

Who does and doesn't pay taxes?

Arun Advani*

University of Warwick, CAGE, and IFS

July 16, 2020

Abstract

We use administrative tax data from audits of self-assessment tax returns to understand what types individuals are most likely to be non-compliant. Non-compliance is common, with one-third of taxpayers underpaying by some amount, although half of aggregate under-reporting is done by just 2% of taxpayers. Third party reporting reduces non-compliance, while working in a cash-prevalent industry increases it. However, compliance also varies significantly with individual characteristics: non-compliance is higher for men and younger people. These results matter for measuring inequality, for understanding taxpayer behaviour, and for targeting audit resources.

I Introduction

A large share of tax owed goes uncollected. In the UK, around 6% of tax liabilities go uncollected each year ([HM Revenue and Customs, 2019](#)), similar in scale to the value of corporation tax receipts. In the US this number is 16.4% ([Internal Revenue Service, 2019](#)); in Canada 11.6% ([Canada Revenue Agency, 2019](#)). This issue has become politically important, with strong public support for reducing both avoidance and evasion ([Tax Justice UK, 2020](#)). It has also captured increasing academic attention, with work on ‘missing income’ and ‘missing wealth’

*This research was funded by the Economic and Social Research Council (ESRC) through the Centre for Microeconomic Analysis of Public Policy (CPP) at IFS (grant number ES/M010147/1), the Tax Administration Research Centre (TARC, grant number ES/K005944/1), the Centre for Competitive Advantage in a Global Economy (ES/L011719/1 and ES/S007121/1) and University of Warwick Impact Acceleration Awards (grant numbers ES/M500434/1 and ES/T502054/1). A previous version of this work circulated as [Advani \(2017\)](#). This work contains statistical data from HM Revenue and Customs (HMRC) which are Crown Copyright. The research data sets 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. The author thanks William Elming and Jonathan Shaw for their earlier involvement in this project, and Carl Emmerson, Paul Johnson, Helen Miller, Jonathan Shaw and Hannah Thompson for helpful comments. Correspondence: Dept of Economics, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL. Email: economics@arunadvani.com

(Zucman, 2015; Alstadsæter et al., 2018, 2019; Guyton et al., 2020). A key input for this research is a clear measure of where in the income distribution under-reporting is taking place, and how much income is under-reported (by income source) across this distribution. This is needed to accurately assign missing income and create corrected income distributions (Alvaredo et al., 2016; Piketty et al., 2018).

In this paper we use data from random tax audits in the UK to study *who* is underreporting income. We combine confidential administrative data on the universe of UK tax filers with a randomised audit programme. We show three main results. First, non-compliance is relatively prevalent, though it is often small. One-third of tax filers under-report income to some extent, but more than half the aggregate under-reporting is by the 2% most non-compliant taxpayers. This highlights the potential for policies like auditing to tackle the most extreme non-compliance. Second, the probability of non-compliance is relatively stable at around one-third across the income distribution. However, the share of tax that is under-reported declines across the distribution. Finally, individual characteristics including age and sex offer predictive power in estimating non-compliance. This is important for assigning missing incomes, particularly when estimated income distributions are then used to look at relative incomes of different groups, such as men and women (Garbinti et al., 2018; Burkhauser et al., 2020). It is also important for policy, since these can be used to target compliance resources such as audits.

Studies of tax enforcement sometimes compare the two main mechanisms observed from tax authorities: sending letters and conducting audits (Pomeranz and Vila-Belda, 2019; Slemrod, 2019; De Neve et al., 2020). Our results speak to why these approaches are not really comparable: they are attempting to solve different problems. ‘Low-level’ non-compliance is very prevalent, highlighting the need for scaleable mass-market interventions like letters (Slemrod et al., 2001; Kleven et al., 2011; Mascagni et al., 2017; Bergolo et al., 2020). But extreme non-compliance is highly concentrated, with most of the revenue loss coming from a small fraction of taxpayers, so more costly interventions such as audits can be targeted here (DeBacker et al., 2018; Advani et al., 2019; Hebous et al., 2019; Sarin and Summers, 2019). The two approaches are complementary, aimed at different types of non-compliance. Given the cost of audit interventions, our results on what characteristics best predict non-compliance and help target audits are also particularly valuable.

In trying to understand heterogeneity in tax compliance behaviour, a large literature has grown up around the idea of non-financial motivations to pay, sometimes called *tax morale* (Andreoni et al., 1998; Luttmer and Singhal, 2014; Dwenger et al., 2016). Early work used surveys to understand *reported* opinions about tax, such as the extent to which tax evasion can be justified, and correlated these with individual characteristics (Alm and Torgler, 2006; Torgler, 2006; Lago-Peñas and Lago-Peñas, 2010; Doerrenberg and Peichl, 2013). More recently lab experiments have been used to relate observed lab compliance with individual characteristics (Cummings et al., 2009; Choo et al., 2016; Guerra and Harrington, 2018). Our results allow a mapping from individual characteristics to behaviour in real-world tax decisions (Alm et al., 2015). We find qualitatively similar patterns to those on reported morale and lab compliance—

compliance is higher among women and increases with age. But we also show that some of the variation in compliance behaviour comes from characteristics of individuals' incomes: incomes that are easier to hide are more likely to be under-reported, and these are more common among men and the young. Evasion behaviour is thus the product of individual characteristics and opportunities to evade.

Beyond tax administration, our results speak to a broad literature in public economics that uses administrative data from tax authorities, rather than surveys, to understand patterns in inequality. Survey data are known to have issues with coverage, and the gap between surveys and national accounts has been rising over time (Atkinson et al., 2011; Ruiz and Woloszko, 2016; Burkhauser et al., 2018a,b; Webber et al., 2020). One reason is that income concepts are sometimes different between national accounts and what surveys are measuring (Advani et al., 2020; Corlett et al., 2020).¹ But even after defining a consistent measure of income, and weighting to account for non-response, there is substantial undercoverage. One proposed solution to ensure consistency with national accounts has been to implicitly assume that by income source, all individuals under-report incomes proportionally—the so-called Distributional National Accounts (DINA) method (Alvaredo et al., 2016; Garbinti et al., 2018; Artola Blanco and Martínez-Toledano, 2019). Our results show that under-reporting is common among a large set of UK taxpayers – one in three filers, one in ten taxpayers – but the majority do not appear to under-report. There is also a very skewed distribution of under-reporting. Hence this under-reporting needs to be corrected directly, before applying something like a DINA approach, and audits are an effective tool in determining how to implement these corrections. This approach complements other recent work focusing on adjustments to the income distribution to capture evasion at the top (Guyton et al., 2020).

The remainder of the paper is organised as follows. Section II provides some context about the aggregate size of the UK tax gap, how it compares internationally, and the relative contributions of different taxes to the gap. Section III outlines the policy context and data sources. Section IV provides evidence on the distribution of non-compliance, both marginally and then by incomes and characteristics. Section V concludes.

II UK tax gap

In the UK the *tax gap* is defined as ‘the difference between the amount of tax that should, in theory, be paid to HMRC, and what is actually paid’ (HM Revenue and Customs, 2019). Measuring what is actually paid is straightforward. Defining the theoretical liability is more contentious. HMRC define it as ‘the tax that would be paid if all individuals and companies complied with both the letter of the law and *our interpretation* of Parliament’s intention in setting law.’ This clearly incorporates (illegal) evasion behaviours, and also incorporates some

¹Another important reason is the rise in capital incomes, which tend to be less well-covered in surveys (Ooms, 2020) and are concentrated at the top (Advani and Summers, 2020b).

avoidance where HMRC deem that to be ‘unlawful’, and this is the definition we will focus on.²

II.1 Aggregate

The aggregate tax gap in the UK was £35 billion for the 2017-18 tax year, or 5.6% of total liabilities (HM Revenue and Customs, 2019). It has remained relatively stable, moving around between 5.3% and 6.5%, for more than a decade. To put these numbers into context, the five biggest taxes in the UK and (in parentheses) the share of revenue they raised in 2017-18 were: Income tax (25.8%), National insurance contributions (18.8%), Value added tax (17.9%), Corporation tax (7.9%) and Council tax (4.6%) (Office for Budget Responsibility, 2019). So the tax gap is equivalent in scale to between the fourth and fifth largest taxes in the UK. In terms of spending it is close to the amount spent on defence or central government spending on education.

Despite this, by international standards the UK tax gap is relatively low. The most recent tax gap estimates for the US, covering 2011-13, found a gap of 16.4% of revenue owed, triple the UK rate (Internal Revenue Service, 2016). In Canada, estimates for 2014 suggest the tax gap was around twice that in the UK, between 10.6-12.6% (Canada Revenue Agency, 2019). Few other countries produce comprehensive tax gap estimates, though some produce numbers for particular tax sources (and particular years) only (OECD, 2019, Table A146). Again the UK performs very well. So the need to understand how to reduce the tax gap, which is important in the UK, is even more necessary in other contexts.

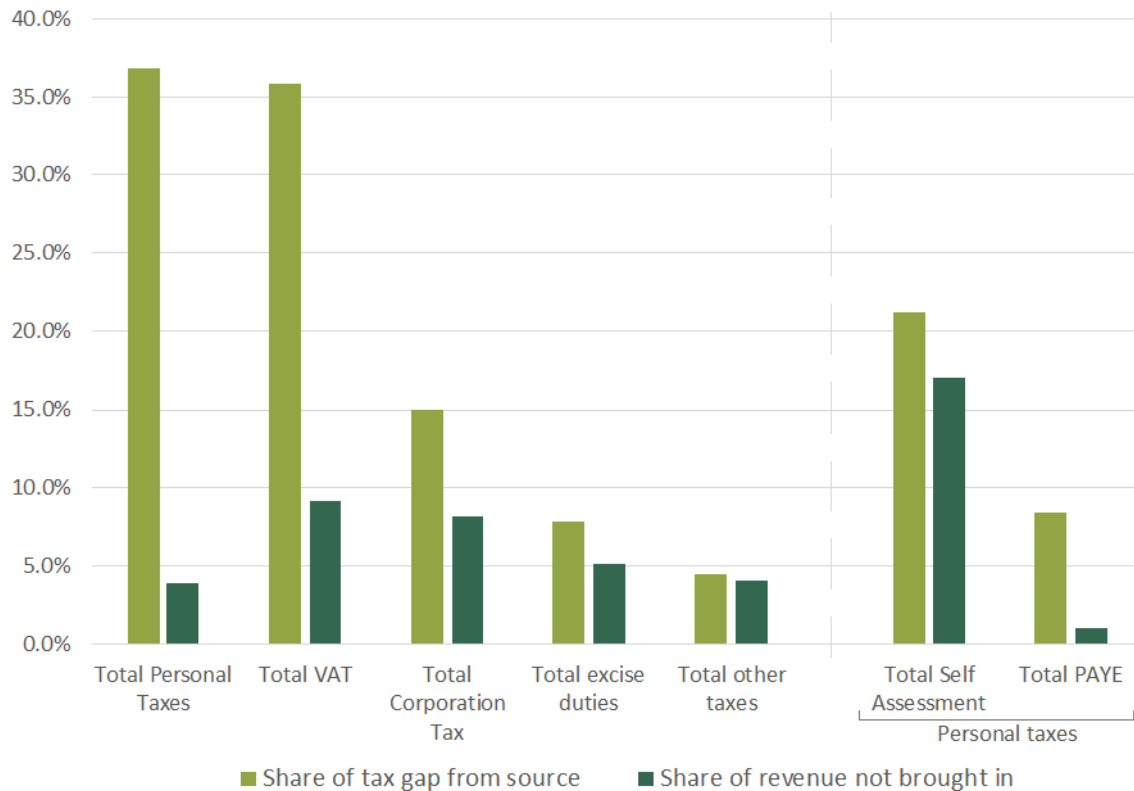
II.2 Structure

To reduce the tax gap, it is useful to think about which areas of tax have the largest gap. Figure 1 shows the share of the tax gap that comes from various sources, and the share of revenue due from that source that is not brought in (the ‘percentage gap’). Three features stand out. First, personal taxes – Income Tax, National Insurance Contributions, and Capital Gains Tax – and VAT make up around a third of the total tax gap each, at £12.9 billion and £12.5 billion respectively (HM Revenue and Customs, 2019). In one sense this is not surprising, since we have already seen that these are the biggest sources of tax revenue. More than half of the remaining gap is from corporation tax.

Turning to the share of revenue due that is not received, this is actually lowest for personal taxes, at only 3.9% (HM Revenue and Customs, 2019). By contrast VAT, which contributes the same share of the aggregate gap, has more than twice the percentage gap: one pound in every eleven (9.1%) is not paid. All other tax types fall somewhere in between. At first glance this might suggest relatively little scope for focusing on individual evasion. While the aggregate amount is large, it is only a small proportion of the revenue due from this source.

²It is worth noting that this definition excludes avoidance behaviours which HMRC interpret as being legal, for example corporate structures that reduce UK tax liabilities. This has led to ‘alternative’ tax gap measures being produced, which try to incorporate these and are consequently substantially larger.

Figure 1: Structure of the UK Tax Gap by Tax



Notes: ‘Share of tax gap from source’ shows the share of the overall tax gap that is attributable to that tax. ‘Share of revenue not brought in’ is the share of all revenue due from that tax that is never collected. Personal taxes comprise Income Tax, National Insurance Contributions and Capital Gains Tax. ‘Other taxes’ includes indirect taxes (Aggregates Levy, Air Passenger Duty, Climate Change Levy, Customs Duty, Insurance Premium Tax, Landfill Tax) and direct taxes (stamp duties, Inheritance Tax and Petroleum Revenue Tax). The share of the total gap from self assessment and PAYE is less than the total gap from personal taxes because the breakdown of personal taxes shown excludes avoidance (1.8%) and hidden economy (5.4%).

Source: HM Revenue and Customs (2019) Table 1.1.

However, within personal taxes there is substantial variation in the percentage tax gap, as seen in the right-hand panel. Almost all tax due through Pay As You Earn (PAYE), the UK’s withholding tax system, is received, with only a 1% percentage gap. By contrast, more than one pound in six (17%) of all tax due through the self-assessment system is not collected. This is by far the largest percentage gap for any tax ‘type’.³ It is also substantial in terms of its contribution to the aggregate tax gap: £7.4 billion of self assessment tax was not paid in 2018, more than one-fifth (21.2%) of the total gap (HM Revenue and Customs, 2019).

The key driver of this difference in percentage tax gap is that PAYE income is ‘third party reported’ and also withheld at source (Kleven et al., 2011; Slemrod and Gillitzer, 2013; Adhikari et al., 2020). Non-compliance therefore needs to be done by or in concert with one’s employer, which is relatively difficult (Kleven et al., 2016; Bjørneby et al., 2018). By contrast self assessment has more scope for non-compliance, since it is self-reported and (partially) self-remitted.

³The only tax component – a subcategory of tax ‘type’ – with a higher percentage tax gap is hand-rolling tobacco duty, which is estimated to be 33%. In cash terms this is only worth £0.7 billion. It is also likely to be over-estimated, depending on what share of hand-rolling tobacco paper is used for tobacco, and what share is used for smoking other substances (Parey and Rasul, forthcoming).

Table 1: **Share of Self-Assessment 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.

Source: Authors' calculations based on HMRC administrative datasets. Reproduced from [Advani et al. \(2019\)](#)

As we will show below, the level of this non-compliance also varies substantially by income source. To understand this better we next explain more about the structure of the UK income tax self-assessment system, and what verification and enforcement mechanisms are available.

III The self-assessment system

III.1 Who files?

Around one in four income taxpayers in the UK file taxes through the income tax self-assessment system (henceforth 'self assessment'): around 10 million people.⁴ In contrast to the US, where essentially all taxpayers are required to file, the UK has an effective withholding system (PAYE) which is able to correctly collect tax for three-quarters of the taxpaying population. Individuals who work in employment (rather than self-employment or partnerships), or who have pension income, will typically have tax withheld at source. If any additional capital income (interest, dividends) is sufficiently small that it falls below the tax-free allowances, then they have no additional tax to pay. Self assessment therefore covers individuals who have income from self-employment or partnerships, who have substantial income from dividends (more than £5000 in 2018) or interest (more than £500 or £1000 depending on income tax bracket), or who have any income from property, since in all these cases correct withholding does not take place. It also covers all individuals with employment income above £100,000 (£40,000 before 2006).

Individuals who are in self assessment are required to annually file a tax return reporting their income (and taxable capital gains). Data from these returns are captured and after

⁴There are around 40 million individuals with income 'assessable for tax'. Not all will pay tax on their income, as they may earn lower than the individual exempt amount (personal allowance). Nevertheless we describe them as taxpayers for simplicity.

processing are available via the HMRC Datalab in the *Valid View* dataset.⁵ Here we obtain information on individuals’ incomes and characteristics such as age, sex, region in which they live, and industry (for those who are self-employed or partners). Table 1 shows share of taxpayers with income from each source.

We combine these with data from the *SA302* dataset, which is produced from individuals’ tax calculations, and contains the tax due based on information about their incomes. We use these data sources over the period 1997 to 2012, and uprate all financial values to 2012 using the Consumer Price Index to account for inflation.

III.2 Verification and Enforcement

As described above, self assessment has proportionally by far the largest tax gap of any income type. To understand this, it is helpful to think of four broad ways in which compliance is achieved in the UK: (i) direct reporting, (ii) third party reporting, (iii) behavioural interventions, and (iv) audits. Understanding how each of these affects different parts of income reported on self assessment will help explain some of the patterns we see in non-compliance.

Direct reporting makes use of systems such as ‘fiscal tills’, so that transactions are recorded directly to the tax authority systems at the same point in time as they are recorded for the individual or business (Slemrod and Gillitzer, 2013). The UK is currently rolling out a more limited version of this – *Making Tax Digital* – which will send quarterly reports to the tax authority based on the entries in an organisation’s accounts. This will apply to businesses, but also self-employed people and landlords who have high enough receipts, and the programme is explicitly intended to reduce the tax gap by improving accuracy and compliance (HM Revenue and Customs, 2020). Although this is less direct than live reporting when transactions occur, it is likely to have lower administrative costs for individuals than live reporting, and is more applicable for cases where transactions don’t go through a till but are just direct transfers into bank accounts e.g. rental payments. Individuals can still keep two sets of accounts, but non-compliance becomes harder, and it becomes particularly difficult to be deliberately non-compliant and later claim it was in error. Although this system can be expected to reduce non-compliance among the self-employed and landlords, it was not at all in place during the period of our data.

Third-party reporting makes use of reports from counterparties to transactions. This can be done systematically, by automatic receipt of information and cross-checking with individual reports; or used in a non-systematic way, by contacting third parties only in cases where an investigation is taking place. The UK does both of these things. Employers directly report employment income to HMRC via the PAYE system, so HMRC can match this against self-reports in self assessment.⁶ This reduces *ex ante* non-compliance since individuals are aware

⁵For historical reasons, after processing the data are available in either the ‘Valid View’ or the ‘Invalid View’ datasets. An individual in a given year can only be in one or other of these data. We always append these to create a single dataset which we describe as *Valid View*.

⁶This is also done regularly with VAT in other countries, where both parties to the transaction have to report it and these reports are matched up (Pomeranz, 2015; Naritomi, 2019; Almunia et al., 2019).

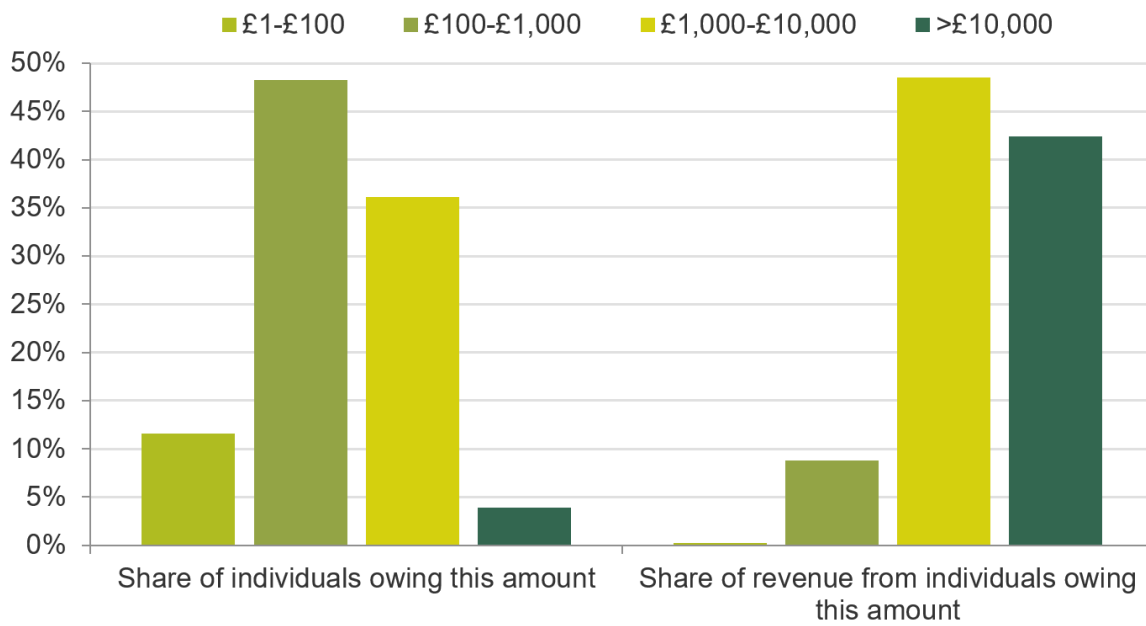
they are highly likely to have misreporting uncovered (Kleven et al., 2011). It also reduces *ex post* non-compliance, both because mismatches on systematically reported items can increase the likelihood a tax report is selected for audit. HMRC can also access bank and credit card records, and non-systematic checking of these third party reports can be used to check the plausibility of returns that are being audited, again reducing *ex post* non-compliance. As will be shown below, variation in which income sources are covered by third party reporting is important in understanding the variation in who is most likely to be non-compliant.

Behavioural interventions are designed to discourage misreporting. These are typically relatively low cost, and are largely used to improve compliance before tax returns are filed. For example, when filing electronically, the computer system can check box-by-box reports against various benchmarks – previous returns, returns of similar individuals, ratio of reported values across boxes – to see whether the declaration seems plausible. If not, the system asks for confirmation. This helps pick up errors, as well deliberate evasion. Since it makes use of benchmarks as well as third party reports, this can reduce non-compliance in areas where third party reports are currently not available—most notably dividend and property income in the UK. However, Fonseca and Grimshaw (2015) caution that interventions which pre-fill the form with information held by the tax authority can *reduce* compliance in some circumstances, if the pre-filled information is inaccurate.

Audits help to improve *ex post* compliance, by looking for erroneous tax returns and having them corrected. They are more costly than behavioural interventions, as they require time from a compliance officer to handle each case. But they are able to look more directly at what is reported, and make use of multiple information sources, to try to get an accurate picture of a taxpayer’s income. There is also evidence they change taxpayer behaviour in the years after audit (DeBacker et al., 2018; Advani et al., 2019).

Although most audits are targeted at taxpayers believed to be non-compliant, HMRC also performs a number of random audits (‘random enquiries’) to ensure all individuals have some chance of audit and to learn about characteristics that predict non-compliance. We make use of audit records from these random enquiries for tax years 1999 to 2009, available 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. Over our sample period there were 38,000 random audits carried out, just below 3,200 a year.⁷

Figure 2: Distribution of tax owed among the non-compliant



Notes: Left hand panel shows the share of individuals found to be under-reporting tax who under-reported by particular amounts. Right hand panel shows the share of all under-reported tax owed that is owed by people under-reporting particular amounts. **Source:** Authors' calculations based on HMRC administrative datasets.

IV Who is underpaying?

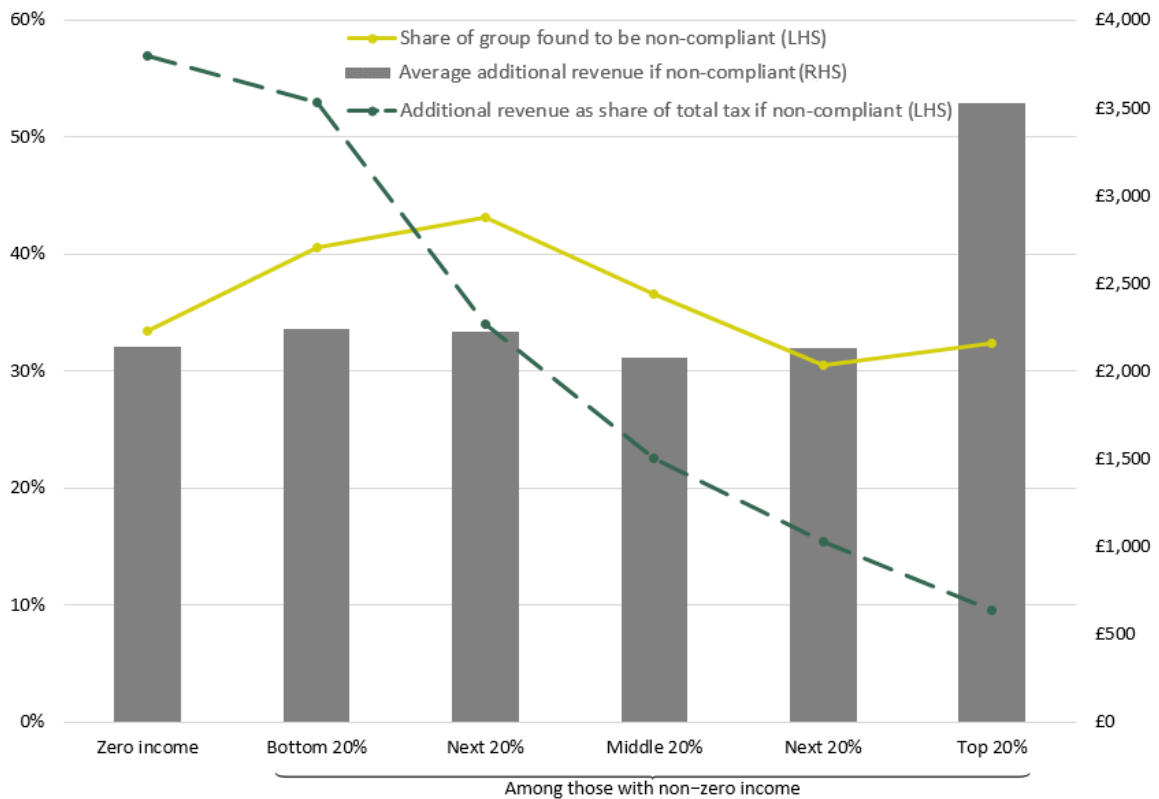
IV.1 Distribution of underpayment

The first key fact to understand in terms of compliance in self assessment is that around one-third (35.7%) of randomly audited individuals were found to be non-compliant. Most of the remainder had correct returns (53.2% of all taxpayers), while around a tenth (11.1%) were recorded 'correct with no underpayment'—either the return had an error that didn't affect tax liability, such a valid expense being claimed but filed under the wrong expense category, or in some cases errors that had led to over payment. Non-compliance is therefore relatively prevalent in the UK self-assessment system.

Despite the wide prevalence of some non-compliance, most of the under-reporting is due to a small minority. Figure 2 shows – among the non-compliant – who owes money and how much. Non-compliant taxpayers are divided into groups according to how much money the audit suggests that they owe: less than £100, between £100 and £1,000, between £1,000 and £10,000, or more than £10,000. 48% of taxpayers owe between £100 and £1,000, with a total of 60% owing less than £1,000. However, this 60% owe only 9% of the total missing tax. In contrast, 42% of the missing tax is owed by the 4% of people who owe more than £10,000. This highlights the enforcement value of audits in pursuing the small minority of individuals who owe extremely large sums of money. Just 200,000 people owe more than half of the self

⁷See Advani et al. (2019) for more details about the audit data including number of cases selected for audit and time trends.

Figure 3: Non-compliance by reported income quintile



Notes: Audited taxpayers grouped into zero income or income quintiles conditional on positive reported income, based on income report the year before audit. **Source:** Authors' calculations based on HMRC administrative datasets.

assessment tax gap – £3.7billion – between them.

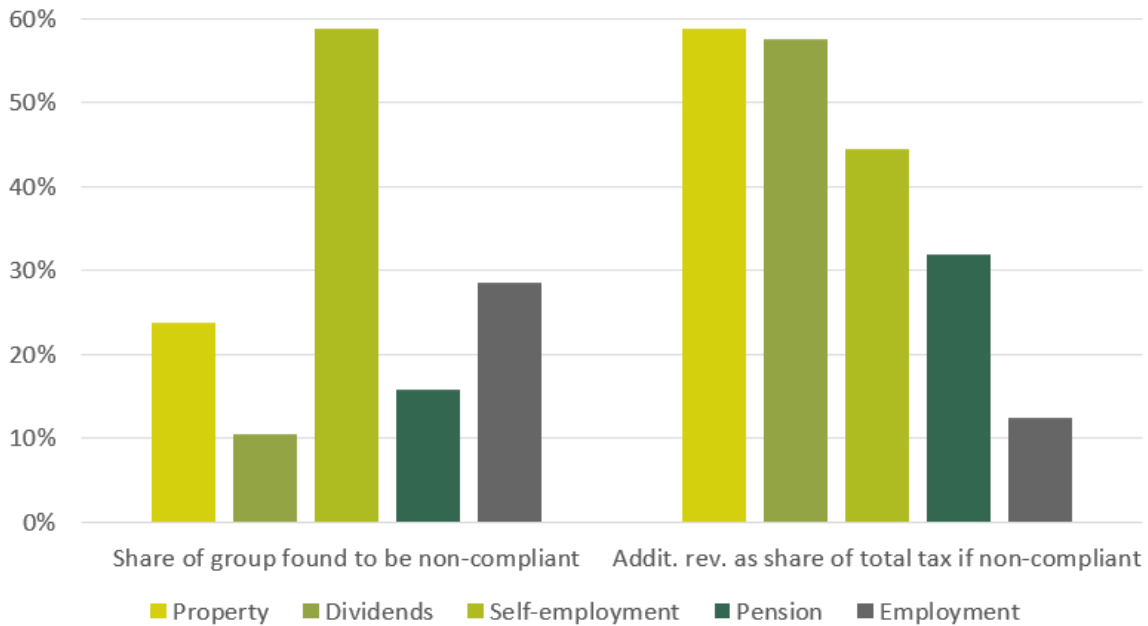
IV.2 Incomes of non-compliers

We next study the income characteristics of the non-compliance in terms of location in the income distribution, income source, and industry.

Evasion across the income distribution To understand how tax compliance varies with reported income, in Figure 3 we divide taxpayers up according to the level of total income they declare the year before they are audited. Around 16% of taxpayers previously declared a zero income, whilst the remaining taxpayers are divided into five equal groups. Across all these groups, the probability of being non-compliant was relatively similar. The average amount of non-compliance in cash terms was also similar, at around £2,200, for all but the highest income group, for whom it was £3,600. However, since total tax owed is generally larger with higher previous incomes, the share of tax misreported is lower for higher income groups.

Studies of inequality often make use of tax data to get measures of income (Piketty and Saez, 2003; Piketty et al., 2018; Smith et al., 2019; Joyce et al., 2019; Advani and Summers, 2020a), and they are increasingly recognising the need to adjust measured inequality to account

Figure 4: Non-compliance by income source



Notes: Based on individuals who report only a single income source (other than bank interest), since audit data don't explain which source(s) was/were erroneous where multiple sources exist. **Source:** Authors' calculations based on HMRC administrative datasets.

for evasion ([Alstadsæter et al., 2019](#)). In the UK these results suggest that inequality is slightly less bad than would be suggested by unadjusted fiscal income, since proportionally more income needs adding to the lower part of the distribution. However, it is worth noting that evidence from the US suggests that some of the more complex types of tax avoidance behaviours are harder to pick up, and so understatement at the very top may be a larger issue ([Guyton et al., 2020](#)).

Evasion by income source Comparing individuals who have income from only one income source other than bank interest gives an indication of how non-compliance varies by the type of income. Figure 4 shows that 59% of taxpayers declaring only self-employment income are found to be non-compliant, double the rate of non-compliance among those declaring only employment income (29%) and almost four times the rate among those declaring only pension income (16%). This is not surprising since – as described previously – self-employment income is almost entirely self-reported, and income is often received from a large number of small transactions, making it is easier both to make mistakes and to obscure deliberately.

More surprising is that non-compliance is so high for those reporting only employment income, even though such income is also reported by the employer. This may be because some other source of income was entirely unreported, rather than because of errors in reporting the employment income, or due to the reporting of illegitimate expenses (which are not third party reported). This highlights the importance of audits even in cases where non-compliance seems more unlikely. In terms of revenue, the largest amount of tax is owed by non-compliant

taxpayers declaring only property income. Although only 24% of those declaring only property income are found to be non-compliant, they owe an average of £3,630. This is almost 60% of the total tax they owe.

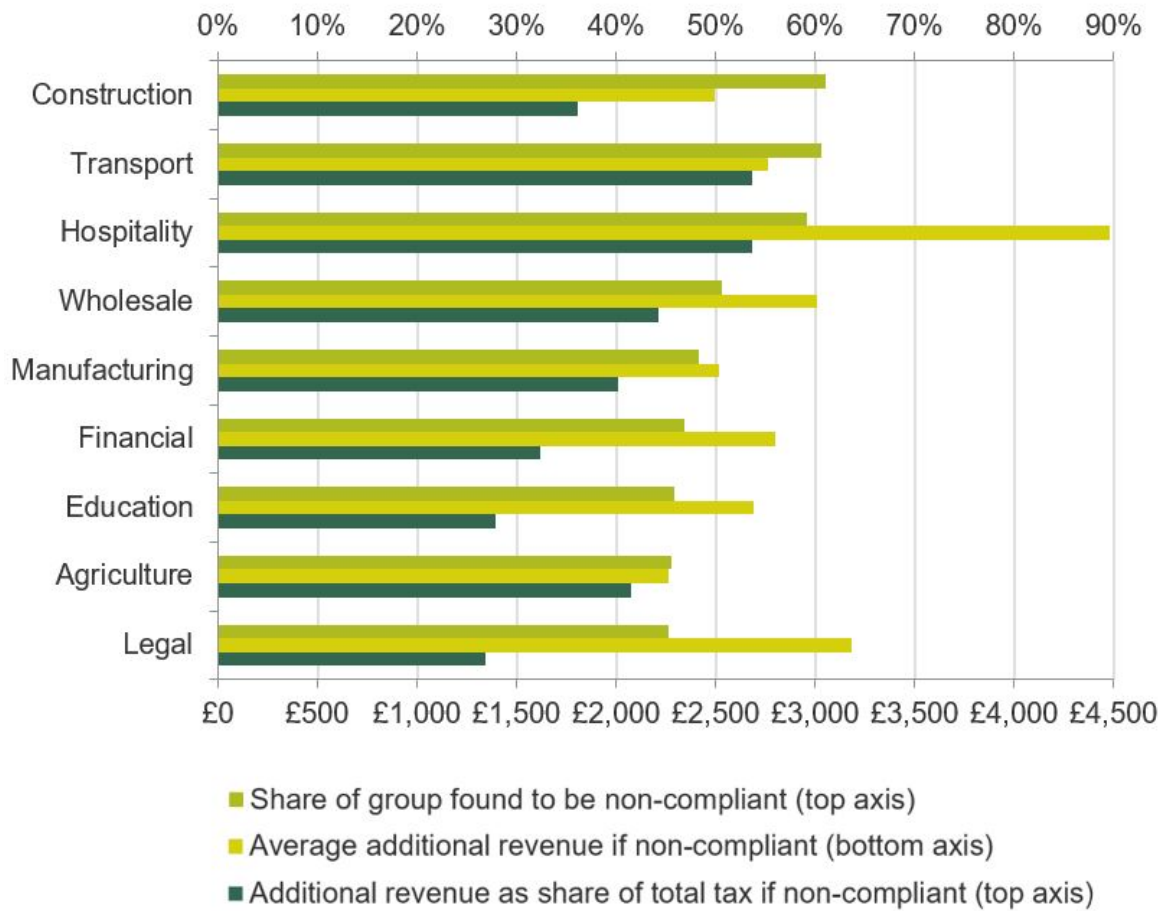
To compare these results to [Kleven et al. \(2011\)](#), they find 14.9% of self-employment income is underreported but only 2.3% of dividend income and 1.1% of employment income, not conditioning on compliance status. The comparable numbers from our context are 26.1%, 5.2%, and 3.5%. Underreporting 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.

Together these results emphasise an understudied aspect of third party reporting: the extent to which it matters for the extensive margin (declaring at all) versus the intensive margin (correct declarations conditional on reporting). Figure 4 orders income sources from lowest to highest extent of third party reporting in the UK.⁸ The right-hand panel shows clearly that third party information is important for the intensive margin, as it makes mismatches in reports easier to pick up. However, it appears to matter much less on the extensive margin, shown in the left-hand panel.

Evasion by industry Industry data are available for taxpayers who have any self-employment income. Figure 5 shows that among those individuals – 43% of all self-assessment taxpayers – non-compliance is highest among those in the construction, transport, and hospitality sectors. In these sectors, around 60% of taxpayers were non-compliant. In revenue terms, non-compliant taxpayers in hospitality owe by far the most, at an average of almost £4,500. Across other industries, the amount owed varies between £2,500 and £3,200. Taxpayers in both hospitality and transport under-reported the largest shares of the total tax they owed, at 54% in both cases.

Again this underlines the importance of third party information: within self-employment this is hardest to get for individuals working in industries where cash payment is common. Construction, transport (taxi and minicab drivers) and hospitality (bed and breakfasts) are all industries where cash is prevalent. This observable variation in where non-compliance is common provides a means for the tax authority to target audits. However, this also highlights the value in understanding how audit effects spillover between individuals, since in a standard Beckerian framework a large ‘campaign’ that increased the probability of audit among taxpayers of a particular type would likely, over time, reduce non-compliance even among the non-targeted. It is also worth noting that as cash use becomes generally less common, that should be expected to reduce under-reporting.

Figure 5: Non-compliance by industry



Notes: Industry classification based on aggregations of reported Trade Classification Number (TCN), among individuals with employment, self-employment, or partnership income. No industry is available for individuals with only pension, dividend, property or savings income. Where multiple income sources with associated TCN exist, the largest reported source is used. **Source:** Authors’ calculations based on HMRC administrative datasets.

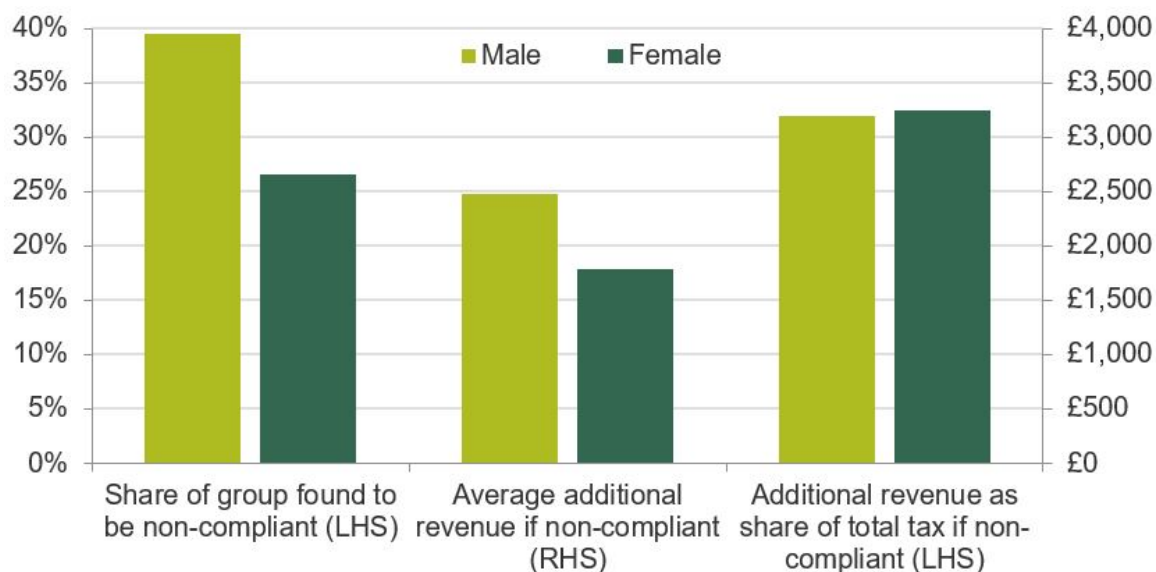
IV.3 Characteristics of non-compliers

While income characteristics change the opportunity for non-compliance, individuals may also vary in their willingness to under-report for other reasons. The literature on tax morale – the intrinsic disutility from misreporting – has studied how this varies across individuals with different characteristics (Luttmer and Singhal, 2014; Besley et al., 2019). For the most part this has relied on reported preferences about tax behaviour (Alm and Torgler, 2006; Torgler, 2006; Lago-Peñas and Lago-Peñas, 2010; Doerrenberg and Peichl, 2013) or lab experiments that elicit willingness to misreport (Cummings et al., 2009; Choo et al., 2016; Guerra and Harrington, 2018). The use of audits lets us instead understand how this varies in a real-world environment.

Knowing how behaviour varies with these characteristics can also be useful for targeting

⁸See Table A2 of Advani et al. (2019) for more details on the level of third party reporting by income source in the UK.

Figure 6: Non-compliance by sex



Source: Authors' calculations based on HMRC administrative datasets.

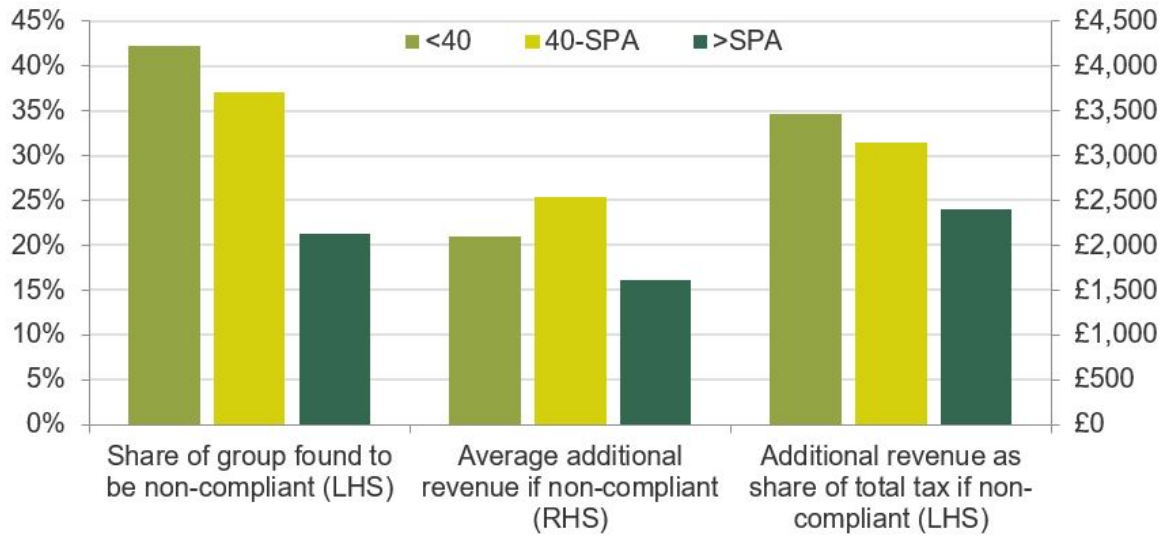
audits, since they are not easily manipulable, unlike characteristics of income. It is important to recognise that there may be ethical constraints to varying audit probabilities based on some of these characteristics.

Compliance by sex Figure 6 shows that men are more likely to be non-compliant. 40% of men are found to owe additional tax revenue, compared with only 27% of women. Non-compliant men also owe more money (£2,470 on average) than non-compliant women (£1,790). However, this difference is driven by differences in male and female incomes. The additional amount owed is 32% of the total tax owed for both non-compliant men and non-compliant women.

Much is made of gender pay gap statistics, including those making use of tax data. These rely on tax data being an accurate measure of incomes, and previous research has noted that the gender pay gap is smaller in tax data than in surveys (Britton et al., 2019). Given the prevalence of non-compliance and gender-disparity in compliance behaviour, these results indicate that the administrative data underestimate the pay gap, driven by the differences in extensive margin behaviour, and this could be part of the reason for the patterns previously observed.

Compliance by age Looking at non-compliance by age, Figure 7 shows that it is broadly similar for those below state pension age (SPA), with around 40% of individuals found to be non-compliant and, among these individuals, around a third of total tax being undeclared. Among those above SPA, non-compliance is much lower, at 21%, and less than a quarter of total tax due was undeclared among the non-compliant in this group. Part of this difference is driven by differences in the types of income received: it is harder to detect both mistakes and deliberate errors in declared self-employment income than in declared pension income.

Figure 7: Non-compliance by age



Notes: Individuals grouped into those aged less than 40, between 40 and State Pension Age (SPA), and those above the SPA. **Source:** Authors' calculations based on HMRC administrative datasets.

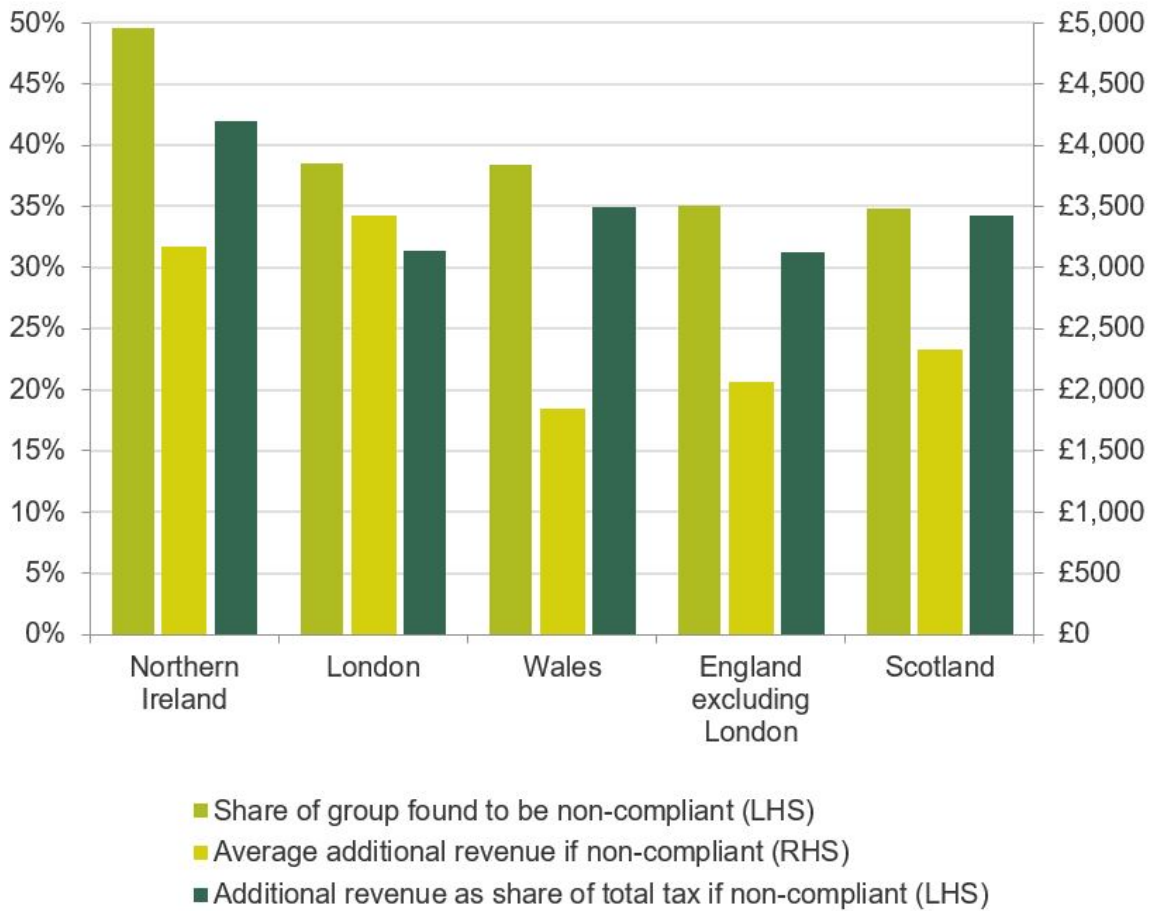
However, even controlling for income sources non-compliance is declining with age.

Our results on sex and age are consistent with [Kleven et al. \(2011\)](#), who find that in Denmark 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.

Compliance by geography Looking across the United Kingdom, Northern Ireland stands out as the region with the highest level of non-compliance. Figure 8 shows that 50% of self-assessment taxpayers from Northern Ireland were found to be non-compliant, compared with between 35% and 39% across the rest of the UK. Northern Ireland is also unusual in having an average of 42% of total tax being unreported among those who are non-compliant, compared with 31% to 35% elsewhere. However, in cash terms, non-compliant taxpayers from London owe the largest amount (£3,420), driven by higher incomes in the capital. Within England excluding London, there is little difference across regions. In contrast to [Kleven et al. \(2011\)](#), being located in the capital city is not associated with lower evasion. These differences are all statistically significant, and hold when other characteristics are controlled for.

With our data we are not able to study further why this regional variation exists. However, if access were ever made to regional information about audits, it would be interesting to study how regional enforcement policy related to regional levels of non-compliance. This question is of particular policy relevance at the moment, as in recent years smaller HMRC compliance offices have been (and are being) closed and replaced with a small number of regional hubs. This policy will create further geographic variation in the cost of engaging in audit work – which can involve site visits – likely increasing the variation in compliance behaviour.

Figure 8: Non-compliance by region



Notes: Regions defined by *Government Office Region* codes, with English regions apart from London being aggregated into a single group. **Source:** Authors' calculations based on HMRC administrative datasets.

The role of tax preparers Around 60% of taxpayers have a tax agent (accountants or other tax preparers). Of these, about 40% are found to be non-compliant, compared with just 29% of those without. Among these non-compliant taxpayers, the average additional tax owed by those with agents is £2,740, around 35% of their total tax owed; for those without it is only £1,400, or 27% of what they owe. Clearly people who have agents are likely to have more complex tax affairs than those without. This leaves more scope both for mistakes and for the use of allowances and other tax reduction methods. The comparison made here is therefore not intended to imply any kind of causal relationship. However, these numbers show that the use of agents, who have experience of preparing tax returns so can reduce mistakes, certainly does not solve the problem of non-compliance. Continued targeting of such taxpayers for audit is valuable.

V Conclusion

The UK tax gap is not large by international standards, but when compared with revenue from other taxes it is sizeable: it sits somewhere between corporation tax and council tax—the fourth

and fifth largest sources of tax revenue. The largest single gap – both in aggregate value and as a share of tax owed – is from taxpayers who file self assessment. Using data from HMRC’s ‘random enquiry’ audit programme, we are able to shed light on *who* it is that owes this money: what types of individuals. The patterns we describe are not intended to be interpreted causally, since some of these characteristics are choices for the taxpayer. Nevertheless, the associations are interesting to understand for four reasons.

First, measures of inequality – whether by age, sex, or income shares across the distribution – often rely on tax data (Piketty and Saez, 2003; Piketty et al., 2018; Smith et al., 2019; Joyce et al., 2019; Advani and Summers, 2020a). Differential non-compliance based on these different characteristics will bias such measures, unless corrected for. Our results allow researchers to sign the bias, and to make corrections by imputing the missing income according to these characteristics. Given the importance of these statistics for both policy and public debate, such corrections are essential to provide an accurate understanding of the levels and trends in inequality.

Second, we generally want to understand the differences in individual behaviour. To the extent that this varies with observables, and those observables can be related to preference parameters such risk aversion, one can study how those parameters relate to observed compliance behaviour in a real world setting. Lab settings typically let one elicit these preferences simultaneously with behaviours, but given the importance of framing effects it isn’t clear that responses to tax experiments in a lab overlap well with real behaviour. For example, Advani et al. (2019) note that although a number of lab experiments find evidence of so-called ‘bomb crater’ effects in compliance (Maciejovsky et al., 2007; Kastlunger et al., 2009), with individuals being less compliant immediately after audit, such behaviours are not seen in administrative tax data—in fact the opposite is observed.

Third, these results underscore the value of having third party information in reducing non-compliance (Kleven et al., 2011). The UK does not have systematic asset registers, so unlike many other countries shares and property income cannot be systematically verified. This increases the scope for non-compliance. A policy move towards increasing access to information for the tax authority would reduce non-compliance in a broad way without having the high marginal costs of an audit.

Finally, our results are relevant for the current targeting of audits, though of course directly using them for targeting is subject to the Lucas critique. They highlight the value of using observable characteristics, even ones that do not directly relate to tax owed, in targeting compliance work. This naturally leads to important ethical and legal questions: is it reasonable to use information about sex in determining who to audit. Administrative datasets on tax do not include ethnicity information, but this seems a clear context in which people may feel uncomfortable about such targeting, if it were to be relevant. These are trade-offs that must be confronted in making policy. The alternative is not obviously less harmful: many people who are unlikely to be non-compliant receive tax audits – costly for the individual audited as well as for the taxpayer – because data that are informative about their likelihood of non-compliance

was not used. In this sense we fit into a new and growing literature on using characteristics to predict legal outcomes ([Skeem and Lowenkamp, 2016](#); [Fagan and Ash, 2017](#); [Amaranto et al., 2018](#)).

To make further progress on policy implications, more needs to be understood about how taxpayers learn and respond to audit risk. [Advani et al. \(2019\)](#) and [DeBacker et al. \(2018\)](#) show that audits shift the behaviour of taxpayers who *receive* audits. But a simple Allingham-Sandmo model of rational non-compliance would imply that as the *risk* of being audited rises for particular groups, compliance by that group should increase even among the non-audited ([Allingham and Sandmo, 1972](#); [Yitzhaki, 1987](#)). For policy this is a positive effect, since non-compliance reduces faster in audit probabilities than if behaviour were static. Quantifying this behaviour is a natural next step in thinking about the number, as well as targeting, of audits.

References

- ADHIKARI, B., J. ALM, AND T. F. HARRIS (2020): “Information Reporting and Tax Compliance,” *AEA Papers and Proceedings*, 110.
- ADVANI, A. (2017): “Who does and doesn’t pay taxes?” Institute for Fiscal Studies BN218.
- ADVANI, A., W. ELMING, AND J. SHAW (2019): “The Dynamic Effect of Tax Audits,” CAGE 414.
- ADVANI, A. AND A. SUMMERS (2020a): “Capital Gains and UK Inequality,” CAGE 465.
- (2020b): “How much tax do the rich really pay? New evidence from tax microdata in the UK,” CAGE 27.
- ADVANI, A., A. SUMMERS, AND H. THOMPSON (2020): “What has been happening to UK top incomes?” CAGE 490.
- ALLINGHAM, M. AND A. SANDMO (1972): “Income Tax Evasion: A Theoretical Analysis,” *Journal of Public Economics*, 1, 323–338.
- ALM, J., K. M. BLOOMQUIST, AND M. MCKEE (2015): “On the external validity of laboratory tax compliance experiments,” *Economic Inquiry*, 53, 1170–1186.
- ALM, J. AND B. TORGLER (2006): “Culture differences and tax morale in the United States and in Europe,” *Journal of Economic Psychology*, 27, 224–246.
- ALMUNIA, M., J. HJORT, J. KNEBELMANN, AND L. TIAN (2019): “Strategic or Confused Firms? Evidence from “Missing” Transactions in Uganda,” *mimeo*.
- ALSTADSÆTER, A., N. JOHANNESSEN, AND G. ZUCMAN (2018): “Who owns the wealth in tax havens? Macro evidence and implications for global inequality,” *Journal of Public Economics*, 162, 89–100.
- (2019): “Tax evasion and inequality,” *American Economic Review*, 109, 2073–2103.
- ALVAREDO, F., A. ATKINSON, L. CHANCEL, T. PIKETTY, E. SAEZ, AND G. ZUCMAN (2016): “Distributional National Accounts (DINA) guidelines: Concepts and methods used in WID.world,” WID.world 2.
- AMARANTO, D., E. ASH, D. L. CHEN, L. REN, AND C. ROPER (2018): “Algorithms as prosecutors: lowering rearrest rates without disparate impacts and identifying defendant characteristics ‘Noisy’ to human decision-makers,” *mimeo*.
- ANDREONI, J., B. ERARD, AND J. FEINSTEIN (1998): “Tax compliance,” *Journal of Economic Literature*, 36, 818–860.

- ARTOLA BLANCO, M. AND C. MARTÍNEZ-TOLEDANO (2019): “On the Long-Run Evolution of Income in Spain, 1900-2017,” Paris School of Economics.
- ATKINSON, A. B., T. PIKETTY, AND E. SAEZ (2011): “Top incomes in the long run of history,” *Journal of Economic Literature*, 49, 3–71.
- BERGOLO, M., R. CENI, G. CRUCES, M. GIACCOBASSO, AND R. PEREZ-TRUGLIA (2020): “Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment,” Tech. Rep. 23631, National Bureau of Economic Research.
- BESLEY, T., A. JENSEN, AND T. PERSSON (2019): “Norms, enforcement, and tax evasion,” Tech. rep., National Bureau of Economic Research.
- BJØRNEBY, M., A. ALSTADSÆTER, AND K. TELLE (2018): “Collusive Tax Evasion by Employers and Employees: Evidence from a Randomized Field Experiment in Norway,” CESifo Working Paper Series 7381.
- BRITTON, J., N. SHEPHARD, AND A. VIGNOLES (2019): “A comparison of sample survey measures of earnings of English graduates with administrative data,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182, 719–754.
- BURKHAUSER, R. V., N. HÉRAULT, S. P. JENKINS, AND R. WILKINS (2018a): “Survey Under-Coverage of Top Incomes and Estimation of Inequality: What is the Role of the UK’s SPI Adjustment?” *Fiscal Studies*, 39, 213–240.
- (2018b): “Top incomes and inequality in the UK: reconciling estimates from household survey and tax return data,” *Oxford Economic Papers*, 70, 301–326.
- BURKHAUSER, R. V., N. HÉRAULT, S. P. JENKINS, AND R. WILKINS (2020): “What Accounts for the Rising Share of Women in the Top 1%?” Working Paper 27397, National Bureau of Economic Research.
- CANADA REVENUE AGENCY (2019): “Tax Gap and Compliance Results for the Federal Corporate Income Tax System,” Report.
- CHOO, C. L., M. A. FONSECA, AND G. D. MYLES (2016): “Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment,” *Journal of Economic Behavior & Organization*, 124, 102–114.
- CORLETT, A., A. ADVANI, AND A. SUMMERS (2020): “Who gains? The importance of accounting for capital gains,” Report, Resolution Foundation.
- CUMMINGS, R. G., J. MARTINEZ-VAZQUEZ, M. MCKEE, AND B. TORGLER (2009): “Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment,” *Journal of Economic Behavior & Organization*, 70, 447–457.

- DE NEVE, J.-E., C. IMBERT, J. SPINNEWIJN, T. TSANKOVA, AND M. LUTS (2020): “How to improve tax compliance? Evidence from population-wide experiments in Belgium,” *CAGE* 458.
- 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.
- DOERRENBERG, P. AND A. PEICHL (2013): “Progressive taxation and tax morale,” *Public Choice*, 155, 293–316.
- DWENGER, N., H. KLEVEN, I. RASUL, AND J. RINCKE (2016): “Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany,” *American Economic Journal: Economic Policy*, 8, 203–32.
- FAGAN, J. AND E. ASH (2017): “New Policing, New Segregation: From Ferguson to New York,” *Georgetown Law Journal*, 106, 33.
- FONSECA, M. AND S. GRIMSHAW (2015): “Does using behavioural prompts in pre-populated tax forms affect compliance? results from an artefactual field experiment with real taxpayers,” Tax Administration Research Centre 015-15.
- GARBINTI, B., J. GOUPILLE-LEBRET, AND T. PIKETTY (2018): “Income inequality in France, 1900–2014: evidence from distributional national accounts (DINA),” *Journal of Public Economics*, 162, 63–77.
- GUERRA, A. AND B. HARRINGTON (2018): “Attitude–behavior consistency in tax compliance: A cross-national comparison,” *Journal of Economic Behavior & Organization*, 156, 184–205.
- GUYTON, J., P. LANGETIEG, D. RECK, M. RISCH, AND G. ZUCMAN (2020): “Tax Evasion by the Wealthy: Measurement and Implications,” in *Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth*, University of Chicago Press.
- HEBOUS, S., Z. JIA, K. LØYLAND, T. O. THORESEN, AND A. ØVRUM (2019): “Behavioral Responses to Tax Audits: Between Deterrence and Approval,” *mimeo*.
- HM REVENUE AND CUSTOMS (2019): “Measuring Tax Gaps 2019 edition,” Official Statistics release.
- (2020): “Overview of Making Tax Digital,” Policy paper.
- INTERNAL REVENUE SERVICE (2016): “Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008-2010,” Publication 1415 (rev. 5-2016), Research, Analysis & Statistics.
- (2019): “Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2011-2013,” Publication 1415 (rev. 9-2019), Research, Analysis & Statistics.

- JOYCE, R., T. POPE, AND B. ROANTREE (2019): “The characteristics and incomes of the top 1%,” Tech. rep., IFS Briefing Note BN253.
- 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.
- KLEVEN, H. J., C. T. KREINER, AND E. SAEZ (2016): “Why can modern governments tax so much? An agency model of firms as fiscal intermediaries,” *Economica*, 83, 219–246.
- LAGO-PEÑAS, I. AND S. LAGO-PEÑAS (2010): “The determinants of tax morale in comparative perspective: Evidence from European countries,” *European Journal of Political Economy*, 26, 441–453.
- 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.
- MASCAGNI, G., C. NELL, AND N. MONKAM (2017): “One size does not fit all: a field experiment on the drivers of tax compliance and delivery methods in Rwanda,” .
- NARITOMI, J. (2019): “Consumers as tax auditors,” *American Economic Review*, 109, 3031–72.
- OECD (2019): *Tax Administration 2017*, OECD Publishing: Paris.
- OFFICE FOR BUDGET RESPONSIBILITY (2019): “Economic and fiscal outlook – March 2019,” Report.
- OOMS, T. C. (2020): “Income from capital, inequality indicators and social policy,” Ph.D. thesis, University of Oxford.
- PARÉY, M. AND I. RASUL (forthcoming): “Measuring the market size for Cannabis: A new approach using forensic economics,” *Economica*.
- PIKETTY, T. AND E. SAEZ (2003): “Income inequality in the United States, 1913–1998,” *Quarterly Journal of Economics*, 118, 1–41.
- PIKETTY, T., E. SAEZ, AND G. ZUCMAN (2018): “Distributional national accounts: methods and estimates for the United States,” *Quarterly Journal of Economics*, 133, 553–609.

- POMERANZ, D. (2015): “No taxation without information: Deterrence and self-enforcement in the value added tax,” *American Economic Review*, 105, 2539–69.
- POMERANZ, D. AND J. VILA-BELDA (2019): “Taking state-capacity research to the field: Insights from collaborations with tax authorities,” *Annual Review of Economics*, 11, 755–781.
- RUIZ, N. AND N. WOLOSZKO (2016): “What do household surveys suggest about the top 1% incomes and inequality in OECD countries?” .
- SARIN, N. AND L. H. SUMMERS (2019): “Shrinking the Tax Gap: Approaches and Revenue Potential,” National Bureau of Economic Research 26475.
- SKEEM, J. L. AND C. T. LOWENKAMP (2016): “Risk, race, and recidivism: Predictive bias and disparate impact,” *Criminology*, 54, 680–712.
- SLEMROD, J. (2019): “Tax compliance and enforcement,” *Journal of Economic Literature*, 57, 904–54.
- SLEMROD, J., M. BLUMENTHAL, AND C. CHRISTIAN (2001): “Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota,” *Journal of Public Economics*, 79, 455–483.
- SLEMROD, J. AND C. GILLITZER (2013): *Tax Systems*, MIT Press.
- SMITH, M., D. YAGAN, O. ZIDAR, AND E. ZWICK (2019): “Capitalists in the Twenty-first Century,” *Quarterly Journal of Economics*, 134, 1675–1745.
- TAX JUSTICE UK (2020): “Taxpayer views on public spending, tax and wealth,” Data retrieved from https://www.taxjustice.uk/uploads/1/0/0/3/100363766/taxpayer_views_on_public_spending_tax_and_wealth_6_april_2020.xlsx.
- TORGLER, B. (2006): “The importance of faith: Tax morale and religiosity,” *Journal of Economic Behavior & Organization*, 61, 81–109.
- WEBBER, D., R. TONKIN, AND M. SHINE (2020): *Using Tax Data to Better Capture Top Incomes in Official UK Income Inequality Statistics*, University of Chicago Press.
- YITZHAKI, S. (1987): “On the Excess Burden of Tax Evasion,” *Public Finance Quarterly*, 15, 123–137.
- ZUCMAN, G. (2015): *The hidden wealth of nations: The scourge of tax havens*, University of Chicago Press.