group of ‘held out’ individuals was not constructed at the time of selection for audit. We therefore draw control individuals from the pool of individuals who actually filed a tax return (i.e. those who appear in SA302). This creates some differences in the filing history between those selected for audit and those who we deem as controls. In a given year, first time filers may be issued a notice to file *after* selection for audit has taken place. They may also end up back-filing one or two returns. Since we cannot directly observe the first year in which a notice to file was issued, in our empirical strategy it is necessary for us to control for the length of time each taxpayer has been



in self assessment. More details – including tests to demonstrate this ensures samples are balanced – are given in Subsection 4.1 below.

### 3 Tax Evasion in the UK

In this section we establish two main results. First, we show that there is significant non-compliance among individual self assessment taxpayers, both in the share of taxpayers who are found non-compliant and the share of tax that is misreported. More than one third of self assessment taxpayers are found to be non-compliant, equal to 12 per cent of all income taxpayers. Second, we demonstrate that third party reporting is important in influencing levels of compliance, but only on the *intensive* margin of how much tax is misreported. Third party information does not predict whether a taxpayer is non-compliant, but conditional on non-compliance it is associated with lower magnitudes of misreporting. Before describing these results in detail, we first provide some descriptives on the probability and timeline of audits.

#### 3.1 Audit descriptives

Figure 1 shows the number of individuals per year who face an income tax random audit over the period 1998/99 to 2008/09.<sup>5</sup> On average over the period the probabilities of being audited are 0.04 per cent (four in 10,000) for random audits and 2.8 per cent for targeted audits.

Table A1 provides some summary statistics for lags in, and durations of, the audit process among random audit cases. As described above, up to and including the 2006/07 return, HMRC had to begin an audit within 12 months of the 31 January filing deadline; since then, HMRC has had to begin an audit within 12 months of the filing date. The average lag between when the tax return was filed and when the random audit was started is 8.9 months, but 10 per cent have a lag of 14 months or more. The average duration of audits is 5.3 months, but 10 per cent experience a duration of 13 months or more. Taken together, this means that the average time between when a return is filed and when the audit is concluded is 14.3 months but there is a long tail for whom the experience is much more drawn out: for almost 10 per cent it is two years or more. This means that individuals will generally have filed at least one subsequent tax return before the outcome of the audit is clear and some will have filed two tax returns. This will be relevant for interpreting the results in Section 4.

---

<sup>5</sup>There are also a small number of partnerships and trusts that are audited, but we exclude these from our analysis.

## 3.2 Evidence of non-compliance

We begin by studying the direct results of random audits, using data on 34,630 completed random audits of on individual self assessment taxpayers from 1998/99 to 2008/09.<sup>6</sup> Table 2 summarises the outcomes of these random audits. More than half of all returns are found to be correct, and 11 per cent are found to be incorrect but with no underpayment of tax, but 36 per cent are ‘non-compliant’ i.e. incorrect and have a tax underpayment.<sup>7</sup> Whilst this is a much higher rate of non-compliance than has been found in other developed country contexts, it should be noted that the self assessment tax population is a selected subset of all taxpayers. In particular, it covers those for whom simple withholding of income at source is not sufficient to collect the correct tax. This may be either because some income cannot be withheld e.g. property or self-employed income, or because PAYE struggles to assign the correct withholding codes as for people with multiple sources of pension income. Despite this, since self assessment taxpayers make up a third of all UK taxpayers, this implies an overall non-compliance rate of 8 to 12 per cent among all taxpayers.<sup>8</sup>

Turning to the intensive margin, the average additional tax owed among the non-compliant is £2,314, or 32 per cent of average liabilities. Since just over a third of random audits find evidence of non-compliance, the average return from an audit is then £826. However, the distribution is heavily skewed: 60 per cent of non-compliant individuals owe additional tax of £1,000 or less, whilst four per cent owe more than £10,000. In terms of total revenue, those owing £1,000 or less make up only 9 per cent of the underreported revenue; the four per cent owing more than £10,000 collectively owe more than 42 per cent of the revenue. Equity concerns around non-compliance are well-known: it is seen as unfair that some are not ‘paying their fair share’. But this is variation in non-compliance is also important for efficiency. Noncompliant individuals previously acted as though there was a lower tax rate. This makes their activities seem more productive than they were, so can lead to resource misallocation.

---

<sup>6</sup>53,400 cases were selected for audit over the period, of which 35,630 were implemented.

<sup>7</sup>Incorrect with no underpayment includes those who, for example, owed no taxes since they had legitimate losses, but who had overstated those losses so would owe less in future years. Anecdotally, it also includes some cases where actual overpayments of tax were made, although we cannot separately identify which.

<sup>8</sup>This is a lower bound, since it assumes everyone who should be in self assessment does register, all non-compliance is picked up at audit, and those who do not need to register are also fully compliant. The range from 8 to 12 per cent depends on the assumptions made about the implementation of audits. If among those selected for audit, implementation of audit were random, this would imply a 12 per cent non-compliance rate. On the other hand, if there is perfect compliance among those for whom audits were not implemented, this would imply an 8 per cent rate.

### 3.3 Third party reporting

Understanding the role of third party reporting has been an important advance in the recent literature: this makes it much easier for a tax authority to detect misreporting by a taxpayer, since tax returns can be automatically checked against third party reports. Existing evidence suggests that third party reporting is important in explaining the relatively high levels of observed compliance (Kleven et al., 2011).

In our context such effects are also visible, but they are seen only on the intensive margin. Figure 2 orders income sources from no third party information on the left to fully third party reported on the right.<sup>9</sup> The left panel of Figure 2 shows that third party reporting is not a good predictor of *whether* a taxpayer is non-compliant (extensive margin). However, as seen in the right panel, conditional on being non-compliant, misreporting is clearly greater when no third party income is available.

To compare these results to Kleven et al. (2011), they find 14.9 per cent of self-employment income is underreported but only 2.3 per cent of dividend income and 1.1 per cent of employment income, not conditioning on compliance status. The comparable numbers from our context are 26.1 per cent, 5.2 per cent, and 3.5 per cent. Underreporting therefore seems to be a greater problem in the UK context than in Denmark, whether third party information is available or not. Similar results apply when looking at non-compliance status. In Denmark the share of taxpayers who are found to be non-compliant is 44.9, 2.2, and 2.6 per cent among those with self-employment, dividend, and employment income respectively, compared with 58.7, 10.5, and 28.5 per cent in the UK.

The extensive margin results imply that there are some other important differences between individuals receiving different types of income, potentially suggesting some effect of ‘tax morale’ – the intrinsic disutility from misreporting (Luttmer and Singhal, 2014). Although our later results across income sources will focus on the the dynamics of any impact compared to its peak, so are not directly affected by this, this pattern will motivate our study of within person effects to ensure that the results are not driven by cross-sectional differences between taxpayers.<sup>10</sup>

---

<sup>9</sup>See Table A2 for more details on the level of third party reporting by income source in the UK.

<sup>10</sup>Appendix B provides more information on how non-compliance relates to individual characteristics.

## 4 Dynamic Impacts of Audits

In this section we establish two main results. First, we show that audits lead to an increase in reported incomes and taxes in subsequent years. Looking at total income and total tax this increase lasts five to eight years after the tax year for which the audit was done. Second, we show variation in this impact by income source. In particular, more autocorrelated income sources (such as pensions) seem to respond permanently to audit. In contrast, income sources which are less autocorrelated, such as self-employment income, more quickly return to baseline. This second result will later help explain *why* we see these dynamic responses. Before describing these results in detail, we first discuss the empirical approach taken. Briefly, we compare individuals selected for random audit with those not selected but who could have been selected. We control for filing history, to account for the way the sample was selected.

### 4.1 Estimation

To understand how audits affect future tax receipts we want to estimate the change in tax paid in the years subsequent to an audit that is caused by the audit. To recover this we make use of the ‘random audits programme’ run by the tax authority (HMRC). This programme selects for audit a random sample of taxpayers from the pool of taxpayers known to be required to file for a given tax year. One can therefore compare those selected for audit with others who were not selected but who could have been.

In each audited tax year we select a sample of individuals who were not audited and could have been. We assign them a “placebo audit” for that tax year. We can then compare them over time to individuals actually selected for audit for that year. Our sample therefore consists of individuals who were selected for random audit in some year between 1999 and 2009, and individuals who could have been selected in those same years but were not.<sup>11</sup> For every individual selected for audit in a given tax year, we draw six control individuals from the population of those who could have been audited in the same tax year.<sup>12</sup>

In practice a little more than two-thirds of those selected for random audit are actually audited. This is largely explained by the high workload faced by the compliance teams implementing audits.<sup>13</sup>

---

<sup>11</sup>Our data on tax returns goes up to 2012.

<sup>12</sup>In principle the entire population of taxpayers who could have been audited could have been used. However, since the data could be accessed only in a secure facility at the tax office, computational constraints given the available hardware limited the sample size that could be used.

<sup>13</sup>A very small share are also ‘deselected’ for audit after being assigned to the random audit group, for example because the taxpayer has died or because there is an ongoing targeted audit.

Additionally, a small fraction of the control group (around two per cent) is also audited. Random audits are selected before targeted audits, and no explicit control group was constructed to ‘hold out’ some individuals from targeting.<sup>14</sup> This will tend to reduce the estimated impacts, since individuals in the control group who are most likely to be non-compliant are audited.

In the empirical work to follow, we focus on the local average treatment effect (LATE), instrumenting receipt of audit with selection for random audit.<sup>15</sup> This is the relevant number for a tax authority thinking about simultaneously expanding the size of the random audit programme and the number of auditors. It gives the average impact  $h$  years after audit for an additional random audit case that might be worked, against which the cost of the audit would be compared.

One limitation of our data is a slight mismatch between our treated and control samples in terms of their probability of filing in previous years, for reasons relating to the audit timeline and when they were first issued a notice to file, as described in Subsection 2.2. This can be seen in Table A3 which documents (unconditional) sample balance between five and one years before audit, for income and tax totals, income components, and individual characteristics. Overall balancing statistics suggest that the samples are fairly well balanced: the p-value of the likelihood-ratio test of the joint insignificance of all the regressors is 0.181, while the mean and median absolute standardised percentage bias across all outcomes of interest are low at 2.4 per cent and 1.7 per cent respectively.<sup>16</sup> However, the likelihood of being in the sample in previous years (‘survival’) differs between our treatment and control groups. This difference is consistent with how the treatment and control groups were selected, so might reflect real differences in the samples. We therefore include controls for presence in the data in the years before audit. Table 3 shows that once we condition on past survival the sample is balanced.

We therefore estimate the following specification:

$$Y_{iht} = \sum_{h=-5}^8 \alpha_h \eta_h + \sum_{h=-5}^8 \beta_h \eta_h D_i + \sum_{s=1999}^{2012} \gamma_s T_s + \sum_{s=-4}^{-1} \delta_s S_{is} + \varepsilon_{iht} \quad (1)$$

---

<sup>14</sup>To our knowledge, in prior work only (Kleven et al., 2011) have an explicit control group. This explains why they can only study a single year after audit – tax authorities are unwilling to hold off on high value audits for multiple years. Hence we compare those selected for a random audit to a ‘business as usual’ group, rather than a pure control group.

<sup>15</sup>In practice this scales up the ‘intention to treat parameter’ – the mean difference between those selected for audit and those who could have been but were not – by 1.5, since the difference in share audited between the groups is two-thirds.

<sup>16</sup>The standardised percentage bias is the difference in the sample means between treated and control groups as a percentage of the square root of the average of the sample variances in the treated and control groups (see Rosenbaum and Rubin, 1985). Rubin’s B and R statistics are also well within reasonable thresholds to consider the samples to be balanced, at 10.8 and 0.983 respectively. Rubin’s B is the absolute standardized difference of the means of the linear index of the propensity score in the treated and control group. Rubin’s R is the ratio of treated to control variances of the propensity score index. Rubin (2001) recommends that B be less than 25 and that R be between 0.5 and 2 for the samples to be considered sufficiently balanced.

where  $Y_{iht}$  is the outcome for individual  $i$ ,  $h$  years after audit (with control observations having  $h = 0$  for the tax year for which they were drawn as controls), when current calendar year is  $s \equiv t + h$ .  $\eta_h$  are indicators for being  $h$  years after audit;  $D_i$  is an indicator for whether the individual is actually audited;  $T_s$  is a calendar time indicator for tax year  $s$ ; and  $\{S_{i,-1}, \dots, S_{i,-4}\}$  are indicators for whether the individual was in the data in each of the four years before audit. The error term,  $\varepsilon_{iht}$ , is clustered at the individual level. Audit status,  $D_i$ , is instrumented by (random) selection for audit,  $Z_i$ . The coefficients of interest are  $\beta_h \forall h$ . These estimate the impact of the audit on the outcome variable  $h$  years after audit, measured as the difference in the mean outcome for those actually audited and those who would have been audited only if selected for a random audit.

## 4.2 Overall impact of audits

Beyond the direct effects of the audit, described in Section 2, we also see clear evidence of *dynamic* effects. Comparing individuals who were randomly selected for audit with individuals who could have been (but were not) selected, those selected for audit on average report higher levels of tax owed in the years after audit. Figure 3 shows the estimated impact on those who were actually audited (i.e. the LATE).<sup>17</sup>

The impact of an audit peaks two years after the tax year for which the audit is conducted. This is consistent with the fact that many audits are not started until after the following year’s tax return has already been submitted.<sup>18</sup> Reported tax among audited taxpayers is significantly greater than among non-audited taxpayers for five years after the audit, and the point estimate appears to decline relatively smoothly, getting close to zero by the eighth tax year after the audited year.

From Figure 3 we can estimate how much revenue audits raise on average by changing the behaviour of audited individuals. Over the five (eight) years after the audited year, the dynamic effects bring in an additional £1,230 (£1,530).<sup>19</sup> This is 1.5 (1.8) times the direct effect of audit, similar in magnitude to the findings of DeBacker et al. (2018). This highlights the policy importance

<sup>17</sup>The difference in the share audited between the treated and control group is around 66 percentage points, so the LATE is around 1.5 times the intention to treat estimate.

<sup>18</sup>In our sample, almost a quarter of audits are not opened for more than 12 months from the date of filing (see Table A1). Additionally, there can be some lag between the tax authority ‘taking up’ a case for audit and notification being received by the taxpayer. If taxpayers each consistently file at the same time every year, this implies at least one quarter would have filed without knowledge of the audit. More than half will have filed without knowing the result of the audit (Table A1).

<sup>19</sup>It is not clear whether this should overestimate or underestimate individual-level non-compliance. Since individuals in the control group may still be selected for a targeted audit, the impact from many of the highest-yielding cases will not be included, suggesting an underestimate. On the other hand, Table A4 shows that audits are opened more quickly for those within the random audit group who were found to be more non-compliant, so it is possible that the remaining unworked cases had even lower misreporting. Then LATE would overestimate the average treatment effect. Note, however, that auditors may have had less time to work on cases opened later. This might have lowered the return on later audits, without implying that these individuals were more compliant. We do not attempt to distinguish these possibilities, since this issue is not our focus.

of studying such dynamic effects: when determining the optimal audit strategy, the revenue raising effects of audits would be grossly understated without considering the impact on future behaviour, which would typically imply too few audits taking place.

Figure 4 shows that a very similar pattern holds for the impact on *total income* reported. Again there is a clear dynamic effect, peaking two years after the audited year and declining to zero by year eight, though not significantly different from zero by year five. This provides additional support to the previous result for tax.<sup>20</sup>

### 4.3 Impact by income source

We repeat the previous estimation separately for different income sources.<sup>21</sup> This will be one way in which we discriminate between different possible explanations for why we see dynamic effects.

Figure 5 shows how the impact of an audit changes over time for the different components of income. Since the magnitudes of these incomes are different, for comparability we rescale them relative to the peak impact for that income source.

We see that, relative to the peak, self-employment income and dividends decline relatively quickly. Three years later point estimates for these are close to zero i.e. reporting is not different to the control group. In contrast, pension income exhibits little decline. Six years later it retains 80% of the impact, and this not statistically different from 100%. This pattern is suggestive of the importance of autocorrelation: income sources which one would expect to be more correlated over time appear to show weaker declines.

Table 4 shows the autocorrelation for each income source. Pension income is highly autocorrelated, since it will typically be an annuity and therefore fixed over time; property income is slightly more volatile, since rents may vary more; and at the other extreme self-employment and dividend income are considerably more volatile. The relative autocorrelations of income sources lines up exactly with their speeds of decline.

There are two caveats to these results. The first is that these measures are noisy, so if confidence intervals were added Figure 5 for each income source, many would overlap. The second is that, as seen in the left hand panel of Figure 2, individuals with different income sources have different propensities for non-compliance. To tackle both concerns together, we compare within individuals who have multiple income sources. This immediately solves the second problem, since our results

---

<sup>20</sup>This result is not purely by construction, since they are based on distinct underlying reported variables.

<sup>21</sup>We do this only for income sources for which at least five per cent of the sample report non-zero amounts, for reasons of sample size. We also exclude interest income, since it is both very small and not everyone needs to report this, making it hard to compare. See Table 1 for information on the share of individuals with each income source.

will be within individual. It will also lead to ten pairwise comparisons.<sup>22</sup> For each pair, our sample is composed of individuals who had both sources sometime in the three years before audit. We then study the relative fall in reporting of each of these income sources four years after the peak. In each case we expect to find the less autocorrelated source falls fastest.

We find this result in eight out of ten cases. If there were no relationship, we should find this to be true in around five of the tests. The probability of this result under the null of no relationship is 5.5%, close to standard significance thresholds. Hence more autocorrelated income sources do seem to decline slower than less autocorrelated ones.

Our interpretation for this result, which we formalise below, is that audits provide the tax authority with information. Once discovered, income from highly autocorrelated sources will be hard to hide again, as deviations from the truth will be easily uncovered. In contrast, declines in less autocorrelated income sources are less informative to the authority because they might well be real for an individual taxpayer.<sup>23</sup> Hence the dynamics seems to come from the provision of information. This is something we know from other settings to be important (Kleven et al., 2011; Pomeranz, 2015), although the value of audits as a potential source of information about future tax has not previously been recognised.

One caveat to this interpretation is that falls in reporting could alternatively be driven by changes in actual income. For example, those who are audited might sell shares to pay fines, reducing dividend income. Whilst this is possible, it seems unlikely. In cash terms, the peak additional income reported for those who have dividend income is £414. Assuming a high end estimate for the dividend yield of 10%, implies £4140 of undeclared shares. Conservatively assuming also that individuals are on the higher rate of income tax, this implies an additional £135 of tax owed. The absolute maximum penalty for misreporting is 100% of the tax due (on top of paying the tax). So selling all these shares (and hence looking like the control group) would be needed only for an individual who is found to have misreported for at least fifteen years, and receives the maximum fine. While such cases might exist, it seems extreme to assume that this is occurring at the average. Hence we think it is unlikely that the observed pattern represents changes in real behaviour, rather than reporting, though we cannot definitively rule it out.

---

<sup>22</sup>Every unordered pair of the five income sources studied.

<sup>23</sup>Viewed in aggregate, falls and rises should be equally likely, since the control group will account for any trends in the income source. Hence we can see that collectively taxpayers with dividend income rapidly start underreporting income again, although we cannot identify which individuals are the ones underreporting.



## 5 Impacts by Audit Outcome

We next consider how dynamic effects vary depending on the outcome of audit. This is important for policy, as it helps distinguish whether merely the process of being audited is enough to impact reported income and tax. We find that those who were found to be correct do not respond, while those for whom errors were found increase reported tax. Being audited per se does not appear to increase reported tax, but those found to have underpaid are 18pp more likely to report higher tax owed after audit. We first describe the approach taken to study this question, since our previous control group cannot help us study effects by audit outcome. We then describe the findings highlighted above.

### 5.1 Empirical approach

Since we now wish to study audit impacts separately by audit outcome, we cannot use the earlier identification strategy. In the “placebo audit” group we cannot observe what audit outcomes would have been, so cannot construct separate control groups for each audit outcome.<sup>24</sup> Instead we now take an ‘event study’ approach. Our sample for each regression is the set of observations who are audited and found to have some particular outcome, e.g. found to be compliant. Within that sample, the timing of audit is random – there is nothing systematic that led individuals to be selected in a particular year within the sample. Hence we can compare the outcome for someone audited and found to have a particular status (e.g. to be compliant) with someone who *will be* audited and found to have the same status.

For our variable of interest we now focus on a binary variable measuring whether tax paid increases, rather than on the sizes of increase, as in Pomeranz (2015). In particular, we estimate a linear probability model where the outcome is whether tax paid in year  $t$  is larger than in the year before audit. Our interest now is understanding *which* individuals – when split by audit status – respond. This outcome is therefore preferred, since it compares individuals to their own history, and is equally responsive to increases for individuals across the distribution of taxes owed. It is also less sensitive to relatively extreme observations, which is more important in our event study approach because the sample size is now much smaller. Whereas previously we had a treatment group of 53,000 individuals, and could draw a large sample of controls from the non-audit population, now the entire sample is those selected for audit. That sample is then further split into subsamples by

---

<sup>24</sup>Gemmell and Ratto (2012) study the effect by audit outcome, comparing each treatment group to the original control group containing people with a mix of possible outcomes. This leaves open the problem that audit outcome is endogenous, and is being conditioned on for only those in the treatment group.

audit outcome status, making results more sensitive to outliers and reducing power. Use of a binary variable removes this sensitivity, without limiting our ability to study which groups respond.

## 5.2 Results by audit outcome

To assess the reasonableness of the approach, we again begin by studying the estimated impact in the years before audit. The first four rows of Table 5 provides the results for the pre-audit period. It can be seen that all the point estimates are close to zero, providing support for the validity of this approach. A second test of validity can be seen from the ‘Not audited’ column. This estimates the effect of being selected for audit on individuals who were never actually audited, nor informed that they had been selected. As expected, again the point estimates are very close to zero.

Turning to the columns, three results can be seen. First, those who were audited and found to have made no errors also do not respond. This is important because it tells us that the dynamic response isn’t driven by the mere fact of audit. Direct audit effects could happen, for example, if the process of audit were sufficiently unpleasant that taxpayers decided to err upwards when uncertain, in the hope of avoiding further audits. One could also potentially have seen negative direct effects in this group. If some taxpayers were incorrectly found to be compliant, they may learn that the tax authority is less effective at detecting non-compliance than they previously believed, and reduce payments. We find neither of these results: on average those whose returns are found correct do not change their reports.

Second, those who are found to have errors are more likely to report higher levels of tax in subsequent years. Even four years later they are 13-14pp more likely to report higher tax owed. Hence the long term effects observed appear to all come from correcting errors made by the taxpayer. Note that even those who made errors but owed no additional tax respond to the audit. This is because the errors made might affect future tax liability. For example, claiming excessively large expenses today might increase the size of a loss on property income that can be carried forward: correcting this increases future tax liabilities. Anecdotally, from speaking to audit officers, in some cases these individuals shift their reports to pay tax in the audit year so that they can smooth out the additional tax liability that they will now face over the coming years.

Third, those who receive a penalty appear to have been driving some of the shape of the dynamics we observed earlier, where we saw a peak two years after the year selected for audit. To illustrate this, Figure 6 displays graphically the point estimates from Table 5. Whilst those with mistakes but no penalty respond immediately, the response for those with a penalty peaks two years after

the year for which the audit is done. This reflects two features of the audit process. Firstly, those who ultimately receive penalties typically take longest to audit, since their underreporting requires more work to detect. The audit settlement date is thus later. If some taxpayers wait until the audit (and uncertainty about detection) is resolved to respond, this will delay the time until they are observed to respond. Second, taxpayers with mistakes but no penalties will have their *original* return corrected, so an immediate response is observed.<sup>25</sup> On the other hand, those who receive a penalty may not have their return corrected: in most cases they instead file a separate form detailing additional tax, interest, and penalties.<sup>26</sup>

## 6 Simple Model of Tax Evasion and Audit Response

To help understand the mechanism underlying the observed results, we consider an extended version of the model by Kleven et al. (2011), which is itself an extension of the model by Allingham and Sandmo (1972).

### 6.1 Model Outline

Taxpayers are risk-neutral, and choose how much tax to evade.<sup>27</sup> The probability of detection is endogenous and is increasing in the difference between reported and true income. The key idea underlying this assumption is that other contextual information the tax authority might have, such as where a taxpayer lives, provides some information on the taxpayer's income, so that reports further from the truth are more likely to be investigated.<sup>28</sup>

The key innovation of our model is to note that incomes from some sources, such as pension annuity income, are very autocorrelated, while other sources, such as self-employment income for a sole trader, are much less stable. By first extending the model of Kleven et al. (2011) to multiple time periods, and then allowing for differential autocorrelation of income sources we are able to distinguish different possible mechanisms for *why* audits are observed to have long term outcomes.

---

<sup>25</sup>Data from the uncorrected return are unfortunately not available, so the details of the corrections cannot be identified.

<sup>26</sup>These data are also not available for analysis.

<sup>27</sup>Relaxing the assumption of risk-neutrality would reduce evasion. It could also introduce a positive correlation between the level of income and evasion, assuming no outside wealth and decreasing absolute risk aversion. This is inconsistent with what we see empirically in our data. Instead we see that there is little variation across the income distribution in the probability of evasion (extensive margin) or the amount of tax evaded in cash terms (intensive margin). The latter result means the *share* of income evaded is falling across the income distribution. See Figure A1 for details.

<sup>28</sup>Other papers using this assumption include Allingham and Sandmo (1972), Yitzhaki (1987), Slemrod and Yitzhaki (2002), Sandmo (2005), and Kleven et al. (2011). It is also consistent with the objective of the tax authority: since not everyone can be audited, larger deviations from the authority's expectation are likely to yield the most revenue.

The importance of autocorrelation is that an audit provides a snapshot measure of the current true level of income. For income sources which are relatively stable, this will also be a good predictor of revenue in the future. For those which are less autocorrelated, the predictive power of this snapshot will decline quickly over time. As we discuss below, different mechanisms make distinct predictions about the dynamics of any audit impact *split by autocorrelation of income source*.

Consider a taxpayer with true income  $\tilde{y}_s$  in year  $s$ . This income can be decomposed into three parts: (i) a third party reported component,  $\tilde{y}_s^{\text{TPR}}$ ; (ii) a self-reported permanent component,  $\tilde{y}_s^{\text{perm}}$ ; and (iii) a self-reported stochastic component,  $\tilde{y}_s^{\text{stoch}}$ . The distinction between third-party reported and self-reported income can explain why evasion rates appear much lower than would be expected given the empirical probability of audit (Kleven et al., 2011). We distinguish within self-reported income sources those which are fixed over time, and those which are time-varying. This is a simplification for expositional purposes, but our main results can be generalised to having multiple self-reported income sources with varying degrees of autocorrelation, as we had in our empirical setting.

In year  $s$  the taxpayer reports an income of  $y_s$ , so evasion is  $e_s := \tilde{y}_s - y_s$ . Evasion is detected with probability  $p(e_s)$ , which is increasing in the level of evasion i.e.  $p'(e_s) > 0$ . This is a composite of both the probability of audit and of the audit successfully detecting evasion.

When evasion is detected, the taxpayer must pay the evaded tax and an additional penalty. The tax is proportional to income, at rate  $\tau$ , and the penalty is proportional to the tax evaded, with penalty rate  $\theta$ . The taxpayer's problem is therefore to choose an evasion rate  $e_s$  to maximise expected net-of-tax income:<sup>29</sup>

$$[1 - p(e_s)] [(1 - \tau)\tilde{y}_s + \tau e_s] + p(e_s) [(1 - \tau)\tilde{y}_s - \theta \tau e_s] \quad (2)$$

Differentiating with respect to evasion,  $e_s$ , gives the first order necessary condition for an interior optimum:

$$[p(e_s) + p'(e_s)e_s](1 + \theta) = 1 \quad (3)$$

Analysis of this problem is straightforward, and the literature studying this (from Allingham and Sandmo (1972) onwards) has focused on comparative static predictions with respect to the probability that evasion is detected, the penalty for evasion, and the marginal tax rate.

---

<sup>29</sup>Here we present the taxpayer's problem as a static decision, independent across periods. This simplifies exposition and is equivalent to assuming that the tax authority can neither audit old tax returns when it selects a taxpayer for audit, nor condition future audit probabilities on the outcomes of past audits. The latter assumption is an accurate description of our empirical context.

In this context the main innovation of Kleven et al. (2011) is to note that attempts to evade tax on third party reported income are highly unlikely to succeed. This is because matching of tax returns and third party reports will detect any discrepancies, and auditing these returns will uncover the evasion. On the other hand, evasion of self-reported income is much less likely to be noticed. To capture this intuition, the authors first note that a taxpayer who wishes to evade should evade first on the sources of income which are relatively less likely to be detected, before switching to the more easily detected. This means that any evasion up to  $\tilde{y}_s^{\text{self}} = \tilde{y}_s^{\text{perm}} + \tilde{y}_s^{\text{stoch}}$  should be of self-reported income, and only after this will taxpayers evade by misreporting  $\tilde{y}_s^{\text{TPR}}$ . Given this optimal structure for any evasion, the probability of detecting evasion is relatively low for  $e_s < \tilde{y}_s^{\text{self}}$ , but then increases quickly once  $e_s$  reaches  $\tilde{y}_s^{\text{self}}$ . Finally, given these detection probabilities, the optimal strategy for a taxpayer is to evade some amount less than  $\tilde{y}_s^{\text{self}}$ .

## 6.2 Implications of Audit

Our insight builds directly on these ideas. Consider an individual who is audited (for the first time) in year  $t$ . Until the time of audit his optimal reporting strategy was identical to that described above: he evaded some amount less than the total amount of self-reported income he received, and none of the misreporting related to third party reported income. Being audited may change his reporting for some combination of the following three reasons: (1) changes in the perceived penalty for evasion; (2) changes in the perceived probability that evasion is detected; and (3) updates to the information held by the tax authority.

In the first of these mechanisms, being audited ( $D = 1$ ) changes the perceived penalty i.e. beliefs about the penalty depend on audit status,  $\theta(D)$ , and vary with audit status,  $\theta(1) \neq \theta(0)$ . If this belief is revised upwards, so  $\theta(1) > \theta(0)$ , then the cost of evasion increases and evasion falls; if it is revised downwards then the opposite occurs. Note that this does not require any particular assumptions on either the initial beliefs or whether updating is rational. It is simply a statement of the direction in which these beliefs about  $\theta$  change. Absent any policy changes which shift beliefs about the penalty rate, this change is permanent so any change in behaviour will also be permanent. One might also expect that updating might be different for those who actually receive penalties, compared to those who don't. The effects we see are not permanent, and they do not differ between the non-compliant who do and do not receive penalties.<sup>30</sup>

The second mechanism supposes instead that the perceived probability that evasion is detected,  $p(e_s, h)$ , is now different because the taxpayer has been audited, and varies with how many years it

---

<sup>30</sup>As noted above, the apparent difference in the year of audit is mainly mechanical.

has been since audit,  $h = s - t$ .<sup>31</sup> If he believes he is initially being monitored more carefully than before, so  $p(e, 1) \geq p(e, 0)$ , this leads to higher compliance immediately after the audit. Alternatively if he believes he is now unlikely to be audited for some time, the so-called ‘bomb crater effect’ (Mittone, 2006; Maciejovsky et al., 2007; and Kastlunger et al., 2009), then  $p(e, 1) \leq p(e, 0)$ , and compliance initially falls. As time since audit increases, the (perceived) effect of having just been audited wears off, so that beliefs about the probability of evasion being detected converge back towards baseline i.e.  $|p(e, h + 1) - p(e, 0)| \leq |p(e, h) - p(e, 0)|$  for  $h > 0$ .<sup>32</sup> Again this does not impose anything on where the perceived audit probability comes from, only on how it is updated. The implication is that as the perceived probability of audit converges back to its initial level, the initial impact on reporting behaviour will decline back to baseline. This convergence will be common across incomes from all sources since the probability is common across them all. Again this mechanism does not explain our findings, since we see a differential decline in reporting across income sources, even within individual. This is not consistent with a response driven purely by a differential probability of audit in subsequent years. Instead it can only be explained by a mechanism which leads to a differential shift in reporting behaviour across income sources over time.

The final mechanism by which audits might affect reporting is that they differentially change the ability to hide certain sources of income. Performing an audit provides the tax authority with more accurate information on a taxpayer’s income at a point in time. In subsequent years, information from the audit will make evasion of less variable income sources easier to detect. To operationalise this, recall our earlier distinction between the permanent and stochastic components of self-reported income,  $\{\tilde{y}^{\text{perm}}, \tilde{y}_{t+h}^{\text{stoch}}\}$ . Once the tax authority performs audits a taxpayer,  $\tilde{y}^{\text{perm}}$  for that taxpayer is observed. After this, evasion of the permanent component of income is easily detected, analogous to the case with third party reported income. Hence taxpayers should now evade by misreporting  $\tilde{y}_{t+h}^{\text{stoch}}$  before any misreporting of  $\tilde{y}^{\text{perm}}$  (or  $\tilde{y}_{t+h}^{\text{TPR}}$ ).

Also, if this information is the mechanism by which misreporting can be uncovered, then as the amount of the information about stochastic income is reduced over time, misreporting becomes easier. In particular, information about past income is useful because the tax authority can compare reported income in some period  $h$  to their expectation of income given the past observation in the audit year. Such deviations will be more informative if past incomes are a good predictor of current

---

<sup>31</sup>For a taxpayer that has not yet been audited,  $h$  will be negative and  $p(e_s, h)$  will take the same value for all  $h < 0$ .

<sup>32</sup>Note that, given the risk-neutrality assumption, *permanent* shifts in the level of  $p(e, h)$  when  $h > 0$  versus  $h \leq 0$  are observationally equivalent to a shift in  $\theta$ . So any permanent shift in perceived audit probability,  $p(e, h) - p(e, 0)$  as  $h \rightarrow \infty$  is observationally equivalent to some shift in the perceived penalty  $\theta(1) - \theta(0)$ . This equivalence would breakdown with risk aversion, which would allow separate testing of these hypotheses.

income i.e. the autocorrelation of stochastic income,  $\rho(h, \text{stoch}) = \text{Corr}(\tilde{y}_{t+h}^{\text{stoch}}, \tilde{y}_t^{\text{stoch}})$ , is high.<sup>33</sup> As  $h$  increases,  $\rho(h, \text{stoch})$  falls, so the ‘amount of information’ about current stochastic income is lower and so misreporting becomes easier. The case of permanent income is the natural limit of this case, where  $\rho(h, \text{perm}) = 1 \forall h$ . This also makes it clear that if multiple sources of stochastic income were available, misreporting should increase more quickly for those sources which have a lower autocorrelation. Hence under this mechanism, the prediction is that the initial impact on reporting behaviour will decline back to baseline, and this decline will be more rapid for income sources which have a lower autocorrelation. This is consistent with our findings, as seen in Figure 5.

## 7 Conclusion

This paper investigated the dynamic effects of audits on income reported in subsequent tax returns. Understanding these effects is important both from the perspective of quantifying the returns to the tax authority from an audit, and for assessing the mechanisms by which audits might influence taxpayer behaviour. To answer this question we exploited a random audit program run by the UK tax authority (HMRC) under which an average of around 4,900 individuals are randomly selected for audit each year. We used data on audits over the period 1998/99 to 2008/09 and tracked responses on tax returns between 1998/99 and 2011/12.

We established five main results. First, we showed that there is significant non-compliance among individual self assessment taxpayers, both in the share of taxpayers who are found non-compliant and the share of tax that is misreported. Second, we demonstrated that in our context third party reporting is important in influencing levels of compliance, but only on the *intensive* margin of how much tax is misreported. Third, we provided evidence of important *dynamic effects*, with the additional tax revenue over the five years post-audit equalling 1.5 times the direct revenue raised by audit. Fourth, we documented that a return to misreporting occurred more rapidly after audit for income sources which were less autocorrelated. Finally, we showed that only those who were found to have made mistakes responded to the audit. With the aid of a stylised model, we demonstrated that the observed dynamics are consistent only with audits revealing information to the tax authority, which makes misreporting certain income sources easier to detect for a period after the audit.

Our results have three main policy implications. First, taking dynamic effects into account substantially increases the estimated revenue impact of audits. The direct effect of an audit is (on average) £830, whilst the cumulative dynamic effect over the subsequent five years is £1230, 1.5

---

<sup>33</sup>In principle what is needed is that the absolute value of this autocorrelation is high, but in practice such autocorrelations are observed to be non-negative.

times the direct effect. This suggests that the optimal audit rate should be increased relative to the situation where there are no dynamic effects.

Second, the variation in dynamic effects observed across different income components alters the way in which targeted audits should be targeted: audits should focus more on individuals reporting types of income with the largest dynamic effects. For example, the peak annual impact on reported self-employment income for each self-employed individual is over £1,000, higher than other components. This suggests focusing more on individuals reporting self-employment income. Likewise, although the maximum annual impact on pension income is lower, it is persistent, so there may be more incentive to target individuals reporting pension of income.

Third, there are implications for setting optimal re-auditing strategies. Impacts for reported self-employment income and dividend income die away after about four years, so it might make sense to revisit these individuals after this sort of horizon. In contrast, the impact on reported pension income seems to persist for at least eight years, implying that there is less need to re-audit these individuals so soon.

Our findings also highlight the importance of further study of the indirect effect of tax compliance audits. One natural direction for further work would be to understand how the dynamic effects differ in the context of targeted audits, which are focused on individuals deemed likely to be non-compliant. A second avenue for exploration is the spillover effect of audits: does auditing taxpayers change the behaviour of other taxpayers with whom they interact. Better understanding these effects is crucial in determining optimal audit policy.



## Bibliography

- ALLINGHAM, M. AND A. SANDMO (1972): “Income Tax Evasion: A Theoretical Analysis,” *Journal of Public Economics*, 1, 323–338.
- ANDREONI, J., B. ERARD, AND J. FEINSTEIN (1998): “Tax Compliance,” *Journal of Economic Literature*, 36, 818–860.
- BEER, S., M. K. E. K. AND B. ERARD (2015): “Audit Impact Study,” in *National Taxpayer Advocate 2015 Annual Report to Congress, Volume 2: TAS Research and Related Studies*, Washington, DC, 67–99.
- BLOOMQUIST, K. (2013): “Incorporating Indirect Effects in Audit Case Selection: An Agent-Based Approach,” *IRS Research Bulletin*, forthcoming.
- CHOO, L., M. FONSECA, AND G. MYLES (2013): “Lab Experiment to Investigate Tax Compliance: Audit Strategies and Messaging,” Research Report 308, HM Revenue and Customs.
- DEBACKER, J., B. HEIM, A. TRAN, AND A. YUSKAVAGE (2018): “Once Bitten, Twice Shy? The Lasting Impact of IRS Audits on Individual Tax Reporting,” *Journal of Law and Economics*, 61, 1–35.
- ERARD, B. (1992): “The Influence of Tax Audits on Reporting Behaviour,” in *Why People Pay Taxes: Tax Compliance and Enforcement*, ed. by J. Slemrod, University of Michigan Press, chap. 5, 95–114.
- GEMMELL, N. AND M. RATTO (2012): “Behavioral Responses to Taxpayer Audits: Evidence from Random Taxpayer Inquiries,” *National Tax Journal*, 65, 33–58.
- HM REVENUE AND CUSTOMS (2016): “Measuring Tax Gaps 2016 edition,” Tech. rep.
- INTERNAL REVENUE SERVICE (2016): “Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008-2010,” Publication 1415 (rev. 5-2016), Research, Analysis & Statistics.
- KASTLUNGER, B., E. KIRCHLER, L. MITTONE, AND J. PITTERS (2009): “Sequences of Audits, Tax Compliance, and Taxpaying Strategies,” *Journal of Economic Psychology*, 30, 405–418.
- KLEVEN, H. J., M. KNUDSEN, C. T. KREINER, S. PEDERSEN, AND E. SAEZ (2011): “Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark,” *Econometrica*, 79, 651–692.

- LONG, S. AND R. SCHWARTZ (1987): “The Impact of IRS Audits on Taxpayer Compliance: A Field Experiment in Specific Deterrence,” Paper presented at the Annual Meeting of the Law and Society Association, Washington DC.
- LUTTMER, E. F. P. AND M. SINGHAL (2014): “Tax Morale,” *Journal of Economic Perspectives*, 28, 149–168.
- MACIEJOVSKY, B., E. KIRCHLER, AND H. SCHWARZENBERGER (2007): “Misperception of Chance and Loss Repair: On the Dynamics of Tax Compliance,” *Journal of Economic Psychology*, 28, 678–691.
- MITTONE, L. (2006): “Dynamic Behaviour in Tax Evasion: An Experimental Approach,” *Journal of Socio-Economics*, 35, 813–835.
- POMERANZ, D. (2015): “No taxation without information: Deterrence and self-enforcement in the value added tax,” *American Economic Review*, 105, 2539–69.
- ROSENBAUM, P. AND D. RUBIN (1985): “Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score,” *American Statistician*, 39, 33–38.
- RUBIN, D. (2001): “Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation,” *Health Services and Outcomes Research Methodology*, 2, 169–188.
- SANDMO, A. (2005): “The Theory of Tax Evasion: A Retrospective View,” *National Tax Journal*, 58, 643–663.
- SLEMROD, J. AND S. YITZHAKI (2002): “Tax Avoidance, Evasion, and Administration,” in *Handbook of Public Economics III*, ed. by A. J. Auerbach and M. Feldstein, Amsterdam: Elsevier.
- TAUCHEN, H., A. WITTE, AND K. BERON (1993): “Tax Compliance: An Investigation Using Individual Taxpayer Compliance Measurement Program (TCMP) data,” *Journal of Quantitative Criminology*, 9, 177–202.
- YITZHAKI, S. (1987): “On the Excess Burden of Tax Evasion,” *Public Finance Quarterly*, 15, 123–137.

## Tables and Figures

**Table 1: Share of Taxpayers with Each Source of Income**

Income component	Proportion
Interest	.587
Employment	.482
Self employment	.375
Dividends	.370
Pensions	.300
Property	.136
Foreign	.048
Trusts and estates	.010
Share schemes	.002
Other	.030

***Notes:** Annual averages for tax years 1998/99 to 2008/09. Includes observations in year selected for audit or placebo audit.*

***Source:** Authors' calculations based on HMRC administrative datasets.*

**Table 2: Random Audit Outcomes**

	Mean	Std. dev.
Proportion of audited returns deemed		
Correct	.532	.499
Incorrect but no underpayment	.111	.314
Incorrect with underpayment (non-compliant)	.357	.479
Mean additional tax if non-compliant (£)	2,314	7,758
Distribution of additional tax if non-compliant		
Share £1-100	.116	.320
Share £101-1,000	.483	.500
Share £1,001-10,000	.361	.480
Share £10,001+	.039	.194
Observations	34,630	

**Notes:** Annual averages for tax years 1998/99 to 2008/09. Includes all individuals with a completed random audit.

**Source:** Authors' calculations based on HMRC administrative datasets.

**Table 3: Sample Balance, Conditioning on Filing History**

Years after audit		-5	-4	-3	-2	-1
Characteristics						
Female	Mean	.274	.276	.278	.282	.287
	Difference	-.005	-.006	-.005	-.006	-.005
	p-value	.236	.212	.292	.234	.338
Age	Mean	49.2	49.3	49.3	49.4	49.5
	Difference	.0	.0	.1	.1	.1
	p-value	.472	.600	.188	.170	.110
In London or SE	Mean	.333	.334	.335	.333	.331
	Difference	-.003	.001*	.003	.002	.002
	p-value	.159	.026	.015	.317	.190
Has tax agent	Mean	.628	.614	.603	.589	.573
	Difference	-.003	-.001	-.001	.002	.002
	p-value	.522	.500	.376	.675	.606
Income and tax totals						
Total taxable income	Mean	35,075	34,670	34,030	32,912	31,755
	Difference	-2	35	-163	71	56
	p-value	.979	.469	.012	.280	.439
Total tax	Mean	9,646	9,539	9,321	8,979	8,635
	Difference	14	12	-40	12	15
	p-value	.982	.288	.061	.261	.887
Income components						
Employment	Mean	22,508	22,534	22,266	21,708	21,145
	Difference	11	-57	-98	112*	43*
	p-value	.758	.023	.152	.049	.05
Self employment	Mean	6,546	6,379	6,200	5,950	5,581
	Difference	56	38	-49	-18	29
	p-value	.298	.435	.033	.161	.684
Interest and dividends	Mean	4,007	3,905	3,895	3,759	3,645
	Difference	-26	16	-27	7	4
	p-value	.667	.189	.235	.958	.086
Pensions	Mean	3,493	3,542	3,561	3,562	3,531
	Difference	-23	-23	-3	4	22
	p-value	.806	.482	.681	.463	.523
Property	Mean	869	844	811	769	726
	Difference	-5	-6	0	6	0
	p-value	.282	.209	.525	.072	.518
Foreign	Mean	194	193	194	181	169
	Difference	1	-1	5	-4	0
	p-value	.627	.137	.526	.117	.925
Trusts and estates	Mean	150	145	145	131	123
	Difference	17	-1	-6	4	3
	p-value	.153	.963	.204	.686	.824
Share schemes	Mean	91	104	68	62	55
	Difference	8*	-8	-1	9	1
	p-value	.019	.711	.783	.683	.243
Other	Mean	80	75	76	73	71
	Difference	3	-3	0	1	2
	p-value	.618	.184	.228	.984	.194

**Notes:** ‘Mean’ is the mean outcome in the control (not selected for audit) group across all years. ‘Difference’ is the coefficient on the treatment dummy in a regression of the outcome on a treatment dummy and dummies for whether the taxpayer filed taxes in each of the four years before audit (or placebo audit for controls). Treatment dummy equals 1 if taxpayer was selected by HMRC for a random audit. p-values are derived from an F-test that coefficients on interactions between treatment and tax year dummies are all zero in a regression of the outcome of interest on tax year dummies, interactions between treatment and tax year dummies, and dummies for whether the taxpayer filed taxes in each of the four years before audit (or placebo audit for controls). This is a stronger test than just testing the coefficient on treatment not interacted. Tests for all outcomes other than ‘survives’ are conditional on survives = 1. Monetary values are in 2012 prices. Standard errors are clustered by taxpayer. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ . **Source:** Authors’ calculations based on HMRC administrative datasets.

**Table 4: Autocorrelation by Income Source**

	Corr(t,t-1)	Corr(t,t-2)	Corr(t,t-3)
Pension income	.946	.904	.864
Property income	.896	.836	.790
Employment income	.862	.769	.690
Interest income	.835	.722	.640
Self-employment income	.832	.728	.644
Dividend income	.813	.723	.657
Sample size	4,506,548	4,506,548	4,506,548

*Notes:* Annual averages for years 1998/99 to 2011/12.

*Source:* Calculations based on HMRC administrative datasets.

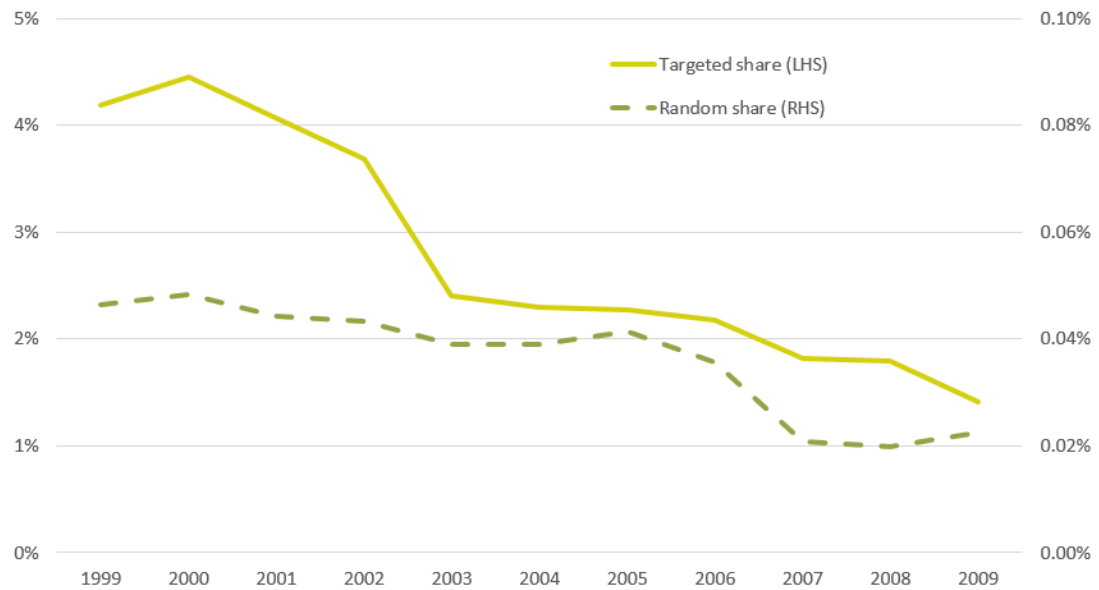
**Table 5: Impact by Audit Outcome**

Years since audit	Overall	Correct	Mistake nonpos	Mistake pos	Pos. yield + penalty	Not audited
-5	-.006 (.013)	-.042* (.018)	.048 (.049)	.033 (.032)	-.014 (.072)	-.002 (.030)
-4	.007 (.014)	-.034 (.019)	.068 (.049)	.050 (.033)	.037 (.068)	-.006 (.030)
-3	.005 (.014)	-.023 (.019)	.058 (.050)	.039 (.033)	.042 (.068)	-.016 (.030)
-2	.022 (.014)	-.005 (.019)	.079 (.050)	.075* (.033)	.032 (.068)	-.008 (.030)
-1	<i>Outcome is difference from -1 so zero by construction</i>					
0	.056*** (.014)	.016 (.019)	.131* (.051)	.179*** (.033)	.092 (.069)	-.014 (.030)
1	.048*** (.014)	.012 (.019)	.109* (.051)	.174*** (.033)	.180** (.069)	-.037 (.030)
2	.042** (.013)	.007 (.020)	.135** (.051)	.152*** (.033)	.207** (.069)	-.052 (.031)
3	.030* (.014)	-.007 (.020)	.135** (.052)	.133*** (.034)	.171* (.069)	-.048 (.031)
4	.031* (.014)	-.0024 (.021)	.134* (.052)	.137*** (.034)	.143* (.070)	-.045 (.031)
5	.033* (.016)	.019 (.023)	.160** (.056)	.119** (.037)	.128 (.074)	-.052 (.034)
N	124,223	46,911	9,519	25,666	6,983	35,144

**Notes:** ‘Overall’ uses the full sample of audited individuals to perform an event study, for whether tax paid is higher in each of the years before/after audit than the year immediately before audit (‘-1’). Coefficients from a linear probability model are shown, with standard errors in parentheses. Other columns split the audited sample by audit outcome: tax return found to be correct; tax return found to have a mistake but which doesn’t change tax liability (or in a small number of cases reduced liability); tax return found to have a mistake leading to increased tax liability, but no penalty charged (i.e. treated as legitimate error); tax return found to have underreported liability and a penalty charged (i.e. deemed to be deliberate); tax return selected for audit but no audit actually implemented (placebo test). \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

**Source:** Authors’ calculations based on HMRC administrative datasets.

Figure 1: Change in the probability of audit over time

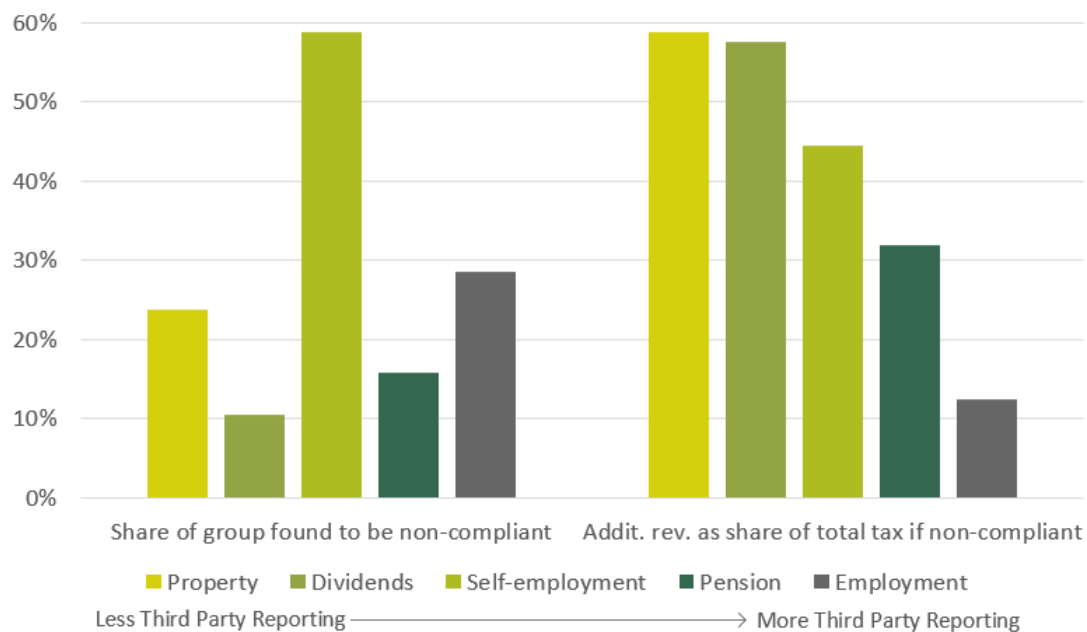


**Notes:** Constructed using data on individuals who received an audit of their self assessment tax return for a tax year between 1998/1999 and 2008/2009, and the full sample of self assessment returns for the same period.

**Source:** Calculations based on HMRC administrative datasets.



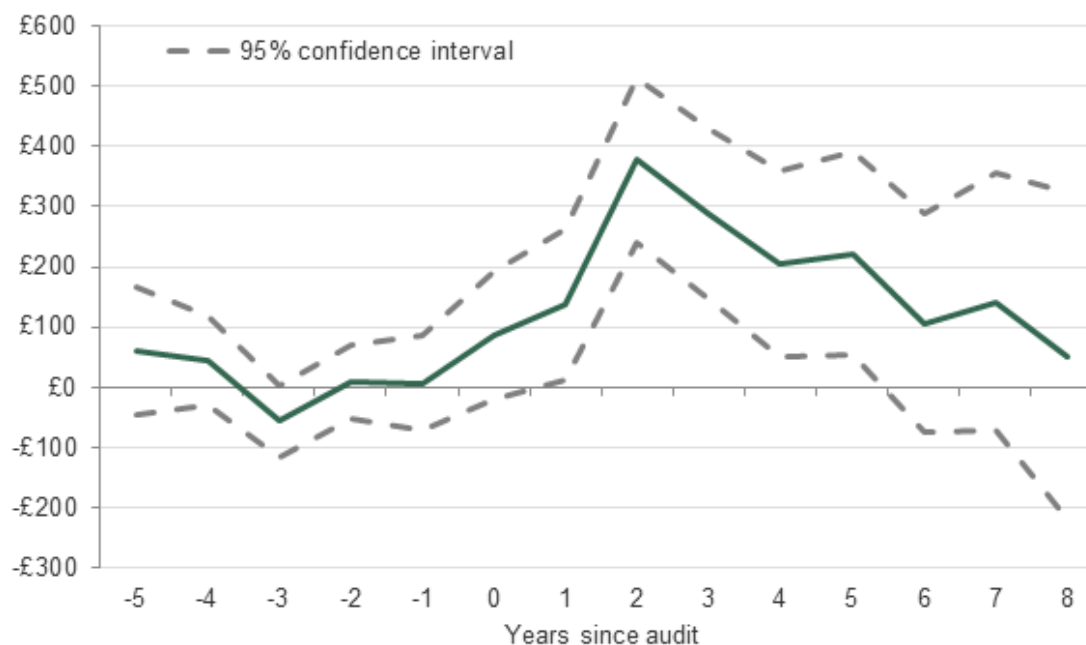
**Figure 2: Third party information doesn't explain the extensive margin of non-compliance, but does explain the intensive margin**



**Notes:** Constructed using data on individuals who received a random audit of their self assessment tax return for a tax year between 1998/1999 and 2008/2009. Income sources are ordered from least third party reporting on the left, to most third party reporting on the right. See Table A2 for details. 'Share of group found non-compliant' is the share of individual taxpayers receiving income from only that income source, plus potentially bank interest, who are found to owe additional tax when audited. 'Additional revenue as a share of total tax if non-compliant' is the additional tax owed divided by total tax owed, averaged across individual taxpayers who owed a positive amount i.e. who were non-compliant.

**Source:** Calculations based on HMRC administrative datasets.

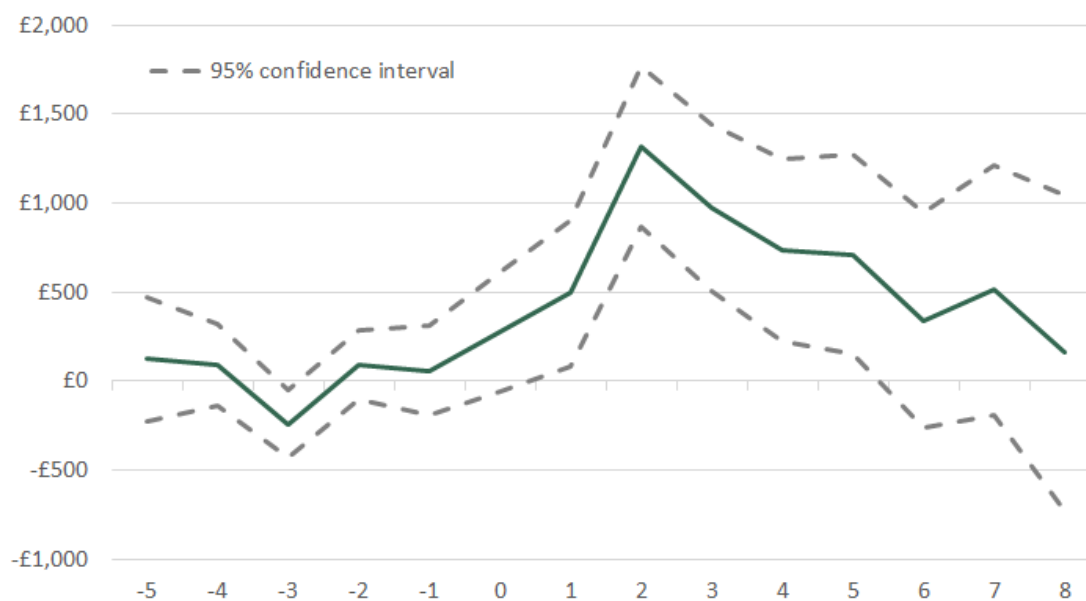
**Figure 3: Dynamic effect of audits on total reported tax owed**



**Notes:** Sample includes individuals selected for a random audit between 1998/99 and 2008/09, and control individuals who could have been selected in the same years but were not. The solid line plots the point estimate for the difference in average 'total reported tax' between individuals who were and weren't audited, for different numbers of years after the audit. This comes from a regression of total reported tax on dummies for years since audit (or placebo audit for controls), dummies for years since audit (or placebo audit for controls) interacted with treatment status, tax year dummies, and dummies for whether the taxpayer filed a return in each of the four years before audit, with audit status instrumented by selection for audit. Standard errors are clustered at the individual level.

**Source:** Calculations based on HMRC administrative datasets.

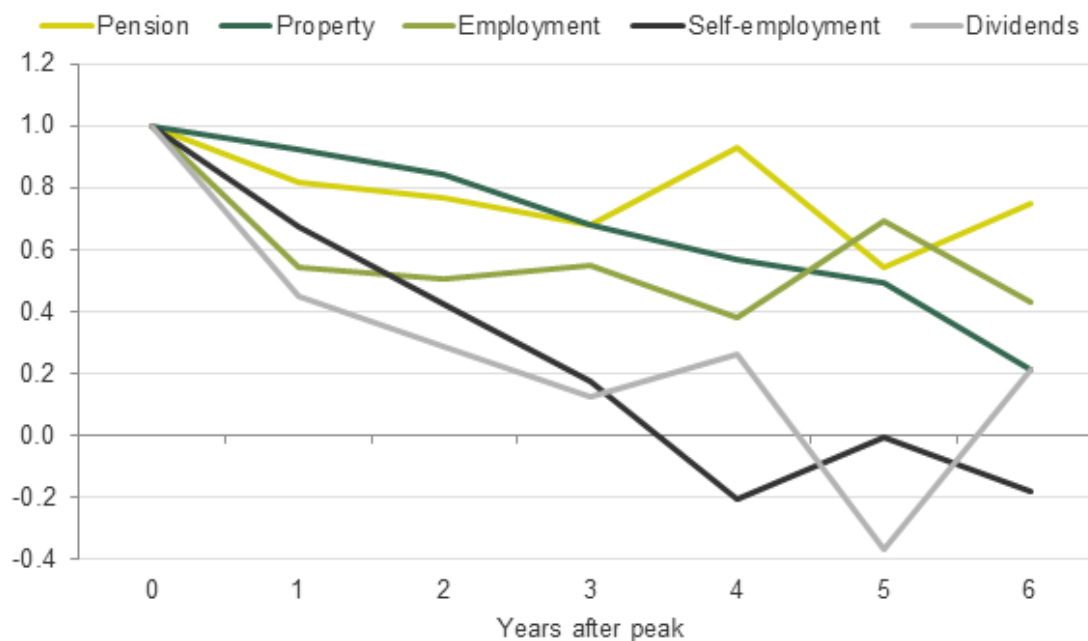
**Figure 4: Dynamic effect of audits on total reported income**



**Notes:** Sample includes individuals selected for a random audit between 1998/99 and 2008/09, and control individuals who could have been selected in the same years but were not. The solid line plots the point estimate for the difference in average 'total reported income' (income from all sources) between individuals who were and weren't audited, for different numbers of years after the audit. This comes from a regression of total reported income on dummies for years since audit (or placebo audit for controls), dummies for years since audit (or placebo audit for controls) interacted with treatment status, tax year dummies, and dummies for whether the taxpayer filed a return in each of the four years before audit, with audit status instrumented by selection for audit. Standard errors are clustered at the individual level.

**Source:** Calculations based on HMRC administrative datasets.

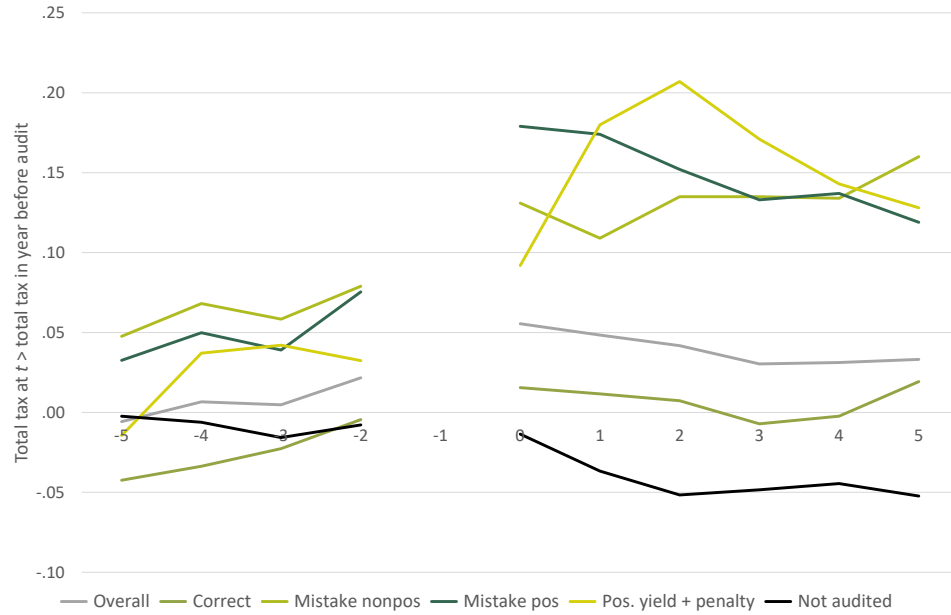
**Figure 5: Relative dynamics by income source: less autocorrelated sources of income see faster declines**



**Notes:** Sample includes individuals selected for a random audit between 1998/99 and 2008/09, and control individuals who could have been selected in the same years but were not. Each line plots the point estimate for the difference in the average of a particular component of income between individuals who were and weren't audited, for different numbers of years after the audit. This comes from a regression of each income component on dummies for years since audit (or placebo audit for controls), dummies for years since audit (or placebo audit for controls) interacted with treatment status, tax year dummies, and dummies for whether the taxpayer filed a return in each of the four years before audit, with audit status instrumented by selection for audit.

**Source:** Calculations based on HMRC administrative datasets.

**Figure 6: Response to audit by audit outcome: only those with errors respond**



**Notes:** Sample includes individuals selected for a random audit between 1998/99 and 2008/09. Each line plots the point estimate for the share of individuals who report higher levels of tax relative to the year before they are audited, for different numbers of years before and after the audit. This comes from a series of regressions of a dummy for ‘reports more tax than the year before audit’ on dummies for years since audit (excluding  $-1$ ) and for tax year, one for each audit outcome status listed. ‘Overall’ uses all the observations selected for audit. ‘Correct’ uses observations for individuals who were audited and found to have no errors. ‘Mistake nonpos’ uses observations for individuals who were audited and found to have errors that did not yield any tax in the tax year for which the return was completed. ‘Mistake pos’ uses observations for individuals who were audited and found to have underreported, but not to the extent that a penalty was charged. ‘Pos. yield + penalty’ uses observations for individuals who were audited and found to have underreported in a way that was deemed deliberate, and hence a penalty was charged. ‘Not audited’ uses observations for individuals who were selected for audit, but for whom no audit was carried out.

**Source:** Calculations based on HMRC administrative datasets.

# Appendices

## Appendix A Additional Tables and Figures

**Table A1: Random audit lags and durations**

	Mean	Std. dev.	Median	75th	90th
Lag to audit start (months)	8.9	4.0	9	11	14
Audit duration (months)	5.3	6.6	3	7	13
Total time to audit end (months)	14.3	7.3	13	17	23

**Notes:** Annual averages for tax years 1998/99 to 2008/09. Includes all individuals with a completed random audit.

**Source:** Authors' calculations based on HMRC administrative datasets.

**Table A2: Third party reporting arrangements in the UK**

Income component	Degree of third-party reporting
Employment	Complete
Interest	Complete
Pensions	Partial, via pension provider
Self employment	None unless an entertainer, sportsman, or contractor in the construction industry
Dividends	None
Property	None

***Source:** Personal communication with Tracey Bowler, Tax Law Review Committee.*

**Table A3: Sample balance**

Years after audit		-5	-4	-3	-2	-1
Characteristics						
Female	Mean	.274	.276	.278	.282	.287
	Difference	-.006	-.004	-.002	-.001	-.002
	p-value	.221	.359	.606	.627	.863
Age	Mean	49.2	49.3	49.3	49.4	49.5
	Difference	.2	.3	.3	.2	.2
	p-value	.756	.586	.390	.610	.057
In London or SE	Mean	.333	.334	.335	.333	.331
	Difference	-.006	.001*	.004*	.002	.002
	p-value	.177	.025	.011	.281	.152
Has tax agent	Mean	.628	.614	.603	.589	.573
	Difference	.000	.002	.001	.003	.005
	p-value	.547	.508	.405	.396	.412
Survives	Mean	.624	.669	.728	.803	.892
	Difference	.032***	.039***	.047***	.050***	.050***
	p-value	.000	.000	.000	.000	.000
Income and tax totals						
Total taxable income	Mean	35,075	34,670	34,030	32,912	31,755
	Difference	881	492	403*	1,051*	1,095*
	p-value	.374	.157	.028	.012	.012
Total tax	Mean	9,646	9,539	9,321	8,979	8,635
	Difference	260	63	82	310	337*
	p-value	.539	.303	.055	.064	.027
Income components						
Employment	Mean	22,508	22,534	22,266	21,708	21,145
	Difference	-31	-136	180*	909**	721*
	p-value	.162	.371	.028	.006	.027
Self employment	Mean	6,546	6,379	6,200	5,950	5,581
	Difference	356	328	173	99	200*
	p-value	.151	.174	.311	.106	.025
Interest and dividends	Mean	4,007	3,905	3,895	3,759	3,645
	Difference	-36	208	18	63	112
	p-value	.767	.432	.700	.578	.580
Pensions	Mean	3,493	3,542	3,561	3,562	3,531
	Difference	176	168	128	148	159
	p-value	.425	.478	.642	.307	.327
Property	Mean	869	844	811	769	726
	Difference	18	-2	37	47	31
	p-value	.813	.952	.576	.498	.134
Foreign	Mean	194	193	194	181	169
	Difference	23	-1	-5	-6	1
	p-value	.759	.240	.956	.317	.766
Trusts and estates	Mean	150	145	145	131	123
	Difference	46	17	8	19	8
	p-value	.245	.367	.290	.125	.367
Share schemes	Mean	91	104	68	62	55
	Difference	17*	24	22**	11	-6**
	p-value	.043	.062	.004	.132	.008
Other	Mean	80	75	76	73	71
	Difference	-1	-4	0	5	6
	p-value	.747	.675	.645	.796	.167

**Notes:** ‘Mean’ is the mean outcome in the control (not selected for audit) group across all years. ‘Difference’ is the coefficient on the treatment dummy in a regression of the outcome on a treatment dummy. Treatment dummy equals 1 if taxpayer was selected by HMRC for a random audit. p-values are derived from an F-test that coefficients on interactions between treatment and tax year dummies are all zero in a regression of the outcome of interest on tax year dummies and interactions between treatment and tax year dummies. This is a stronger test than just testing the coefficient on treatment not interacted. Tests for all outcomes other than ‘survives’ are conditional on survives = 1. Monetary values are in 2012 prices. Standard errors are clustered by taxpayer. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ . **Source:** Authors’ calculations based on HMRC administrative datasets.



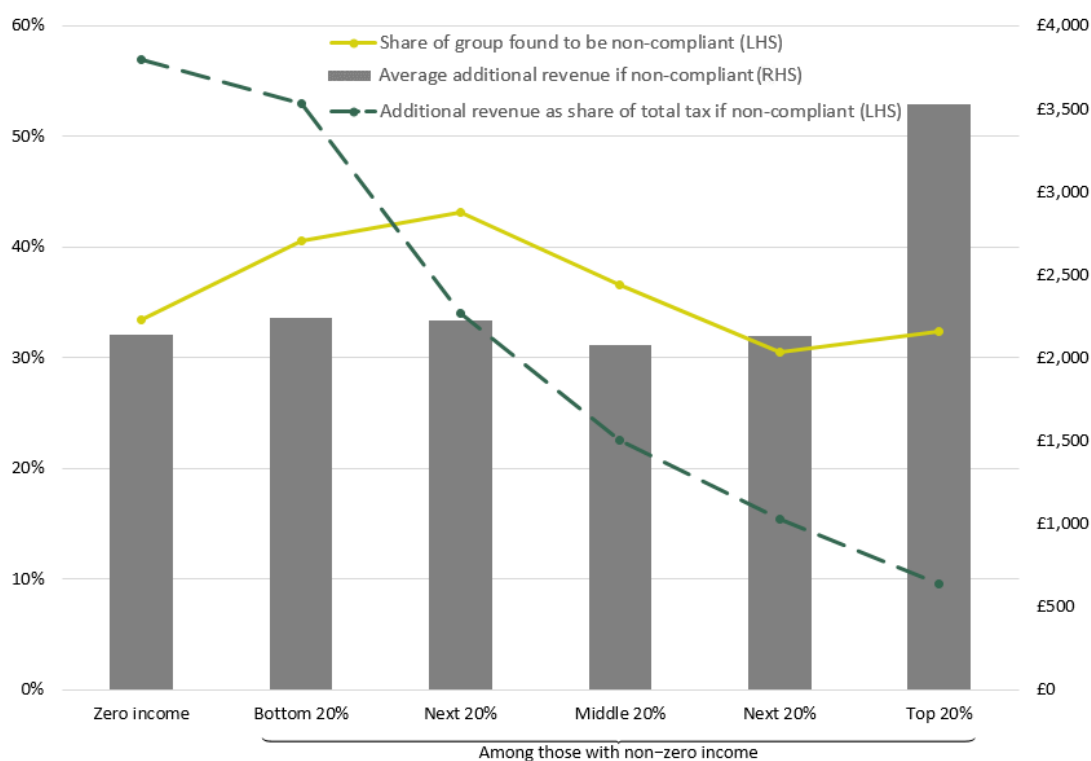
**Table A4: Cases opened quickly uncover more underreported tax**

Lag between filing and audit start (months)	Mean underpaid tax (£)	Standard Error (£)	N
$\leq 2$	1460	249	614
3	1120	145	921
4	953	86	1,338
5	882	97	1,369
6	1040	100	1,409
7	896	78	1,454
8	834	76	1,564
9	923	96	1,591
10	784	86	1,906
11	784	99	2,286
12	600	73	1,101
13-21	693	129	3,006

**Notes:** Annual averages for tax years 1998/99 to 2008/09. Includes all individuals with a completed random audit. Audits opened within two months of filing are grouped together to satisfy disclosivity restrictions.

**Source:** Authors' calculations based on HMRC administrative datasets.

**Figure A1: The share of income which is misreported is falling across the reported income distribution**



**Notes:** Constructed using data on individuals who received a random audit of their self assessment tax return for a tax year between 1998/1999 and 2008/2009. Income grouping is done based on previous year's reported income. 16.2% of individuals report having zero income in the previous year. The remaining individuals are divided into five equal sized bins based on their previous income: quintiles conditional on reporting non-zero income. 'Share of group found to be non-compliant' is the share of individual taxpayers receiving income from only that income source, plus potentially bank interest, who are found to owe additional tax when audited. 'Average additional revenue if non-compliant' is the average total tax in 2012 £ that was not reported among those individuals for whom some tax was not reported (the non-compliant). 'Additional revenue as a share of total tax if non-compliant' is the additional tax owed divided by total tax owed, averaged across individual taxpayers who were non-compliant.

## Appendix B Additional Evidence of Observable Heterogeneity in Compliance

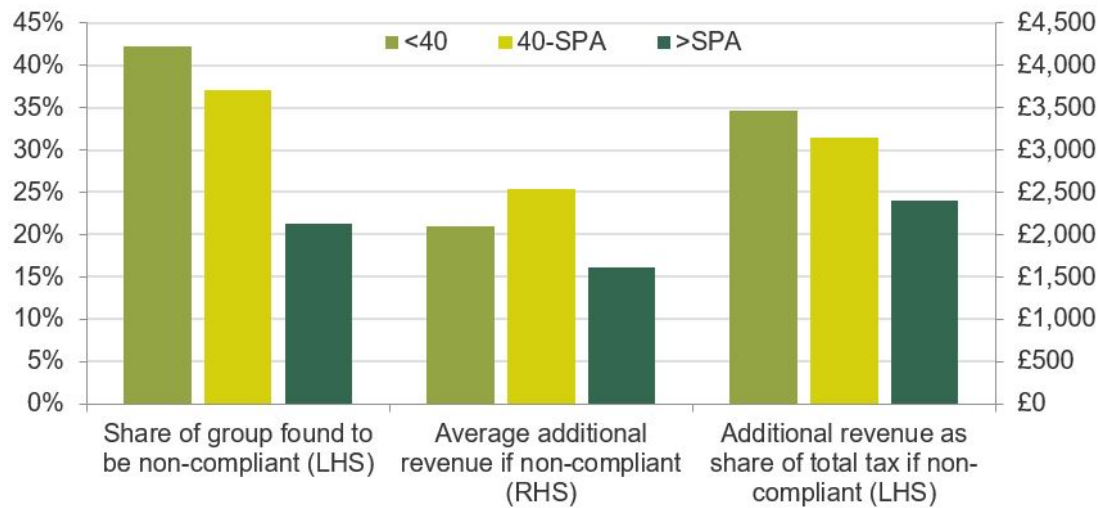
In this section we describe the heterogeneity in the direct impact of audits.

As seen in Table 2, 36 per cent of taxpayers are found to be non-compliant i.e. to have underreported tax due. Among these taxpayers, 32 per cent of the tax they owe was not reported, around £2,320 in cash terms (2012 prices). Figures 2 and A1 already provide evidence on how these results look across different income sources and across the distribution of reported income.

Figures B1 and B2 show how non-compliance varies by age and sex. These results are consistent with Kleven et al. (2011), who find that being older and being female are negatively associated with evasion. However, while they find that only the result on sex is statistically significant, these differences are all significant in our case. Table B1 shows that these results also hold conditional on additional predictors of non-compliance, including age, sex, income source, industry, and region. Note that these are not to be interpreted causally, since many of these characteristics are choices for the taxpayer. Nevertheless, the associations are interesting to understand, and are relevant for current targeting of audits, though of course using them for targeting is subject to the Lucas critique.

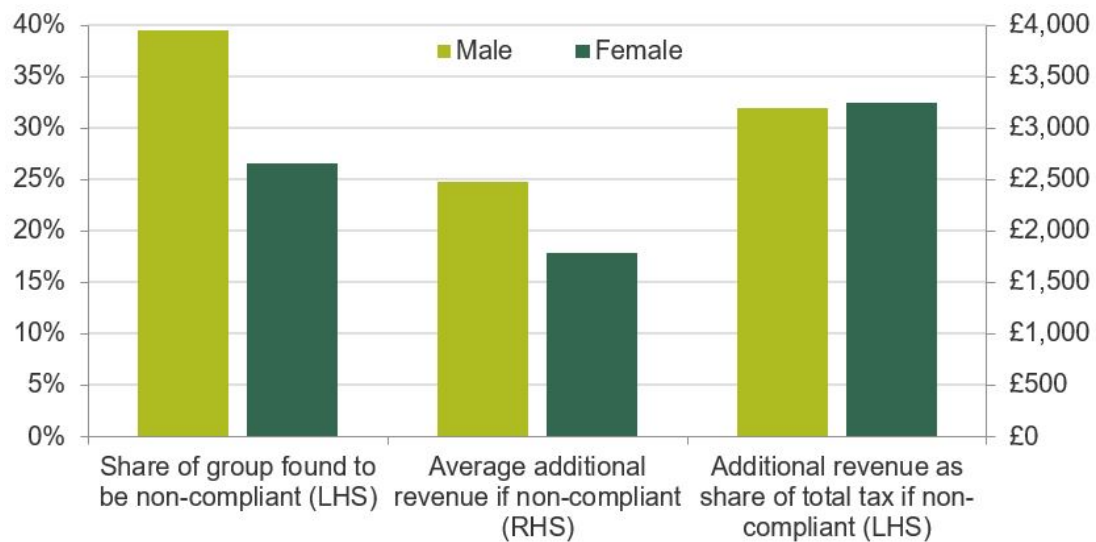
In contrast to Kleven et al. (2011), being located in the capital city is not associated with lower evasion. Breaking this down further, the share of Londoners found to be non-compliant is higher than in the other nations of the UK (where England is calculated excluding London) except Northern Ireland (see Figure B3). Again these differences are all statistically significant, and hold when other characteristics are controlled for.

Figure B1: Non-compliance by age



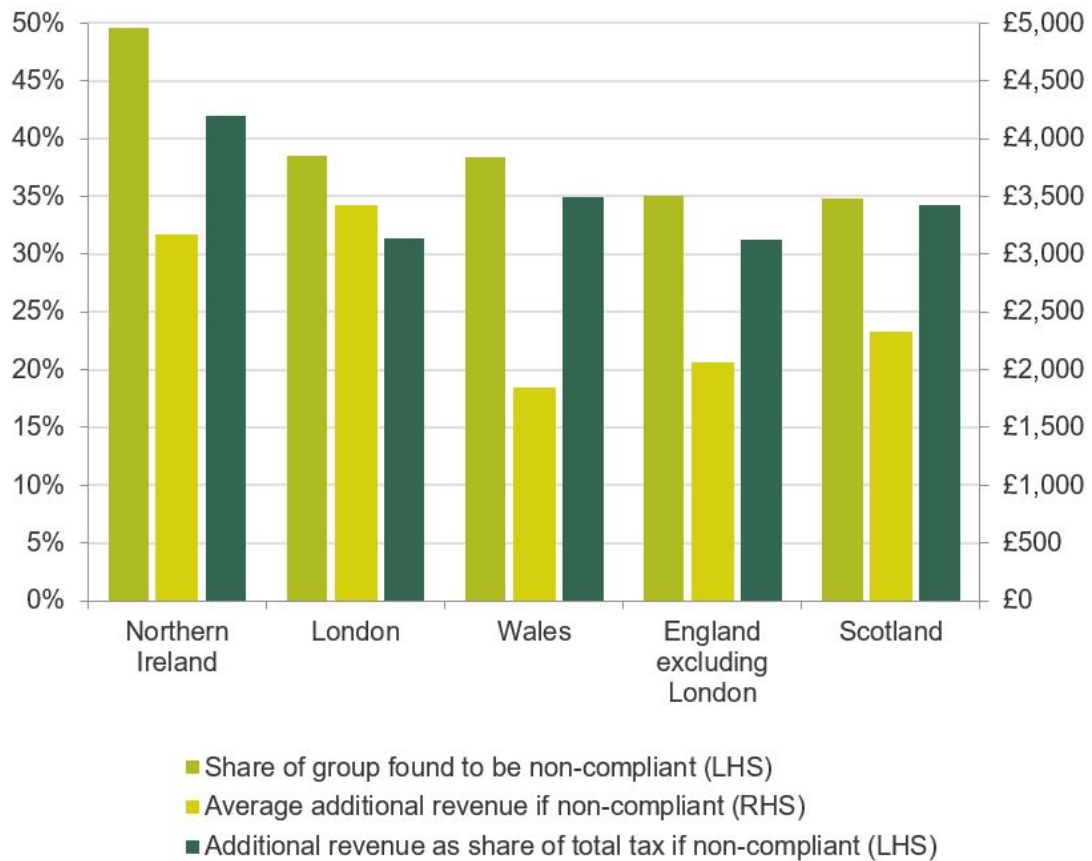
**Notes:** Constructed using data on individuals who received a random audit of their self assessment tax return for a tax year between 1998/1999 and 2008/2009. SPA is the state pension age, which in the UK was 65 for men and 60 for women over this period. 'Share of group found to be non-compliant' is the share of individual taxpayers receiving income from only that income source, plus potentially bank interest, who are found to owe additional tax when audited. 'Average additional revenue if non-compliant' is the average total tax in 2012 £ that was not reported among those individuals for whom some tax was not reported (the non-compliant). 'Additional revenue as a share of total tax if non-compliant' is the additional tax owed divided by total tax owed, averaged across individual taxpayers who were non-compliant.

Figure B2: Non-compliance by sex



**Notes:** Constructed using data on individuals who received a random audit of their self assessment tax return for a tax year between 1998/1999 and 2008/2009. 'Share of group found to be non-compliant' is the share of individual taxpayers receiving income from only that income source, plus potentially bank interest, who are found to owe additional tax when audited. 'Average additional revenue if non-compliant' is the average total tax in 2012 £ that was not reported among those individuals for whom some tax was not reported (the non-compliant). 'Additional revenue as a share of total tax if non-compliant' is the additional tax owed divided by total tax owed, averaged across individual taxpayers who were non-compliant.

Figure B3: Non-compliance by region



**Notes:** Constructed using data on individuals who received a random audit of their self assessment tax return for a tax year between 1998/1999 and 2008/2009. Region is either London or the nation of the UK (where London is excluded for England) in which the taxpayer's home address is located. 'Share of group found to be non-compliant' is the share of individual taxpayers receiving income from only that income source, plus potentially bank interest, who are found to owe additional tax when audited. 'Average additional revenue if non-compliant' is the average total tax in 2012 £ that was not reported among those individuals for whom some tax was not reported (the non-compliant). 'Additional revenue as a share of total tax if non-compliant' is the additional tax owed divided by total tax owed, averaged across individual taxpayers who were non-compliant.

**Table B1: Predictors of non-compliance**

Dependent Variable: Taxpayer is found non-compliant

Sample: Self-assessment taxpayers subject to random audit on tax returns for 1999-2009

Standard errors (in parentheses) clustered at individual level

	(1)	(2)	(3)
Female	-.417*** (.031)	-.417*** (.031)	-.425*** (.032)
Age	-.013*** (.001)	-.014*** (.001)	-.014*** (.001)
Age <sup>2</sup>		-.000*** (.000)	-.000*** (.000)
Has self-employment income	.831*** (.051)	.818*** (.051)	.787*** (.051)
Has employment income	-.152*** (.033)	-.176*** (.033)	-.213*** (.034)
Has property income	.368*** (.040)	.362*** (.040)	.384*** (.040)
Has pension income	-.161*** (.039)	-.081 (.042)	-.100* (.043)
Manufacturing	-.085 (.119)	-.108 (.119)	-.063 (.122)
Construction	.263** (.094)	.258** (.094)	.297** (.097)
Wholesale	.073 (.106)	.052 (.106)	.081 (.108)
Hospitality	.468** (.145)	.448** (.145)	.484** (.148)
Transport	.372*** (.111)	.345** (.112)	.383*** (.114)
Legal	-.179 (.134)	-.194 (.134)	-.150 (.137)
Financial	-.060 (.098)	-.077 (.098)	-.030 (.101)
Education	-.069 (.111)	-.085 (.111)	-.028 (.113)
Other industry	-.139 (.100)	-.147 (.100)	-.100 (.102)
Industry not recorded	-.428*** (.098)	-.430*** (.098)	-.363*** (.100)
Bottom income quintile	.350*** (.045)	.338*** (.045)	.344*** (.046)
Second income quintile	.469*** (.045)	.456*** (.045)	.454*** (.045)
Third income quintile	.426*** (.046)	.414*** (.046)	.396*** (.047)
Fourth income quintile	.509*** (.049)	.488*** (.049)	.461*** (.050)
Top income quintile	.653*** (.049)	.627*** (.050)	.605*** (.051)
Constant	-1.009*** (.103)	-.929*** (.104)	-.951*** (.107)
Regional FE	No	No	Yes
Observations	30,252	30,252	29,323

**Notes:** \*\*\* denotes significance at .1%, \*\* at 1%, and \* at 5% level. The sample comprises all self assessment taxpayers who were subject to a random audit on a tax return for tax years 1998/99 to 2008/09. Observations are at the individual taxpayer level. The outcome measures whether the taxpayer was found to have underreported tax due for the tax year in which they were audited. Table shows coefficients from a logistic regression of the outcome on demographics, income sources, industry, and income quintile based on reported income level for the previous year. Income quintile is calculated based on the previous years' reported income, among those reporting positive income the previous year. The excluded group is those declaring no income in the previous year. The excluded industry category is agriculture. Column (2) includes a square term in age, to allow for non-linearity in this effect. Column (3) additionally allows for regional level fixed effects. Regions are London, England excluding London, Scotland, Wales, and Northern Ireland.