

***BEHAVIORAL RESPONSES TO TAXPAYER AUDITS:
EVIDENCE FROM RANDOM TAXPAYER INQUIRIES***

Norman Gemmell and Marisa Ratto

This paper argues that random audit programs provide income taxpayers with information that alters their perceptions of, and hence their behavioral responses to, audits. Comparing samples of randomly selected audited and non-audited UK taxpayers, the evidence confirms predictions that audited taxpayers found to be “compliant” reduce their subsequent compliance. The reverse response is observed for taxpayers found to be “non-compliant”. The results highlight the importance of testing separately for the responses of taxpayers facing different opportunities and incentives to evade tax to avoid conflating their different effects, and to reveal both positive and negative indirect revenue effects from random auditing.

Keywords: tax evasion, audit perceptions, behavioral responses

JEL codes: H26, H30, K42

Norman Gemmell, The Treasury, Wellington, New Zealand

(norman.gemmell@treasury.govt.nz)

Marisa Ratto, Université Paris-Dauphine, (LEDa-SDFi), Paris, France

(maria_luisa.ratto@dauphine.fr)

I. INTRODUCTION

This paper investigates whether UK taxpayers alter their compliance behavior in response to the process of random audits, using a sample of individual and business taxpayers. When taxpayers are audited, two effects on their declared tax liabilities might be expected. First, where non-compliance is detected, there is a direct effect in the form of additional tax yield, including any fines, for the year of the inquiry. Second, there are possible indirect effects in the form of changes to future compliance behavior. These behavioral responses take two forms: (i) changed behavior by audited taxpayers themselves (sometimes labeled a “corrective” or “preventive” effect); and (ii) spillovers to non-audited taxpayers, sometimes labeled a “deterrent” effect. This paper first focuses on the former “preventive effect” on audited taxpayers.¹

Despite an extensive literature on tax evasion in general, the literature on how taxpayers’ compliance behavior responds to auditing is more limited. In part this reflects researchers’ lack of access to confidential taxpayer unit record data. However, Slemrod, Blumenthal and Christian’s (2001) investigation of individual taxpayers’ responses to the *threat* of an audit, stimulated a number of subsequent empirical and experimental studies that seek to identify how taxpayers’ compliance behavior responds to various compliance regimes and social settings. Alm and McKee (2004, 2006), for example, use experimental methods to assess the relative merits of random versus risk-based audits, and to examine how individuals compliance decisions respond to announcements regarding audited selection and to the “productivity” of an audit.² Nevertheless, as Alm, Jackson and McKee (2009, p.392-3) note, it remains the case that: “the ways by which taxpayers become aware about audit intensity and then respond to their assessments of enforcement efforts are ‘open questions’ (Plumley

¹ Alm, Jackson and McKee (2009) prefer the term “direct deterrent effect” to refer to effects on those who are audited, and “indirect deterrent effect” to refer to spillover effects on the non-audited.

² Their findings suggest that the announcement of audits increases the compliance rate of those who are told that they will be audited, but decreases the compliance rate of those who know, by elimination, that they will not. An

(1996)”. Alm Jackson and McKee (2009) address some of these “audit response” questions using experimental methods. Our objective here is to address similar questions but using actual, rather than experimental, taxpayer data. This also allows us to address responses to the audit process which have only recently begun to be considered in the experimental literature, such as the impact of differences in the “quality” of audit outcomes on taxpayer responses.

Focusing on the preventive effects in (i) above, we argue that the predicted responses by taxpayers faced with random auditing, are likely to depend both on their expected, endogenous probability of audit, and the expected or perceived “quality” of those audits (e.g. how much evasion they detect). Neither of these features has figured prominently in previous analyses. Based on unique access to UK taxpayer data, our results provide direct evidence on the behavioral responses to the *experience*, as opposed to the *threat*, of an audit. We argue that these responses are not uniquely signed in theory and can be expected to depend on the compliance/non-compliance verdict of previous audits and the tax authority’s previous success in identifying the amount of non-compliance. Our empirical evidence suggests that random auditing generates both positive and negative preventive effects that differ across taxpayers in ways consistent with our predictions. In addition, since predicted signs differ for different groups of audited taxpayers, we show that testing for a common response across all audited taxpayers is likely to conflate these positive and negative responses and hence increase the likelihood of incorrectly reaching a “no response” verdict.

The remainder of this paper is organized as follows. Section II briefly reviews the recent theoretical and empirical literature on responses to taxpayer auditing. Section III considers how the experience of an audit is expected to influence taxpayers’ future compliance, by changing their perceptions of being audited and their expectations regarding the amount of discovered tax evasion. This motivates our empirical analysis in Section IV

increase in the announced audit effectiveness has a positive effect on compliance only for taxpayers whose audit has been announced.

which describes the data used and our empirical methodology. Results are discussed in Section V. Section VI considers the relevance of these results for the choice between random and targeted audits, and some conclusions are drawn in Section VII.

In brief, our results demonstrate that the verdict of a previous audit – in particular whether a taxpayer is identified as “compliant” or “non-compliant”, and the amount of undeclared tax yield identified – affects whether subsequent compliance increases or decreases and the extent of any increased compliance. Distinguish empirically between so-called “compliant” and “non-compliant” audited taxpayers turns out to be crucial for correctly identifying their responses. In addition, responses seem to be larger for audited business taxpayers (who generally have greater opportunities to evade) than for personal taxpayers.

II. EXISTING LITERATURE ON BEHAVIORAL RESPONSES TO TAX AUDITS

In the standard portfolio model of tax evasion an individual chooses the optimal amount of income to evade in order to maximize expected utility. The tax rate and the fine rate are exogenously decided by the tax authority, and taken as given by the individual. The probability of detection is typically modelled as a fixed subjective probability, as perceived by the taxpayer³. It is assumed that the audit is 100% successful so that, if audited, an evading individual is caught and fined for the full amount of evasion.

As is well-known, in this model an increase in the audit probability unambiguously decreases the amount of evaded income as it increases the expected cost of being caught. The assumption of a fixed probability of detection is, however, rather strong. In the original model, Allingham and Sandmo (1972) also consider the case where the probability of detection varies with declared income and find the same result that an increase in the probability of detection decreases tax evasion.

³ See Sandmo (2005) for an overview.

However, as Alm, Jackson and McKee (2009) acknowledge, in reality individuals are likely to form expectations based on a variety of available information. In addition, audits are not always 100% successful in the sense that they may detect less than the full amount of evasion. This can be expected to affect audited taxpayers' perceptions of the "quality", as well as the likelihood, of future audits.

Recent theoretical contributions have modelled taxpayers' perceptions of being audited as different from the true probability of audit. Kim (2005) considers a model in which taxpayers have a private signal about the true probability of being audited. The author analyzes the role of audit misperceptions on compliance behavior and shows that an increase in the noise of the private signal about the audit has a positive effect on compliance behavior only when the true probability of detection falls below the cut-off equilibrium value which defines evaders and non-evaders. Hence if the true probability is very low, compliance will be higher the greater the imprecision of the taxpayers' private signal. In his model the probability of an audit corresponds to the probability of detection with the audit 100% successful, as in the standard model. Snow and Warren (2007) extend the standard portfolio model of tax evasion to a two-period framework, where the taxpayer has some initial expectations about being audited and updates these expectations on the basis of past audit experience, following a Bayesian updating rule. In their model, being audited in the past increases the perceived probability of being audited in the future. Also in this model, however, when an audit occurs evasion is fully detected and the updating rule is unaffected.

Turning to the empirical literature, the available evidence on the effects of taxpayer audits shows contrasting results on the magnitude of their tax compliance impact. We can distinguish two main approaches: studies that consider the effects of experiencing an audit and studies that consider the effects of an increased probability, or threat, of an audit. Among the former, early work by Long and Schwartz (1987) and Erard (1992) using Tax Compliance Measurement Program (TCMP) data, found little effect of audits on future behavior. On the

other hand, Dubin, Graetz and Wilde (1990), Tauchen, Witte and Beron (1993), Plumley (1996) and Dubin (2007), using time series cross-sectional information from individual tax returns aggregated at state level, found that audits had a large positive compliance effect with the indirect effects of an audit outweighing the direct effect⁴.

These recent findings based on individual data find some support from experimental evidence. For example, some laboratory studies in the 1980s⁵ showed that subjects being audited in earlier rounds of experiments improved their compliance in later rounds. One interpretation of this is that being audited one year raises taxpayers' perceptions of the chances of being audited subsequently. However, more recent experimental studies by Mittone (1999, 2006), Maciejovsky, Kirchler and Schwarzenberger (2007), and Kirchler (2007), specifically designed to examine the dynamics of compliance after an increase in the audit rate, suggest a sharp *decrease* in compliance immediately after an audit, followed by an improvement after a number of rounds. They interpret this as evidence that a recently audited taxpayer believes the likelihood of a subsequent audit is small (even when participants know that audits are truly random). After a few rounds however, the presumed likelihood of an audit again increases, raising compliance.

The Slemrod, Blumenthal and Christian (2001) study of audit threat is one of only a few to examine actual taxpayer responses, using tax returns from the Minnesota Department of Revenue. A group of randomly selected taxpayers (the "treatment group") was informed by letter that their up-coming tax returns would be "closely examined". The effect of receiving the letter was estimated using a difference-in-difference approach: the change in declared income before and after the letter was sent was compared between the treatment group and a control group of randomly selected taxpayers who did not receive the letter. Results suggested that the "audit threat" effect varied depending on the level of income and on

⁴ For example, according to Plumley (1996), who uses a rich dataset across US states for the years 1982-1991, the indirect effects of audits produce eleven out of twelve dollars of additional revenue.

⁵ For example Spicer and Hero (1985), Benjamini and Maital (1985), Webley (1987).

opportunities to evade. Among the low and middle income taxpayers, the audit notice had a very large, positive impact on high opportunity taxpayers (12% increase in reported tax for middle-income, high-opportunity taxpayers, and 145% for low-income, high-opportunity). However, tax liability for high-income taxpayers appeared to fall significantly compared to the control group. A plausible explanation provided by the authors was that high-income taxpayers, having more complicated tax affairs and being more likely to have professional assistance with their tax matters, may have interpreted the audit threat information differently to other taxpayers. In particular, they suggest that such taxpayers may have believed that the ultimate outcome of an audit is more manipulable than for other taxpayers and depends more on their initial report. More precisely, expected income upon audit may not be a monotonically increasing function of declared income, but it may reach a maximum at some level of declared income which is less than their actual income. Hence, facing a certain audit, they would optimally report less income than before in order to maximize expected income in the audit state.

These various results challenge the simple view that a tax audit, or increased audit threat, necessarily raises a taxpayer's compliance in future. Individuals form perceptions not only about the probability of being audited but also, conditional on being audited, about the probability of being caught evading and the extent of detected evasion. If an audit does not fully identify evaded tax, then perceptions of being caught in future could be lowered, potentially reducing compliance. Slemrod, Blumenthal and Christian (2001) show that audit threat responses differ across types of taxpayers and can affect taxpayers' perceptions about being audited in different ways, depending on their income and opportunities to evade. In our analysis below, where we examine observed responses to an *actual* audit, we similarly argue that distinguishing between alternative taxpayer subgroups is important, but where these subgroups differ in the official information revealed to them from previous audits.

This is in line with Alm, Jackson and McKee (2009), who use experimental approaches to examine how “official” information from the tax authority and “unofficial” information from other taxpayers affects compliance choices. In particular they show that responses to post-audit information are conditioned by how well-informed taxpayers are (regarding audit rates) prior to auditing – a result that is broadly consistent with our evidence below.

III. HOW THE AUDIT EXPERIENCE INFLUENCES FUTURE COMPLIANCE: OUR APPROACH

In this paper we evaluate the impact of an audit on future compliance, by treating taxpayers’ behavior as a function of their perceptions of being audited and the amount of tax evasion they expect to be uncovered. Since taxpayers are not normally informed of the audit rule adopted by the tax authority, their ex-ante perceived probability of audit may not correspond to the actual probability of being audited. We also allow for the case where the tax authority may not (fully) discover the true amount of evasion, which is private information of the taxpayer. Crucially, taxpayers’ expectations of being audited and of the amount of discovered evasion are unlikely to be fixed, but rather change in response to the audit experience, as we clarify below. Our approach, then, differs both from the standard portfolio model, where the subjective probability of audit is fixed and the audit is 100% successful, and from the more recent literature, which does not allow partial audit success. As we show, our predictions are that being audited in the past does not necessarily increase future compliance.

In our approach the perceived probability of being identified as non-compliant is the product of the perceived probability of being audited and the perceived conditional probability that evasion is detected by the auditor, once an audit takes place. We assume that individual taxpayers adapt their beliefs about these probabilities and detection rates in the light of experience. In the present context, given data constraints, and in line with the findings

of Lochner (2007), we assume that the only information used to update perceptions arises from the individual's own audit experience.⁶

We begin by clarifying the updating process of taxpayers' perceptions after receiving an audit. The result of an audit will affect both the expected probability of a future audit and the expected probability that all future evasion is detected. In particular, the expected probability of being audited in the future increases if tax evasion has been discovered (even if only partially) by a previous audit. This is because those taxpayers are more likely to be added to a risk-based audit pool⁷. Alternatively, if no evasion is discovered, two cases arise: either there was indeed no evasion, or the audit completely failed to uncover it. In either case, if the individual thinks that future random audits are conducted "with replacement" (that is, where the probability of being *randomly* selected in current or future periods is independent of past selection for audit), then there will be no change in the expectation of being audited in future.

However, if the individual thinks that future audits are conducted "without replacement" (such that being randomly selected for audit last period reduces the probability, possibly to zero, of being randomly selected subsequently), or the audit completely failed to discover tax evasion, then the expected probability of being audited in the future will decrease. That is, zero detected evasion at time $t-1$ reduces the probability of being selected "randomly", or via a risk-based audit program, at time t .

For perceptions of audit success, when an audit discovers less than an individual expects (based on their private information), the perceived probability of success of the audit

⁶ In our empirical analysis below of randomly selected "audited" and "control" groups this is analogous to assuming that any other relevant characteristics are equally reflected in both random groups. Empirical studies have attempted to shed some light on the factors affecting offenders' beliefs of being caught. Lochner (2007), for example, examines which factors influence individuals' perceptions about the probability of arrest among young offenders and shows that while beliefs are largely unresponsive to most outside influences, such as demographic characteristics, neighborhood conditions, local official arrest rate, they do respond to an individual's own experience with crime and the police.

⁷ At least in the UK, evidence of non-compliance from random audits can provide a trigger for possible further monitoring by the risk-based audit program. In addition, information on non-compliance obtained from the results of random audits may be used to inform the selection rules of risk-based auditing. As a consequence a taxpayer who is found to be non compliant is more likely to receive a targeted audit in the future.

in the next period will decrease. The converse holds when an audit discovers more than an individual expects, for example where a taxpayer was confident, prior to audit, that a particular form of evaded tax had a low probability of discovery. Alternatively the taxpayer may believe, prior to an audit, that once selected for audit $x\%$ of evasion will be discovered. If this percentage is confirmed by the audit, the perceived probability of audit success remains unchanged.

The matrix below summarizes the possible effects of random audits on the various components. The first column represents the known outcome of a random audit at time t which classifies taxpayers into two categories: compliant (zero evasion) and non-compliant (positive evasion). Consider the first row of the matrix. If the taxpayer was truly compliant or was non-compliant but believes that the next period's audit sample is selected "with replacement", expectations about the probability of being audited in the next period will remain constant. If, instead, the taxpayer was actually non-compliant but escaped detection and believes random audits are with non-replacement, expectations of an audit at $t+1$ will decrease. The expected detection rate (column 3) remains constant if the individual was truly compliant, but decreases if the individual was actually non-compliant.

Figure 1
Effects of Random Audits on Expected Audit Probabilities and Detection Rates

Category (from random audit)	Expected audit probability	Expected detection rate	Compliance:
Compliant	FALL (non-replacement) CONSTANT (replacement)	FALL or CONSTANT	FALL or CONSTANT
Non-compliant	RISE (targeted in future) CONSTANT (replacement)	RISE or FALL	RISE or (FALL?)

The second row of the table shows taxpayers classified by audit at t as non-compliant. These might expect to be specifically targeted in the future, such that expectations of a $t+1$ audit increase. However, if only random audits are conducted (and with replacement), an unchanged probability of audit in the next period might be expected. The expected rate of detection rises or falls depending on whether the proportion of evaded income discovered by the audit is greater or less than the individual expected prior to audit at t .

The overall effect on future tax evasion can be seen in column 4, and depends on the combination of the effects on the expected audit probability and on the expected detection rate. For those previously classified as “compliant”, future compliance is expected to fall or remain constant. For those classified as “non-compliant” compliance is expected to rise, unless the audit identified less tax evasion than the taxpayer expected, and this is sufficient to counteract any increase in the future audit probability via risk-based audits.

In our empirical analysis we analyze the effects of random audits on different groups of taxpayers, using individual level data from UK taxpayers’ income tax returns. With limited taxpayer-specific information on the “other factors” likely to affect taxpayers’ perceptions, we consider a simplified model in which only past experience with the audit process affects individuals’ expectations. Note that we cannot distinguish whether the individual believes that future random audit samples are selected with or without replacement, as we do not have this information. Hence, if an individual receives a random audit, changes in their perceptions depend only on whether the audit rule might change – for example if the individual is caught evading and therefore expects to be specifically targeted via a risk-based selection procedure in the future.

IV. THE UK TAXPAYER DATA & EMPIRICAL METHODOLOGIES

In seeking to identify the preventive effects discussed in the previous section our analysis needs to distinguish between those taxpayers labelled as “compliant” and “non-

compliant” by the audit process and identify how this information changes different taxpayers’ expectations of future detection (via both the likelihood and detection rate of future audits). We have data on the direct yield (additional tax raised plus penalties) from random audit enquiries in the year 2000, separated into compliant (zero direct yield) and non-compliant (positive direct yield) – on completion of the audit process. Clearly this only captures “true” (non)-compliance to the extent that the audit succeeds. We should note that this distinction between compliant and non compliant as determined by the audit process may not be very accurate, in that taxpayers with positive but low yield values might be considered “compliant”, for example because the low yield reflects genuine compliance mistakes rather than intended evasion. With this view, in Section V.D, we consider whether the behavioral response to a random audit of taxpayers with larger yield differs from those with a lower yield. Inevitably, no similar information is available for the non-audited.

We also cannot observe directly how taxpayers’ subjective probabilities of audit change in subsequent years. However, from our previous analysis of the expected audit probability and detection rate for those labelled as “compliant” and “non-compliant” by the 2000 audit, we can predict the direction of change in indirect yield as described above. In addition since some taxpayers have to submit their subsequent years’ tax returns after knowing their 2000 return has been randomly selected but before knowing the audit outcome, we may expect to observe different responses related to these differences in information available to the taxpayer.

A. Sample Selection

UK non-corporate taxpayers are liable for personal income tax (PIT) and National Insurance Contributions (NICs) – a social security “tax”. For “standard” employees not subject to higher marginal rates of PIT, both PIT and NICs are paid via the pay-as-you-earn

(PAYE) scheme and such taxpayers are not required to complete a personal tax return.⁸ Those that do complete a tax return are referred to as Self-Assessment (SA) taxpayers. Within this set of SA taxpayers, the UK revenue authority, HM Revenue & Customs (HMRC) undertakes random audits on a number of different “segments”. Three segments are examined here: Medium (M) and Small (S) Businesses, and Personal (i.e. individual) taxpayers (P).⁹ By definition, the business segments are not incorporated (otherwise they would be taxed instead via the corporate income tax). In practice, though the M- and S-segments may be thought to be similar, many small businesses (those with a turnover of less than £15,000) are akin to personal taxpayers with some self-employment income. They are therefore often more like the self-employed within the P-segment, but with a higher proportion of total income coming from self-employment.

For the present analysis, we identified all those taxpayers within each segment who had been audited in 2000 via the random selection process. The preventive effects of an audit may take some time to feed through, and may depend on when the audit was completed. We therefore chose initially to compare declared tax in the three years after 2000 with the three years before. To minimize the likelihood that previous or future audits influence observed reactions to the 2000 audit, we excluded all taxpayers who were audited in these six years (1997-99, 2001-03).¹⁰ We then randomly selected a sample of “control” taxpayers who were not audited in 2000 but otherwise matched the selection criteria of the audited taxpayers.¹¹

⁸ Of course, all taxpayers, even if they are taxed at source via PAYE, are expected to complete a tax return if they believe they are likely to be liable for additional tax.

⁹ Segment definitions are as follows. *Personal* (P): Individuals/Non-Business/Non-Complex). Self-assessed taxpayers, but with no income from trade, land, property or partnership. Individuals have no directorships and total non-business income less than £100,000. *Small Businesses* (S): Self-assessed taxpayers with income from trade and land property but no recorded partnerships. Total annual turnover from all trades and property must be under £15,000. *Medium Businesses* (M): As for small businesses, but with total annual turnover in the range £15,000 to £250,000.

¹⁰ A taxpayer audited again *after* 2000, may respond differently to the 2000 audit because audits can take a number of years to complete and some taxpayers’ responses may be conditioned by audit outcomes, as discussed below.

¹¹ For example, controls were selected from identical taxpayer segments and, like the audited, were deselected if they had been audited in 1997-99 or 2001-03.

This yields a balanced panel where, for each household, we observe all six years of data¹² for declared tax, total income its components (e.g. self-employment income, employment income).¹³ Two further distinctions within each taxpayer segment are relevant for our subsequent investigations: whether audited taxpayers were (i) identified as “compliant” or “non-compliant”; and (ii) whether their tax returns for years after 2000 were submitted before or after the outcome of the 2000 audit was known.¹⁴

On (i), the “compliant” are defined as those audited taxpayers for whom the 2000 audit led to no additional direct yield; the “non-compliant” are those for whom positive additional yield was deemed liable (additional PIT or NICs, plus any interest/penalties) as a result of the audit. On (ii) our data identifies whether audited taxpayers submitted their 2001, 2002 or 2003 returns before or after the end of the 2000 inquiry. Table 1 shows the numbers of taxpayers in each sample (segment), and distinguishes the “compliant” from the “non-compliant”.

Each segment consists of around 2,500 – 3,000 individuals of which around 20-25% were audited. Within the audited, the ratios of compliant to non-compliant are quite different however. For the S and P segments, there are relatively large numbers of compliant taxpayers (with only 136 (= 18%) non-compliant P-segment taxpayers out of a total of 776 audited, for example). The audited M-segment, however, reveals 55% (320/575) non-compliant.

Table 2 reports the year in which the 2000 audits were completed. This shows that only 42% of enquires had been completed by the time audited M-taxpayers submitted their 2001

¹² We eliminated from the original data set those taxpayers who left SA over the period 1997-2003. This raised some concerns about attrition issues. However, those taxpayers who left SA did not seem to have left due to the treatment. For example, for the M segment, we initially had a total of 3715 taxpayers, of which 80% were in the control group and 20% in the treatment group. After eliminating those who left SA over the period, we were left with 2770 observations, of which 79.06% were in the control group and 20.94 in the treatment group. Hence, despite having lost 25% of the original sample, the relative size of the control and treatment groups stayed the same, suggesting that leaving SA was not related to being subject to an audit.

¹³ The full set of income components includes: self-employment income, employment income, partnership income, UK land and property income, UK savings, UK pension income, income from investment in shares, foreign income, alimony, other income, total allowable deductions. We also have data on age, gender and trade description, but due to many missing values, these latter variables were not used.

tax returns. By 2003 this had risen to 90%, with the remainder completed within the next two years. This pattern is noticeably different from the P-taxpayers, where almost 80% submitted their 2001 returns after knowing the outcome of the audit. The S-taxpayers reveal an intermediate case with 65% in 2001.

The three segments also reveal differences in the time profile of their direct yield. Table 2 shows (an index of) the average amount of additional tax obtained from the audits that were completed in each year, by segment. To respect the need for confidentiality of sensitive information we do not report the amount of direct yield obtained per non-compliant taxpayer but instead report an annual index for each segment where 2001 = 100. The table shows that, whereas for small and medium businesses, on average the longer running audits tended to be associated with higher direct yield (e.g. more than four times the average yield per (non-compliant) audit in 2004-05 compared to 2001), the reverse is true for the P-segment. This suggests that it takes longer to identify the extent of tax evasion in businesses compared to individuals, with the “easy” (low evasion) cases generally completed first in the case of businesses.¹⁵

Finally, if both audited and controls are selected truly randomly, any observed differences between these two groups prior to audit should be due to random factors only. For example, on average, income should be the same in both groups. In Sub-section *B* we report tests for possible non-random differences between groups. This issue is also relevant to our use of difference-in-difference (D-i-D) methods in Section 4 to separate out audit impacts. Though our combined audited group was selected randomly, the compliant and non-compliant sub-groups were not, the split being determined by the various aspects of the audit process. These sub-groups could therefore differ from the controls, even before the 2000 audits for systematic reasons. Below we discuss how we deal with this.

¹⁴ We do not consider the 2000 return since data for 2000 includes any additional tax mandated by the audit and fines imposed for under-declaration. These represent *direct* effects of the investigations.

B. Empirical Methodologies

This sub-section describes the application of D-i-D methods to audited and control group data to measure the impact of auditing. It then considers how the approach is affected by separating the audited into compliant and non-compliant sub-groups. First, we test whether the declared tax (including NICs) of audited and control groups can be considered to differ for random reasons only, based on a t -test and χ^2 -test for group mean and median differences respectively and the Wilcoxon rank-sum test. For all three segments these suggest that, using a 10% confidence interval, the inter-group differences prior to 2000 are not statistically significant confirming that, to a reasonable level of confidence, the two group means *can* be considered the same prior to 2000.¹⁶

Test results for the M-segment (though not for the others) provide some evidence of significantly higher tax paid by the audited after 2000 – even though our hypotheses propose that the compliant and non-compliant audited sub-groups could deviate in either direction from the controls as a result of auditing. This evidence is *suggestive* of a possible preventive effect of auditing. However, the key tests are whether, *ceteris paribus*, the “after versus before” difference for the audit group exceeds that for the control group and whether the impact of auditing differs as hypothesized between the two compliant and non-compliant sub-groups.

To examine this we use a differences-in-differences regression, applied to the two 3-year periods before and after 2000 (1997-99 and 2001-03). This takes the form:

$$[1] \quad y_{it} = \beta_0 + \beta_1 D_T + \beta_2 D_G + \alpha D_G D_T + \varepsilon_{it},$$

¹⁵ Note that, though we use the term “evasion” here, this includes some very small amounts of direct yield that may result from unintended errors on tax returns. These are discussed further below.

¹⁶ Results are available from the authors.

where i = taxpayer; t = time; y = declared tax (including NICs); D_T = dummy for time (= 0 before 2000; = 1 after); D_G = dummy for an individual being in a group (1 = audited, 0 = control); and ε_{it} = random error term.

Interpretation of the coefficients is as follows: β_0 = expected average outcome for the controls before 2000; β_1 = difference in expected average outcome before/after 2000; β_2 = difference in expected average outcome between audited and controls *before the audit*; and α = difference in the change in the outcome for the audited group compared to the control group (the D-i-D parameter measuring the effect of the intervention. That is, $\hat{\alpha} = (\bar{y}_{1,1} - \bar{y}_{1,0}) - (\bar{y}_{0,1} - \bar{y}_{0,0})$, where \bar{y} is average declared tax, and where the first and second subscripts refer to D_T and D_G respectively).

Equation (1) can readily be re-written to capture the impact on tax declared for each post-audit year separately. We use OLS regressions for this first part of the analysis and cluster at the taxpayer level. Ideally we would control for personal characteristics. As already mentioned, our dataset includes age, gender and trade description; however, given that there are substantial missing values for those characteristics, to avoid losing large numbers of observations we do not control for age, gender and occupation¹⁷. We also decided not to include the components of total income separately in the regressions since these come from the same declaration. Income component data show that the shares of the main components of total income vary across the three sample segments and over the six years. These shares are also potentially affected by auditing and possibly in different way for the “compliant” and “non-complaint”¹⁸.

¹⁷ The most problematic variable was *trade description*, for which we had 2754 missing values (out of 16464 observations) for the M-segment, 7440 missing values (out of 15942 observations) for the S-segment and 17304 missing values (out of 18150 observations) for the P-segment.

¹⁸ For example, some types of income are likely to be seen as easier to hide or, once non-compliance is identified, more likely to be checked by HMRC in future (e.g. via being added to the risk-based pool). Hence, the audited might be expected to change the amounts they declare. Further, if an audit reveals that, for example, self-employment income in particular is evaded, future reporting of this component could change.

To examine whether the impact of random audits is different for compliant and non-compliant audited we use a similar regression to (1), but allowing for two audited groups. In this case the specification becomes:

$$[2] \quad y_{it} = \beta_0 + \beta_1 D_T + \alpha_1 D_C D_T + \alpha_2 D_{NC} D_T + \gamma_i + \varepsilon_{it},$$

where now there are three groups (controls, compliant audited, C, and non-compliant audited, NC) and D_C (D_{NC}) is a dummy for being in the compliant audited (non-compliant audited) group. In this case α_1 (α_2) measures the difference in the change in the mean declared tax between compliant audited (non-compliant audited) and *all* non-audited controls.

In the case of (1), where both groups (audited and controls) are selected randomly, we do not need to allow separately for differences in unobservable characteristics across groups; e.g. via individual fixed effects – captured by γ_i in (2). Random selection should ensure there are no systematic group differences. In (2), however, we include γ_i since the two audited sub-groups are not randomly selected, but rather are selected on the basis of known taxpayer characteristics (their compliance) identified by the audit. Within the control group we cannot distinguish the compliant from the non-compliant. However we can allow for the impact on each audited sub-group (compliant versus non-compliant) to differ compared to all non-audited. We therefore adopt the (individual)¹⁹ fixed effects (FE) estimator to allow for the possibility of any systematic differences between the three groups prior to audit. This specification allows the compliant audited and non-compliant audited to differ on observable and unobservable characteristics that are constant over time. It may be, for example, that the differences between a truly compliant and a non-compliant taxpayer are related to taxpayer attitudes or use of a tax advisor. Effectively our FE estimator treats any such differences impacting on compliance levels as constant over the time period we consider. This

¹⁹ We need to allow for the fact that individuals will have correlated errors with themselves, as we have a panel dataset, and hence consider individual fixed effects. If we considered group fixed effects, we would obtain biased standard errors.

assumption would be violated if trends in the control and audited groups changed *in different ways* after 2000 for *reasons unconnected with auditing*.²⁰

We also argued in Section III that the closure of the inquiry would give taxpayers information that might affect their expectations of future audit or the extent of future non-compliance detection. We therefore also examine whether there is evidence of more robust preventive effects of auditing when each taxpayer's "post-audit" declared tax is measured only from the tax returns submitted after inquiry closure.²¹

Finally, we consider whether *the amount* of discovered tax evasion affects the behavioral change. In order to test whether taxpayers charged with larger direct yield (arising from their discovered non-compliance) have a stronger response to the audit, we include an interaction term in (2) between the dummy identifying the non-compliant after the audit ($D_{NC}D_T$) and the amount of their direct yield.

V. DIFFERENCE-IN-DIFFERENCE RESULTS

This section first considers whether the randomly audited taxpayers as a whole differ from the random control group (Sub-section A). We then examine how far outcomes differ when the audited are split into compliant and non-compliant, and when preventive effects are estimated using only "after audit closure" data (Sub-section B). Sub-section C calculates the implied revenue gains/losses associated with the preventive effects identified. Finally, in Sub-section D we examine how the amount of discovered tax evasion affects the behavioral responses of audited non-compliant taxpayers.

A. Audited versus Controls

²⁰ In an appendix to an earlier version of this paper (available from the authors) we explore some of these issues. Since this did not identify any non-audit sources of difference, we regard the fixed effects model as reliably identifying audit effects.

²¹ Taxpayers are likely to acquire differing amounts of information during the audit process regarding the final audit outcome, and they may adjust their expectations accordingly before closure. It is therefore unclear how far audit "closure" provides additional information.

Our hypotheses above suggested that, within the audited group, differently signed responses to audit might be expected depending on a number of characteristics, such that the net response across all audited taxpayers is ambiguous. We therefore began by using (1) to test whether the audited and control groups display a significant D-i-D parameter, comparing the two periods 1997-99 and 2001-03 (top part of Table 3). For each of the three segments this produced values of α in (1) that are positive for the M-segment and negative for the S- and P-segments, though all three are insignificantly different from zero.²² That is, changes in the audited group (as a whole) after audit were not statistically larger or smaller than those observed for the control group. This is perhaps not surprising, given the different sub-group responses expected by the audited. However, together with the evidence on different sub-groups that we present below, it serves to highlight the importance of testing audited sub-groups separately to avoid conflating, and hence potentially failing to identify, these different responses.

B. Separating the Compliant and Non-compliant Audited

Results when the audited are split into “compliant” and “non-compliant” components are shown in Table 3. The first set of D-i-D parameters are obtained when comparing average declared tax in 2001-03 with 1997-99; the second set compare declared tax in the first year after inquiry closure with the average over 1997-99.²³ Because many enquiries last for several years we have insufficient data to consider more than one year after closure.

In principle, the parameters from regressions of the form in (1) and (2) represent average additional declared tax in £s per year. However, HMRC are understandably sensitive

²² For example, for the M-segment, we find positive α s for 2001-03 ($\alpha = 128$) and for each year individually when they are included separately ($\alpha_{01} = 42$, $\alpha_{02} = 77$, $\alpha_{03} = 261$). However, t -ratios are 1.13 (for 2001-03); and 0.33, 0.57, 1.62 (for 2001 to 2003 respectively).

²³ Where the data indicate that a taxpayer’s audit closed during, say, 2001 but he/she submitted their 2001 tax return prior to that date, their 2002 tax return is treated as the first “after closure” year. For the period before the inquiry we use the mean of declared tax (including NICs) over the period 1997-1999. After 2000, for those who were audited we consider the first tax return following the inquiry closure. For example, in the M-segment, 42% of the 2001 returns were submitted after the inquiry closed, so for those cases the “year after” is 2001. Similar

to public release of information on the amount of direct tax yield from individual audits. Hence, to preserve confidentiality around yield data, those shown in Table 3 have been re-scaled by a common, undisclosed factor. The *t*-ratios shown in the table are those obtained for the original parameters (hence standard errors cannot be inferred from Table 3). The relative size of parameters in Table 3 therefore indicates the relative magnitude of indirect effects by the various taxpayer categories; *t*-ratios indicate the statistical significance of the original parameters from zero.

The results in Table 3 are revealing. In all cases positive parameters are obtained for the non-compliant and negative parameters for the compliant.²⁴ Considering first the “before/after 2000” regression, for the P and S-segments the negative parameters are larger and statistically significant. Unsurprisingly perhaps, the negative preventive effects for the compliant are less robustly estimated for medium businesses where they are a smaller fraction of the audited. The point estimates suggest a behavioral response amounting to a tax loss of around 140-190 units per compliant taxpayer per year. The reverse is the case for the positive preventive effects for the non-compliant where these are largest (342 units per year) and most robustly estimated for medium businesses.

Results are broadly similar for the “post inquiry closure” regressions with 4 of the 6 parameters now statistically significant including a larger negative (and significant) effect by compliant medium businesses. This is also now similar in magnitude to the positive effect of the non-compliant, amounting to around 18% of declared tax by these groups on average. This larger parameter for the complaint may reflect a greater capacity on the part of medium businesses on average (compared to small businesses or individuals) to change their

numbers for 2002 and 2003 are 41% and 7% respectively. For the control, we construct a weighted average of their 2001-2003 returns using as weights the same proportion of returns as for the audited (42%, 41% and 7%).

²⁴ We dropped 5 observations for the S-segment (3 from the audited compliant group, 1 from the audited non-compliant and 1 from the control group), which experienced a drastic decrease in declared tax in 2002 and 2003. Close examination of the data revealed that these cases experienced a very large increase in employment income in 2001 (above £100,000) which was followed by a very sharp decrease (greater than £80,000 a year) in the next

compliance behavior once an audit is complete. They are also likely to have greater confidence that the compliance label means that another near-future audit is unlikely.²⁵

With this exception, the results in Table 3 do not indicate that the magnitudes of preventive effects are necessarily larger on average once inquiries are closed. This may partly reflect the dominance in the sample of the many small business/personal audits whose audits closed quickly in 2001 or 2002 (Table 2). Hence, for those taxpayers our first post-2000 observation is also often the first “post inquiry closure” observation.²⁶

C. How Big Are Preventive Effects?

The implied “preventive yield” from the random audits can be calculated based on the parameters in Table 3. We use the “one year after closure” results (since these are the most robustly estimated), together with data on the numbers of taxpayers in each category. These are shown in Table 4 as a percent of each category’s declared tax, on average, in 2000. The table also shows the preventive yield obtained using parameters plus/minus one standard error from the mean point estimate.²⁷

This shows, for example, that for Personal taxpayers identified as compliant, declared tax fell, due to audit, by about 7% on average over 2001-03. For the non-compliant Personal, the estimates are less precise (because parameters are less well identified), but these predict on average a similarly sized increased preventive yield. For medium and small businesses, the positive preventive effects are much larger, at around 18-24% of declared tax, while the

two years. These changes were due to either a change in employment in 2001 or to bonuses received in 2001, neither of which seem likely to be audit-related.

²⁵ In view of the relatively large compliance costs imposed on audited businesses, HMRC is generally reluctant to audit such businesses again, when previously found to be compliant, should they be randomly selected again soon thereafter.

²⁶ We noted earlier that the expected probability of a future audit might be expected to decline immediately following an audit, and perhaps rise again subsequently if the experimental results of Maciejovsky, Kirchler and Schwarzenberger (2007) apply here. However, we are limited in our ability to examine this issue using the 2000 audit data since we only have complete data on our samples to 2003 and many audits, especially for medium business, were not completed till then.

²⁷ Publicly available data on total amounts of direct tax yield from audits are given in NAO (1999; p.R59) for the 1998-99 tax year. They show that self assessment enquiries yielded £81 million from around 248,000 non-business taxpayers (average = £327) and £98 million from around 121,000 business taxpayers (average = £808).

negative preventive effects of the compliant are also large on average for medium businesses (-17%), and for small businesses (-11%, but with a wider margin of error).

D. Controlling for the Amount of Discovered Tax Evasion

So far we have defined a taxpayer as “compliant” if the 2000 audit led to zero yield. Clearly however a small amount of positive yield could indicate mistakes (that both the taxpayer and the revenue authority recognize as such), rather than deliberate evasion. In this case, such low-yield taxpayers may be regarded as, and behave like, zero-yield “compliant” taxpayers.

Hence some *positive* amount of tax evasion disclosed by the audit may provide a better threshold by which to distinguish the behavioral responses of “compliant” and “non-compliant” taxpayers. Using the data on the amount of direct yield from the year 2000 audits, we test (a) whether taxpayers with a “small positive yield” (defined below) behave similarly to zero-yield taxpayers; and (b) whether in general larger yield values (perhaps above some threshold) signal a greater degree of non-compliance and hence a greater behavioral response. We first introduce in (2) an interaction between the dummy identifying the non-compliant taxpayers (positive yield) after the audit with the amount of total yield recovered by the tax authority from these taxpayers.

We examine this for the case where taxpayers know the outcome of the inquiry (when their yield is also known). To explore possible “threshold” levels of direct yield we first examine the quintiles of yield distribution and then more arbitrary ranges of direct yield based on £ values. Results are shown in Table 5 for the M segment²⁸. Results in the top half of the table use the quintiles of the direct yield distribution. These show that the impact of an audit is negative on those taxpayers with an amount of direct yield in the first quintile of the distribution, although results are not statistically significant, and a stronger and positive effect

²⁸ Results for the S and P segments were not statistically significant.

on the (more obviously) non-compliant whose direct yield puts them in the 4th and 5th quintiles of the distribution.

Having reviewed yield data, we also examined a number of more arbitrary threshold yield levels. The lower half of the table reports results for yield (£) ranges: 1-150, 151-350, 351-1000, 1001-3000 and >3000 (the omitted range being from 1-150). Results indicate that for those non-compliant taxpayers charged with a direct yield below £150, declared tax in the year following the inquiry closure decreased, similarly to the compliant taxpayers, although the result is not statistically significant. But for those taxpayers charged with a direct yield above £1000, declared tax increased, with a larger effect for those in the range above £3000. Hence, our point (a) above, on whether a small positive yield might induce a similar behavior to other zero-yield taxpayers, receives some limited support (but with results that are not statistically robust). On point (b) above - whether higher yield generates a greater positive compliance reaction - there is some support for the M segment.

Similar estimates for the S and P segments were not statistically significant. This may not be surprising once we look at the distribution of direct yield by quintile for the three (M, P, S) segments in Table 6. This shows an index of direct yield with the lowest quintile (Q1) for the P-segment set equal to 1. (The level of yield for this group was sufficiently small that it likely reflects mistakes rather than deliberate evasion). It is clear that the M-segment demonstrates considerably greater detected evasion (yield) across each quintile and in total. Moreover, especially for the S- and P-segments, there is little yield except from the 5th quintile, and perhaps the 4th quintile for the S-segment. So it is perhaps not surprising that there is little evidence of statistically robust yield effects for the S- and P-segments – especially since the numbers of taxpayers in each quintile for those segments are also lower.

In conclusion, we find positive changes in declared tax for those taxpayers who have been discovered under-declaring tax, and negative changes for those found to be compliant over the three segments. We interpret these changes as changes in compliance. In an

Appendix available from the authors, we look for any changes in trends in total income and the shares of self-employment, and employment, income components – the most important sources of income for the M- and S-segments. We find very similar patterns across the three segments and across control, compliant audited and non-compliant audited. Hence we conclude that the differences we find in declared tax are not explained by differences in income trends.

VI. SHOULD THE UK’S RANDOM AUDITS BE REPLACED BY MORE RISK-BASED AUDITS?

The evidence above suggests there may be significant negative preventive effects associated with taxpayers identified by random audits as “compliant”, especially for individual taxpayers where a large fraction of the randomly selected are found to be compliant. This might suggest that it would be preferable to substitute more risk-based audits for the random audit program. This would both target the “non-compliant” better, hence raising revenue per audited taxpayer, and avoid providing information to the so-called “compliant” audited which allows them to reduce their compliance. However, these arguments must be weighed against the potentially important role of indirect deterrent effects from the impact of auditing on *non-audited* taxpayers. Alm and McKee’s (2004) experimental evidence, for example, established that an enforcement agency could do better by using a combined risk-based and random audit program, for given compliance resources.

Information for our UK taxpayer sample does not allow us to model the indirect, deterrent effects of the random audit program. However, using publicly available data on the numbers of taxpayers in various categories sheds some light on the possible consequences of reallocating audit resources away from random audits in order to expand the risk-based program.

During the period covered by our empirical analysis in Section V, there were approximately 27 million income taxpayers, of whom around 9 million were self-assessed

(SA) taxpayers. Most of the remaining 18 million were pay-as-you-earn (PAYE) taxpayers, considered by HMRC to be sufficiently low risk that they are not obliged to complete a self-assessment tax return – they are expected to be compliant or engage in only minor non-compliance. Of the 9 million SA taxpayers, many would similarly be regarded as “low-risk”; they are required to complete an annual return simply because they are liable for the top marginal tax rate which is not applied at source to all income (e.g. bank deposit interest). As a benchmark, consider the case where half the SA sample (=4.5 million) form the pool of taxpayers who would be targeted by HMRC’s risk-based audit program, as potentially “high-risk”. This probably exaggerates the numbers of taxpayers with a high risk of non-compliance since our random samples considered above reveal only around one-third as underpaying tax and for many of those the level of their tax underpayment was small (Table 6).

From the UK National Audit Office (NAO, 2000), it is known that around 369,000 SA tax returns were audited via the risk-based program in 1999, with an additional 6,000 selected via the random program. Our sample evidence suggests that perhaps around 2000 (one-third) were non-compliant and may therefore be targeted by risk-based audits. Thus the replacement of random audits with more risk-audits could be expected to lead to only a small increase in risk-selected taxpayers: from 369,000 to 371,000 or, at most, 375,000.

To illustrate the magnitudes involved, these data imply that, for every 10,000 UK taxpayers, there were less than 2 “low-risk”, and less than 1 “high-risk”, who were audited randomly. This compares with 137 (per 10,000) audited taxpayers via the risk-based program, and a remaining 9860 taxpayers who were not audited at all. Hence removing random audits merely allows an additional 2 “high-risk” taxpayers (now 139) per 10,000 taxpayers, to be audited, for given compliance resource. Hence any *direct* revenue yield is likely to be small.

Additional *indirect* revenue effects potentially arise from two sources: greater compliance from those feeling a threat of audit from the risk-based program, and less compliance from those who consider themselves unlikely targets of that program. For the

latter case, 22.5 million “low-risk” taxpayers would now face a zero probability of being audited. Though we cannot know the previous probability of audit perceived by those taxpayers, if based on publicly available information on the ratio of audited to non-audited taxpayers, this probability could be between 0.022% (6000/27 million) and 1.39% (375,000/27 million). However, a range of studies find that taxpayers often *over*-estimate the probability of audit; for example, Alm, McClelland and Schultz (1992), Snow and Warren (2007). With no random program it seems plausible that such “low-risk” taxpayers would now perceive a zero audit threat and engage in (greater) non-compliance.

On the other hand, a compensating increase in revenue may arise from the increased probability of risk-based audit for the 4.5 million “high-risk” taxpayers. However this increase is negligible - rising from 1.37% (369,000/27 million) to, at most, 1.39% (375,000/27 million). That is, a rise in probability of 0.02 percentage points for 4.5million “high-risk” taxpayers, compared to a fall of 0.02 percentage points or more (to 0%) for the remaining 22.5 million “low-risk taxpayers. While we cannot be sure how these opposing influences would net out, there is a strong possibility of reduced overall compliance and revenue. Many so-called low-risk taxpayers are clearly not *innately* compliant, but likely comply in response to incentives, degrees of risk-aversion etc. With a large proportion of such taxpayers in the UK system, it is plausible that moving to a system with effectively zero probability that minor evasion will be detected, could generate a substantial revenue response in aggregate.

VII. CONCLUSIONS

This paper has argued that a random audit process provides income taxpayers with information that can be expected to alter their perceptions, and hence their behavioral responses to audits. In particular, taxpayers’ perceived probabilities of audit in future and the predicted extent of detected evasion can be expected to change. Three key determinants of these perceptions are likely to be: whether the verdict of a previous random audit was that the

taxpayer is compliant or non-compliant, whether random audits are known or perceived to be conducted with or without replacement, and the existence or otherwise of a targeted audit program.

Comparing samples of randomly selected audited and non-audited UK taxpayers, our evidence suggests, as predicted, that previously audited taxpayers who were found to be compliant, reduced their subsequent compliance. The reverse response was observed for taxpayers previously found to be non-compliant. For some taxpayers these responses were especially evident following the closure of the random inquiry; that is after a compliant/non-compliant verdict had been reached – and communicated to the taxpayer. These results serve to highlight the importance of testing for the responses of the so-called compliant and non-compliant sub-groups separately to avoid conflating their different responses.

We also found that behavioral responses (both positive and negative) by small and medium businesses were greater than those for personal taxpayers. This is consistent with the argument of Slemrod, Blumenthal and Christian (2001) that different opportunities to evade are associated with differing levels of compliance, with business taxpayers in the UK generally having greater opportunity to evade. However, whereas the Slemrod, Blumenthal and Christian (2001) experiment effectively increased the probability of audit for all participants, our actual audit evidence has been shown to either increase or reduce the perceived probability of detection, with consequent positive and negative responses respectively. For the medium businesses (for whom direct yield from investigations is generally higher), we found evidence of a greater positive compliance reaction from those taxpayers charged with high levels of direct yield.

The data available to assess the possible impact of replacing the UK's random audit program with more risk-based audits are limited. However, the data are suggestive that, consistent with evidence for the US from Plumley (1996), and the experimental evidence of Alm, Jackson and McKee (2009), the *indirect* revenue effects of audits likely dominate direct

revenue effects. Further, the indirect effects associated with a shift to a risk-based only audit system in the UK *could be* net revenue-negative.²⁹

ACKNOWLEDGEMENTS

We are grateful to HMRC's Analysis Department for giving us the opportunity to work on the data used in this paper and especially to former Director, Professor David Ulph. For helpful comments on earlier versions of this paper, we would also like to thank two referees and the editors of this *Journal*, and participants at seminars at the Centre for European Economic Research (ZEW), Mannheim, Germany, and the School of Economics, Victoria University of Wellington, New Zealand.

²⁹ In addition, eliminating random audits could result in any benefits they provide in fine-tuning the “red flags” that guide risk-based auditing being foregone.

REFERENCES

- Allingham, M.G. and Sandmo, A., 1972. "Income tax evasion: a theoretical analysis." *Journal of Public Economics* 1, 323–338.
- Alm, J., McClelland, G.H. and Schultz, W. D., 1992. "Why do people pay taxes?" *Journal of Public Economics* 48, 21-38.
- Alm, J. and McKee, M., 2004. "Tax compliance as a coordination game." *Journal of Economic Behavior & Organisation* 54, 297-312.
- Alm, J. and McKee, M., 2006. "Audit certainty, audit productivity and taxpayer compliance." *National Tax Journal* 59(4), 801-816.
- Alm, J., Jackson, B.R. and McKee, M., 2009. "Getting the word out: enforcement information dissemination and compliance behaviour." *Journal of Public Economics* 93, 392-402.
- Benjamini, Y. and Maital, S., 1985. "Optimal tax evasion and optimal tax evasion policy. Behavioral aspects." In W.Gaertner and A. Wenig (ed), *The Economics of the Shadow Economy*, 245-264. Berlin: Springer-Verlag.
- Dubin, J., Graetz, M. and Wilde, L., 1990. "The effect of audit rates on the federal individual income tax, 1977–1986." *National Tax Journal* 43, 395–409.
- Dubin, J., 2007. "Criminal investigation enforcement activities and taxpayer non-compliance." *Public Finance Review* 35(4), 500-529.
- Erard, B., 1992. "The influence of tax audits on reporting behavior." In J. Slemrod (ed), *Why People Pay Taxes. Tax Compliance and Enforcement*, 95-114. Ann Arbor: University of Michigan Press.
- Kim, Y., 2005. "Audit misperception, tax compliance and optimal uncertainty." *Journal of Public Economic Theory* 7(3), 521-541.
- Kirchler, E., 2007. *The Economic Psychology of Tax Behavior*. Cambridge University Press. Cambridge, UK.
- Maciejovsky, B., Kirchler, E. and Schwarzenberger, H., 2007. "Misperceptions of chance and loss repair: on the dynamics of tax compliance." *Journal of Economic Psychology* 28, 678-691.
- Lochner, L., 2007. "Individual perceptions of the criminal justice system." *American Economic Review* 97, 444-460.
- Long, S.B. and Schwartz, R.D., 1987. "The impact of IRS audits on taxpayers compliance: a field experiment in specific deterrence". Paper presented at the Annual Meeting of the Law and Society Association, Washington D.C.
- Mittone, L., 1999. "Psychological constraints on tax evasion: an experimental approach." Unpublished PhD thesis, Bristol University, UK.

Mittone, L., 2006. "Dynamic behavior in tax evasion: An experimental approach." *Journal of Socio-Economics* 35, 813-835.

National Audit Office (NAO), 2000. *Appropriation Accounts 1998-99 Volume 16: Class XVI Departments of the Chancellor of the Exchequer*. NAO: Report of the Comptroller and Auditor General.

Plumley, A., 1996. "The determinants of individual income tax compliance." *Internal Revenue Service Publication 1916 (Rev. 11-96)* Washington, DC.

Sandmo, A., 2005. "The theory of tax evasion: a retrospective view." *National Tax Journal* 58(4), 643-663.

Slemrod, J., Blumenthal, M. and Christian, C., 2001. "Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota." *Journal of Public Economics* 79, 455-483.

Snow, A. and Warren A., 2007. "Audit uncertainty, Bayesian updating and tax evasion." *Public Finance Review* 35(5), 555-571.

Spicer, M.W. and Hero, R.E., 1985. "Tax evasion and heuristics: a research note." *Journal of Public Economics* 26, 263-67.

Tauchen, H.V., Witte, A.D. and Beron, K.J., 1993. "Tax compliance: an investigation using individual TCMP data." *Journal of Quantitative Criminology* 9, 177-202.

Webley, P., 1987. "Audit probabilities and tax evasion in a business simulation." *Economics Letters* 25, 267-70.

Table 1
Sample Numbers [and Percent] of Audited and Controls

	Sample:		
	Medium (M)	Small (S)	Personal (P)
Control	2169 [79%]	2002 [75%]	2249[75%]
Audited	575 [21%]	655 [25%]	776 [25%]
of which:			
“compliant”	255	447	640
“non-compliant”	320	208	136
TOTAL	2744	2657	3015

Table 2
Case Closure and Average Direct Yields by Year

Year	Percent of returns submitted after audit closed			Index of average direct yield per non-compliant taxpayer * (2001 = 100)		
	Medium	Small	Personal	Medium	Small	Personal
2001	42	65	79	100	100	100
2002	83	96	98	168	176	52
2003	90	99	99	174	249	35
2004-05	100	100	100	426	312	4

* Based on average yield per non-compliant audited taxpayer.

Table 3
Difference-in-Difference Parameters
(*t*-ratios in parentheses)

Audited sub-group	Medium	Small	Personal
“Before/after 2000” regression:			
<i>All audited:</i>	128 (1.13)	-50 (-0.70)	-115 (-1.29)
<i>Non-compliant:</i>	342*** (2.67)	166 (1.50)	219 (1.44)
<i>Compliant:</i>	-143 (-1.01)	-151* (-1.90)	-187** (-2.40)
<i>No Obs.</i>	16,464	15,942	18,150
“Before 2000/after inquiry closure” regression:			
<i>Non-compliant:</i>	330*** (3.18)	321*** (3.37)	170 (1.23)
<i>Compliant:</i>	-312*** (-2.95)	-95 (-1.41)	-182*** (-2.60)
<i>No Obs.</i>	16,110	15,900	18,132

Notes: Regression parameters in this table have been re-scaled by a common, undisclosed factor to preserve confidentiality. The *t*-ratios shown in parentheses are the actual estimated values. *** significant at 1%, ** significant at 5%, * significant at 10%

Table 4
Preventive Yield
(As % of Declared Tax)

Audited sub-group	Medium	Small	Personal
Total corrective yield (using “after inquiry closure” parameters):			
<i>Non-compliant:</i>	+ 18%*	+ 24%*	+ 5%
<i>Compliant:</i>	- 17%*	- 11%	- 7%*
+/- 1 standard error			
<i>Non-compliant:</i>	± 6%	± 7%	± 4%
<i>Compliant:</i>	± 6%	± 7%	± 2%

* Based on D-i-D parameter significantly different from zero.

Table 5
Difference-in-Difference Parameters: Controlling for Direct Yield

Audited sub-group	Medium Businesses Parameter (t-ratio)	
“Before 2000/after inquiry closure” regression:		
<i>Compliant</i>	-121*** (-2.95)	-312*** (-2.95)
<i>Non-compliant</i>	-109 (0.50)	-258 (-0.91)
<i>Non-Compliant*Direct Yield Q2</i>	421 (1.36)	-
<i>Non-Compliant*Direct Yield Q3</i>	320 (1.03)	-
<i>Non-Compliant*Direct Yield Q4</i>	793*** (2.57)	-
<i>Non-Compliant*Direct Yield Q5</i>	670** (2.16)	-
<i>Non-Comp*Direct Yield Range £151-350</i>	-	543 (1.50)
<i>Non-Comp*Direct Yield Range £351-1000</i>	-	462 (1.40)
<i>Non-Comp*Direct Yield Range £1001-3000</i>	-	869*** (2.50)
<i>Non-Comp*Direct Yield Range >£3000</i>	-	968*** (2.55)
<i>No Observations</i>	16,110	16,110

Notes: Regression parameters in this table have been re-scaled as in Table 3. The *t*-ratios shown in parentheses are the actual estimated values. Q = quintile; *** significant at 1%, ** significant at 5%.

Table 6
Average Direct Yield by Segment Quintiles *

Quintile	Segment:		
	Medium	Small	Personal
Q1	9	4	1
Q2	25	10	5
Q3	57	19	12
Q4	139	46	28
Q5	473	216	255
ALL	141	58	60
Non-compliant taxpayers:	320	209	136

* Values have been scaled such that Personal Q1 = 1. Average yield (per non-compliant taxpayer) for this group was sufficiently small that it likely reflects mistakes rather than deliberate evasion.