

# Response to a Comment by Auten and Splinter on “Tax Evasion at the Top of the Income Distribution: Theory and Evidence” \*

Daniel Reck

London School of Economics

Max Risch

Carnegie Mellon University

Gabriel Zucman

University of California Berkeley

October 22, 2021

## Abstract

In this article we respond to a comment by Gerald Auten and David Splinter on our recent working paper. Auten and Splinter’s comment does not challenge the basic argument of our paper about the difficulties in estimating the tax gap at the top of the income distribution with data from random audits alone. Rather, they assert that certain methodological choices created an “inherent upward bias” in our benchmark illustrations of how accounting for the limitations of random audit data affects estimates of the tax gap at the top. We agree that, partly due to data limitations that we point out in our paper, the top tax gap is quantitatively uncertain. However, here we show that principled reasoning and further data analysis contradict the assertions that led Auten and Splinter to argue for an upward bias, and a number of factors ignored by Auten and Splinter suggest that our benchmark estimates of top-end evasion may actually be under-estimated on net. We respond to each of the issues they raised and ultimately conclude that, while it is important to always bear in mind that tax gap estimates are uncertain, our original benchmark estimates are in fact reasonable and conservative.

---

\*Corresponding Author: Daniel Reck, [d.h.reck@lse.ac.uk](mailto:d.h.reck@lse.ac.uk).

*The views expressed here are those of the authors and do not necessarily reflect the official view of the Internal Revenue Service. This project was conducted through the Joint Statistical Research Program of the Statistics of Income Division of the IRS. All data work for this project involving confidential taxpayer information was done on IRS computers, by IRS employees, and at no time was confidential taxpayer data ever outside of the IRS computing environment. Reck and Risch are IRS employees under an agreement made possible by the Intragovernmental Personnel Act of 1970 (5 U.S.C. 3371-3376).*

# 1 Introduction

We posted a working paper, entitled “Tax Evasion at the Top of the Income Distribution: Theory and Evidence” as an NBER working paper in March of 2021.<sup>1</sup> The paper was joint work with ourselves and to IRS coauthors, John Guyton and Patrick Langetieg. The main findings of this paper, which remains a working paper subject to revision before publication, are as follows:

1. Data from National Research Program (NRP) random audits, as they exist currently, substantially under-measure tax evasion at the top of the income distribution.
2. Evasion via under-reporting of pass-through business income and via offshore financial accounts are two important forms of evasion at the top that are substantially under-measured by these random audits.
3. These two forms of evasion matter quantitatively for aggregate statistics like total under-reporting and the total tax gap at the top, and top 1% fiscal income shares adjusted for under-reported income.

In May 2021, Gerald Auten and David Splinter circulated a comment on this paper (Auten and Splinter (2021), henceforth AS). The general thrust of this comment is that several of the assumptions we made to arrive at our quantitative benchmark estimates regarding the third point above created an “inherent upward bias” in our estimates of overall evasion at the top of the income distribution. This supposedly in turn led us to over-state the impact of accounting for under-reporting on top 1% fiscal income shares. The comment does not dispute the first two findings above.

In this article, we respond to these assertions. We agree that there is substantial uncertainty about the exact quantitative magnitude of the income tax gap at the top of the distribution, and we thank Auten and Splinter for pushing us to think harder about this uncertainty along a number of dimensions. We tried to convey the uncertainty in our quantitative results via sensitivity analysis in our working paper. Much of the uncertainty is a consequence of the first two findings above: evasion that is difficult to detect by audits is inherently difficult to quantify. However, we argue that our benchmark estimate is not clearly, unambiguously, or inherently biased toward estimating too much evasion at the top. Rather, we show that 1) some issues raised as potential source of upward bias by Auten and Splinter actually reflect unknowns that could bias our estimates in either direction (with no compelling reason to expect bias in a particular direction); 2)

---

<sup>1</sup>The NBER working paper is available here: <https://www.nber.org/papers/w28542>. All cross-references here to results from the working paper correspond to the NBER version of this paper, posted in March 2021. Readers experiencing difficulty finding this version are welcome to email Daniel Reck.

As we noted in the working paper, we use the term evasion here to denote any form of non-compliance with tax law, whether intentional or unintentional. This usage is consistent with other academic writing in economics (see e.g. Kleven et al. (2011)). What we call evasion here is typically called “non-compliance” in official IRS publications and legal writing, while the term evasion is reserved for intentional, criminal conduct. We take no stand on what portion of the tax gap reflects intentional, criminal conduct versus other non-compliance.

other issues they raise do not stand up to scrutiny, and 3) several factors that they ignore suggest that our current estimates of evasion at the top may be too low. We conclude that while these aggregate statistics are uncertain and further research should attempt to refine them, our benchmark estimates remain a reasonable and conservative illustration of the importance of evasion un-detected by random audits for the tax gap at the top of the distribution and other aggregate statistics.

Regarding broader, ongoing discussions about how to measure the tax gap, we emphasize that nothing in our study contradicts that random audit programs are immensely useful for research on the tax gap. Measuring tax evasion is very difficult. The most popular approach worldwide pins its hopes on random audit data. Random audit data have highly advantageous features: they are *representative* of the full population and *comprehensive* in attempting to capture all types of individual non-compliance. The broad lesson of our paper, however, is that sophistication makes it difficult to detect some significant forms of evasion via random audits. Specifically, our paper documents evidence that these limitations apply to NRP random audits *as they are currently conducted in the US.*<sup>2</sup> The limitations of random audit data stem from the limitations of the underlying audit procedures (and not, for instance, from the random selection of cases). The procedures used in NRP random audits are the same as general audit procedures for most operational audits of individual tax filers, so the same concerns about detecting sophisticated types of evasion could apply to operational audit data, though there are some specialized audit programs that take aim at more sophisticated evasion. The root cause of the difficulty in arriving at quantitative estimates for our third finding above is that there is no available data that retains all the benefits of random audit data *and* overcomes their limitations in capturing top-end evasion. Until more perfect data may be gathered, perhaps by a more perfect random audit procedure, the best approach available to measuring the tax gap is to supplement random audit data with additional, potentially non-representative or partial data on other forms of evasion. Doing so requires confronting some uncertainty in estimates of the tax gap, especially at the top of the distribution, which we seek to do in a transparent and principled fashion.

In what follows, we divide the specific comments by Auten and Splinter, into three categories and respond to each in turn. The first category involves distributional questions around the procedure called Detection Controlled Estimation (DCE), which is used in official tax gap statistics to account for types of evasion detected by some but not all auditors conducting random audits. The second category concerns how we quantify and distribute evasion in pass-through businesses. These two categories comprise the bulk of the issues raised by AS about our benchmark estimates. We address each of the miscellaneous other issues raised by AS in a residual third category. Finally, we discuss some bigger picture issues, in an attempt

---

<sup>2</sup>We note that random audits in other countries face similar challenges [Alstadsaeter et al. \(2019\)](#), suggesting that our findings are not entirely unique to the 2006-2013 NRP in the US.

to place our work and this comment in the context of the broader theoretical and empirical literature on tax evasion and inequality.

## 2 On Detection Controlled Estimation

**Overview.** A sizable share of the tax gap in official statistics (IRS, 2007, 2016, 2019; Johns and Slemrod, 2010), about one half to two thirds of the tax gap for individual income under-reporting, is attributable to DCE. The core motivation for employing DCE is that there is a sizable difference between estimated evasion detected in an average National Research Program (NRP) random audit versus NRP audits conducted by the most thorough auditors. The most thorough auditors are estimated to find about three times as much under-reporting as the average auditors.<sup>3</sup> Official statistics therefore estimate the total amount of individual income tax assessed in a counterfactual scenario in which the full population were audited *by the most comprehensive auditors*. Thus total individual income under-reporting and the individual income tax gap are about three times larger than would be suggested from random audits without this adjustment.

The DCE methods we employ are based on Johns and Slemrod (2010), which details the methods used in the 2001 Tax Gap study (IRS, 2007), and applies them for distributional analysis exactly like we have done in our working paper. Specifically, we divide taxpayers into two categories: one group has either reported sole proprietorship income (Schedule C income) or total positive income (TPI) greater than \$100,000, and the second group is the complement of the first. Within each group, detected income under-reporting for each type of income is scaled at the micro level by one of two multipliers, depending on whether the income in question is designated “high-” or “low-visibility” income (essentially whether the income is supported by substantial third-party information reporting). Based on estimated auditor effects models, returns with sole proprietorship income or TPI > \$100,000 get multipliers of 3.358 for low-visibility income and 2.340 for high-visibility income; returns with no sole proprietorship income and TPI < \$100,000 get multipliers of 4.148 for low-visibility income and 2.009 for high-visibility income. Because a large share of detected under-reporting is low-visibility income for those with business income or losses, the 3.358 multiplier applies to a large majority of all under-reporting. Finally, the above apply to detected under-reporting, while we make no DCE adjustments in the case of detected over-reporting (or equivalently we use a multiplier of 1 in these cases). For further details, we refer readers to Johns and Slemrod (2010). For what follows, we refer to this basic approach as the “uniform multiplier” method, because holding the taxpayers group and type of income fixed, all detected under-reporting is scaled up by the same multiplier by DCE adjustments.

---

<sup>3</sup>One concern with this approach is that what we call the “most thorough auditors” might in fact be over-aggressive and assess tax beyond what the law implies. All tax gap statistics are based on auditors’ initial recommended adjustments, without further adjustments for disputes from taxpayers. We acknowledge this concern, but our goal here is to take the official macro amount of the tax gap as given and estimate a reasonable distribution for this amount (and then add evasion that is likely missed at the top).

This approach to DCE was designed to estimate total under-reporting in the full population, but it was not intended for distributional analysis. There is widespread agreement that this creates substantial uncertainty about how to accurately distribute overall DCE adjustments by rank in the income distribution (see our working paper, [Johns and Slemrod \(2010\)](#); [DeBacker et al. \(2020\)](#), and AS). Our main limitation is that we do not observe estimated auditor fixed effects for each return, so we do not apply different multipliers to different tax returns depending on the specific auditor who audited those returns. We dealt with the resulting uncertainty in our working paper by showing results with DCE following [Johns and Slemrod \(2010\)](#), but also making it plain exactly how DCE shaped the benchmark estimates and providing readers enough information to disregard DCE in examining all of the key aggregate statistics if they wished to do so. However, because official estimates of the tax gap include DCE, we opted to include our attempt at distributing DCE adjustments in the benchmark scenario.

Auten and Splinter assert that the uniform multiplier approach over-estimates evasion at the top. They rightly point out that taxpayers with large amounts of detected under-reporting will on average have been audited by more thorough auditors, so they should receive a smaller multiplier than the uniform multiplier method allocates to them. They present a three-person example in which this is the dominant bias with the uniform multiplier approach. However, this potentiality is far from the only source of bias in the uniform multiplier approach. Mainly, there is a countervailing source of bias from the fact that returns with low detected evasion receive a multiplier that is too small. Without more information, there is no compelling reason to expect that the upward biased DCE adjustments will dominate the downward-biased adjustments, as AS imply (without a formal proof). We noted this bi-directional ambiguity in our working paper, as did [Johns and Slemrod \(2010\)](#) and [DeBacker et al. \(2020\)](#).<sup>4</sup>

To understand better why the net direction of the bias is unclear even for estimates of the tax gap for the top 1%, consider a simple example in which there are two types of auditors: very thorough auditors who find 100% of evasion every time and more typical auditors who always find 20% of true evasion. Suppose, purely for illustrative purposes, that the latter group substantially outnumber the former: 90% of auditors are typical types and 10% are thorough types. In this situation, 28% of true population evasion ( $= 0.1 * 1 + 0.9 * 0.2$ ) is detected by auditors, implying a population multiplier of 3.6. However, the correct multiplier for a given taxpayer would be 1 in the case of a thorough auditor and 5 in the case of a typical auditor.

---

<sup>4</sup>In addition, any noncompliance undetected by the most thorough auditors will not be accounted for by DCE. Insofar as that systemic non-detection has a distributional bias, that bias will be compounded by the uniform multiplier approach. Consequently, application of the uniform multiplier approach will dampen the loading of DCE at the very top of the income distribution, where we show in this paper that much of the misreporting goes systemically undetected in the NRP. Empirically, this concern is what makes us suspicious of the sharp drop-off in estimated rates of under-reporting within the top 1% of the distribution in the DCE-adjusted estimates.

To build intuition, suppose further that we know each taxpayer's rank in the true income distribution. That is, we know *ex ante* which taxpayers were in the top 1% of the distribution. By assumed random assignment of auditors, we know that 28% of true evasion by the true top 1% of the distribution will be detected by examiners. Multiplying detected evasion in the true top 1% by 3.6 would therefore provide an unbiased estimate of the extent of total under-reporting in the top 1% by true income.

Of course in reality we do not know which taxpayers belong in the top 1% of the true income distribution and we must estimate this. The above logic implies that, maintaining the underlying assumptions under which DCE yields an unbiased estimate for the full population, problematic re-ranking is the sole source of bias here. But it is reasonable to believe in our example that accounting for bias from re-ranking *magnifies* the extent of downward bias from the uniform multiplier approach. To see why, note that with the uniform multiplier method, highly non-compliant individuals with typical auditors tend to be ranked too low in the distribution (creating downward bias in estimated non-compliance at the top), while those highly non-compliant individuals with thorough auditors will be ranked too high (creating upward bias). With a small share of thorough auditors, the downward bias will dominate the upward bias. The exact magnitude of both biases would of course depend on the joint distribution of income and non-compliance.

Our example illustrates just one reason the direction of the bias is unclear, which is that the distribution of auditor effects is unknown. Another potential reason, which is somewhat outside the scope of the original underlying model of DCE (Feinstein, 1991), is that the effect of auditor quality is plausibly heterogeneous through the income distribution, so that the auditor quality effect is not a true fixed effect throughout the distribution.<sup>5</sup> For example, if thorough auditors are better at detecting under-reporting of financial capital income or complex pass-through business income<sup>6</sup>, while most auditors are equally able to detect under-reporting in sole proprietorships, then DCE adjustments should disproportionately load onto the top of the income distribution, where financial capital and pass-through business income are prominent.

Our intent here is not to claim that the uniform multiplier approach does in fact lead to under-estimation of the tax gap or overall under-reporting of income at the top. We only seek to convey that the sign and magnitude of the bias resulting from the uniform multiplier approach depends on many more factors than the one possibility highlighted by AS, and we see no compelling *ex ante* reason to believe the approach is “inherently upward biased,” as AS assert.

Nevertheless, given all this uncertainty it is eminently reasonable to conduct some sensitivity analysis,

---

<sup>5</sup>Splitting on whether TPI > \$100,000 in the definition of groups in the method we use reflects this type of possibility, but one can imagine a number of ways to enrich the method along these lines.

<sup>6</sup>In practice very little pass-through evasion is found due to infrequent examination of pass-through business tax returns, as we discuss in Section 3.2 of the working paper. Rarely, auditors do audit pass-through business returns in the course of an NRP audit, and apart from this there are occasional large adjustments without a business level-audit, e.g. disallowances of passive losses due to anti-avoidance rules. So the substantive point here stands.

to see how feasible alternatives to the uniform multiplier approach affect the estimated tax gap at the top, which we turn to now.

**Auten and Splinter’s “Gradient” Method.** AS propose scaling detected non-compliance at the micro level by a gradient of multipliers based on the principle that those with large detected non-compliance should receive smaller multipliers. They implement this using varying multipliers ranging from 8 for those with very little exam-detected under-reporting to 1 for those with the most under-reporting (see their Table A2).<sup>7</sup> AS implement this approach using SOI data with imputed under-reporting based on un-released work in progress (Auten and Langetieg, 2020). Additionally, they implement this approach using overall incomes only: rather than adjust each line-item one-by-one as conventional DCE methods do, they assign multipliers based on the ratio of exam-corrected to reported income, and apply said multipliers to overall income under-reporting.

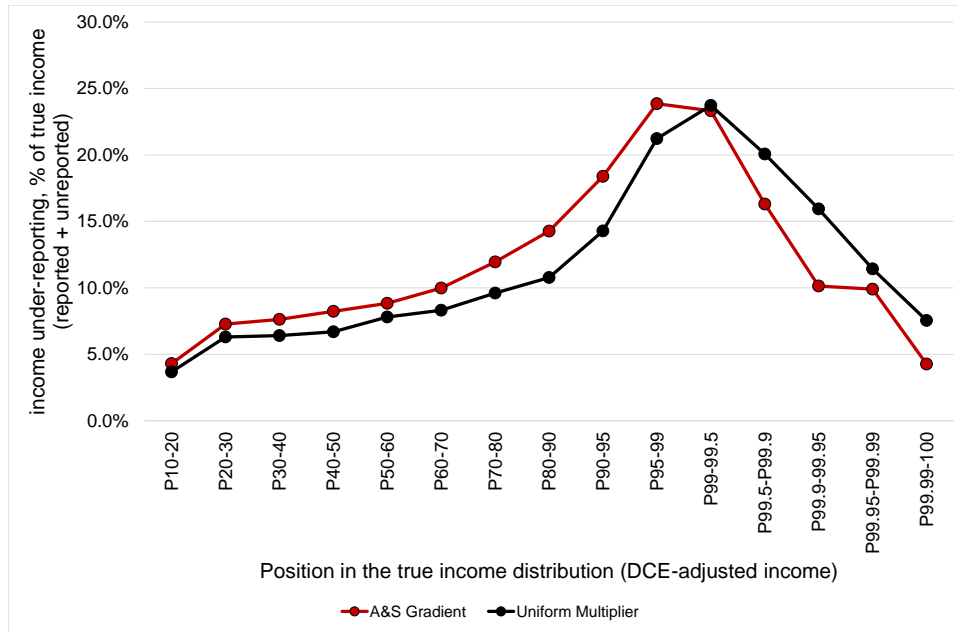
In Figure 1, we implement AS’ gradient-based multiplier approach on overall income with our micro data from the NRP, and we compare the results to the uniform multiplier estimates employed in our working paper. As AS suggest, this estimation method increases under-reporting in the bottom 95% and decreases under-reporting in the top 1%. We note, however, that the qualitative profile of evasion is similar to the uniform method: estimated under-reporting peaks at just over 20% of true income around the 95th to 99th percentile of the distribution, and then falls sharply within the top 1%. These estimates would imply a reduction in the top tax gap and the top 1%, but only insofar as we accept this steep drop-off in under-reporting within the top 1% - the broader implication of our findings being that one should not necessarily accept this drop-off as fact.

Nevertheless, setting aside the issue of what might be un-detected in the NRP data even by top auditors, Figure 1 does suggest that the uniform multiplier method may be upward biased with respect to evasion at the very top of the distribution. However, we are not content to stop at Figure 1 for two reasons. First, the specific set of multipliers AS used is ad hoc, without empirical justification. Second, they apply gradient-based multipliers to overall income rather than specific line items, which creates some conceptual difficulties. For example, it may be that a taxpayer for whom an auditor only found modest under-reporting

---

<sup>7</sup>The gradient we use here to replicate their findings comes from the second row of Table A2 in AS. AS also report results for a “steep gradient” and a “flat gradient.” If we adopt these multipliers, we find that the flat gradient leads to far less total under-reporting than the DCE-adjusted population total under-reporting, and the steep gradient leads to far more total under-reporting. When using these alternative gradients, AS then re-scale population non-compliance not to match DCE-adjusted population under-reporting from the NRP data, but rather a different total under-reporting based on a problematic interpretation of NIPA data, which we discuss further in Section 4. Moreover, setting aside the proper population total to re-scale toward, this type of re-scaling itself is problematic, because it mechanically modifies the multipliers applied to individuals’ exam-detected under-reporting in an opaque fashion. In other words, the multipliers actually used to scale exam-detected evasion are not those appearing in the third and fourth rows of Table A2 due to the additional re-scaling to match a lower total amount of under-reporting. Finally, they apparently do not account for the re-ranking effects implied by how re-scaling would modify the multiplier applicable to each individual, which would mechanically flatten the profile of evasion by rank in the income distribution. We therefore focus here on the main “Gradient” multipliers, for which the population total under-reporting is similar to the appropriate population total.

FIGURE 1: AUTEN & SPLINTER GRADIENT DCE VS UNIFORM MULTIPLIER



of wage income (e.g. unreported tips), by virtue of being assigned a large multiplier, is implicitly allocated substantial Schedule C evasion. In what follows, to examine these issues systematically, we develop a flexibly parameterized gradient-based method, and we apply this type of approach to overall income and to each line item separately.

**A Parameterized Gradient-Based DCE Method.** We should begin with a principled way to specify DCE adjustments as a function of exam-detected mis-reporting. In part because it has an attractive property which we shall derive shortly, we use the following functional form for the appropriate DCE adjustment as a function of exam-detected under-reporting:

$$\frac{y_{i,DCE} - y_{i,rep}}{y_{i,exam}} = \alpha_{yg} + \beta_{yg} \frac{y_{i,exam} - y_{i,rep}}{y_{i,exam}}, \quad (1)$$

where  $y_{i,rep}$ ,  $y_{i,exam}$ , and  $y_{i,DCE}$  are taxpayer  $i$ 's reported, exam-corrected, and DCE adjusted income, respectively, and  $\alpha_{yg}$  and  $\beta_{yg}$  are parameters that may be specific to the type of income being considered ( $y$ ) or the group of taxpayers ( $g$ ). When implementing on overall income, similar to AS and Figure 1, we use overall income for  $y$  and a single group of taxpayers. When implementing this approach line-by-line



we implement DCE adjustments separately for each type of income, with two groups of taxpayers based on whether the taxpayer reported Schedule C income or TPI>100k). Note also that equation (1) implicitly requires that the taxpayer is under-reporting:  $y_{i,exam} > y_{i,rep}$ . As ever, in the other case we use a multiplier of 1, i.e. we do no DCE adjustment.

In words, equation (1) implies that the rate of DCE-adjusted mis-reporting is linear in exam-detected mis-reporting when both are scaled by exam-corrected income. We further require that  $\alpha_{yg} \geq 0$  and  $\beta_{yg} \geq 0$ . It is straightforward to derive that the multiplier applied to individual  $i$ 's mis-reporting of income type  $y$ , i.e.

$$m_{iy} \equiv \frac{y_{i,DCE} - y_{i,rep}}{y_{i,exam} - y_{i,rep}} \text{ is} \quad m_{iy} = \beta_{yg} + \alpha_{yg} \frac{y_{i,exam}}{y_{i,exam} - y_{i,rep}} \quad (2)$$

Importantly, the last term in equation (2) is the inverse of the exam-detected rate of mis-reporting, thus  $\alpha_{yg} \geq 0$  implies that individuals with higher degree of exam-detected mis-reporting will receive smaller multipliers. This functional form therefore captures the key feature of a gradient-based approach. To discipline the values of  $\alpha_{yg}$  and  $\beta_{yg}$ , we impose that the population DCE multiplier must equal the appropriate population multiplier – when we do this line-by-line, this multiplier is one of the four multipliers listed above.

$$E[y_{i,DCE} - y_{i,rep} | i \in g, y_{i,exam} > y_{i,rep}] = M_{gy} E[y_{i,exam} - y_{i,rep} | i \in g, y_{i,exam} > y_{i,rep}] \quad (3)$$

Multiplying both sides of equation (1) by  $y_{i,exam}$ , taking expectations, and applying equation (3), we obtain the following restriction on  $\alpha_{yg}$  and  $\beta_{yg}$

$$M_{gy} = \alpha_{gy} \frac{E[y_{i,exam} | i \in g, y_{i,exam} > y_{i,rep}]}{E[y_{i,ex} - y_{i,rep} | i \in g, y_{i,exam} > y_{i,rep}]} + \beta_{gy} \quad (4)$$

Equation (4) allows us to reduce the parameter space to a single dimension that nests all possible implementations of equation 1 that keep population DCE totals fixed. With a little bit of work, we find that the functional form for DCE adjustments in equation (1) and the restriction in equation (3) are jointly equivalent to the following re-parameterization, for a single parameter  $p$ :

$$\frac{y_{i,DCE} - y_{i,rep}}{y_{i,exam}} = M_{gy} \left\{ p \underbrace{\frac{y_{i,exam} - y_{i,rep}}{y_{i,exam}}}_{\text{ind. evasion rate}} + (1-p) \underbrace{\frac{E[y_{i,exam} - y_{i,rep} | i \in g, y_{i,exam} > y_{i,rep}]}{E[y_{i,exam} | i \in g, y_{i,exam} > y_{i,rep}]}}_{\text{pop. evasion rate}} \right\}, \quad (5)$$

where  $\alpha_{yg} \geq 0$  and  $\beta_{yg} \geq 0$  imply that  $p \in [0, 1]$ .<sup>8</sup> Equation (5) has some appealing properties, because

<sup>8</sup>In theory we could use different values of  $p$  for different  $y$ 's or  $g$ 's, but in practice when implementing line-by-line we assume the same value of  $p$  for all types of income and groups of taxpayers, so that the only thing that is group or income-type specific is the

we can view the DCE adjustment applicable to a given taxpayer as a weighted sum of 1) exam-detected mis-reporting for that taxpayer, and 2) average exam-detected mis-reporting in the population, both scaled by the appropriate population DCE multiplier. Implicitly, then, the parameter  $p$  encodes the reliability of exam-detected mis-reporting as a signal of total mis-reporting, exam-detected and non-exam-detected, in a fashion that is at least aesthetically similar to an Empirical Bayes estimator (Efron and Morris, 1975).<sup>9</sup> When  $p = 1$ , or equivalently when  $\alpha_{yg} = 0$ , we obtain the uniform multiplier approach; when implementing this approach line-by-line we therefore obtain identical results to the uniform multiplier approach when  $p = 1$ .

**Results.** Panel 2a of Figure 2 shows our results from implementing this parameterized gradient-based DCE on overall income, for various values of  $p$ . Panel 2b of Figure 2 shows the same results implementing this approach line-by-line.

We observe that as in Figure 1 and AS, implementing this approach to DCE on overall income tends to decrease the estimated rate of evasion in the top 1%. The numbers AS used in their gradient specification resemble our results with perhaps  $p \approx 0.8$ . At relatively small values of  $p$ , a re-ranking effect causes a reversal of the decline in evasion rates within the top 1% at the very top of the distribution, but at such low values of  $p$  the multipliers applied to those with very little exam-detected under-reporting reach somewhat extreme levels so we do not suggest taking these values very seriously.

In sharp contrast, when implementing this approach line-by-line, we observe in Panel 2b of Figure 2 that the implied bias from the uniform multiplier approach goes in the opposite direction. For  $p < 0.75$  or perhaps slightly lower, these results would imply that the sources of upward and downward bias in the uniform multiplier approach are almost exactly offsetting, while for smaller values of  $p$  the uniform multiplier approach leads to a downward-biased estimate of top 1% evasion. This occurs for two reasons, which are related to our discussion above. First, individuals at the top of the distribution with relatively little detected under-reporting are allocated larger adjustments of some types of income, and second, the use of line-by-line adjustments allocates under-reporting for a given type of income to relatively fewer individuals, which strengthens the re-ranking effect. For example, Schedule C DCE adjustments are only applied to those found under-reporting Schedule C income by examiners, rather than potentially being applied to any taxpayer found under-reporting overall income in Panel 2a. This implies a stronger re-ranking effect (see also the discussion around Figure 4 below). As before, we caution readers of making too much of the estimates for very low values of  $p$  due to the extreme multipliers implied by such an approach for some taxpayers.

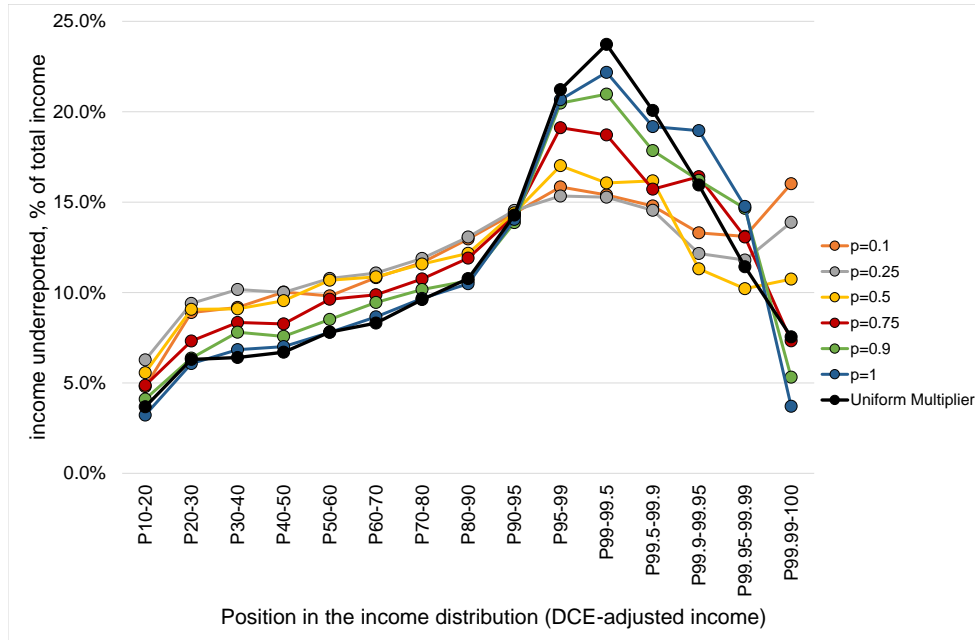
---

population multiplier  $M_{gy}$ .

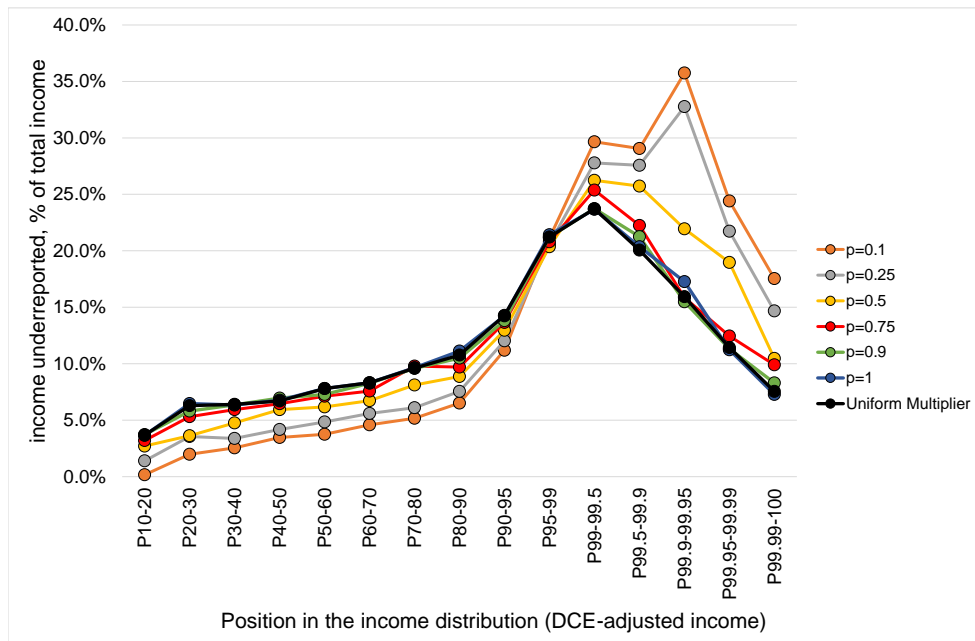
<sup>9</sup>We conjecture that the resemblance of equation (5) to an Empirical Bayes estimator is not merely aesthetic, but we have not developed a deeper theory of DCE adjustments that admits an Empirical Bayes type interpretation of the parameter  $p$ .

FIGURE 2: PARAMETERIZED GRADIENT-BASED DCE METHOD

(a) Overall income



(b) Line-by-line



In summary, the bias in the uniform multiplier approach applied by this sensitivity analysis tends in opposite directions depending on whether the adjustment is applied to overall income or line-by-line. But the method that seems to us closest in spirit to the underlying model and to the manner in which DCE adjustments are computed in official statistics – a line-by-line adjustment for two distinct groups of taxpayers – suggests that the uniform multiplier approach implies negligible or slight downward bias in estimated top 1% under-reporting.

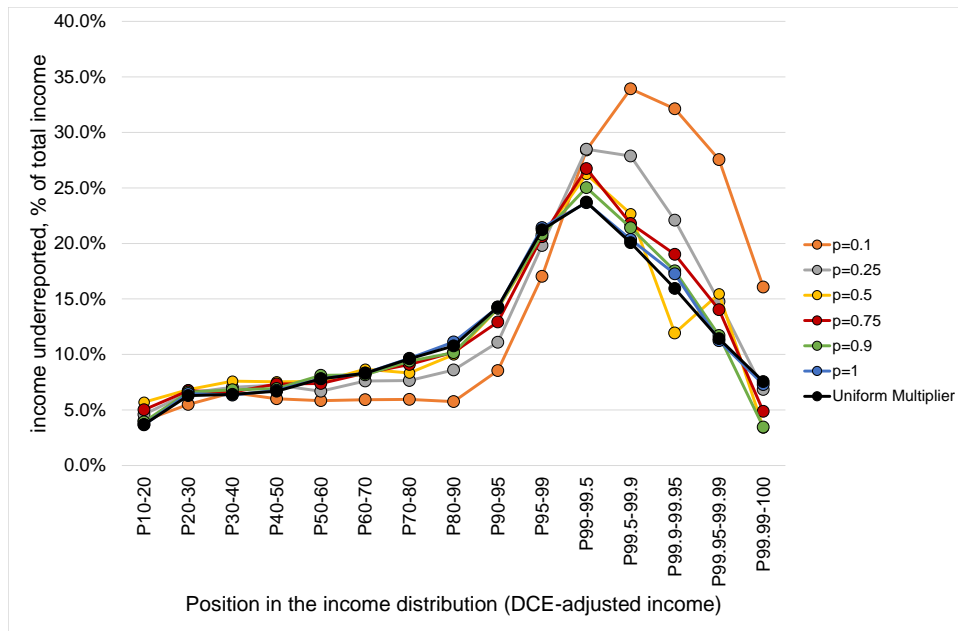
Finally, we report the results of one simpler sensitivity check on the distribution of DCE adjustments, which is motivated by the primary concern expressed by AS and DeBacker et al. (2020) with the uniform multiplier approach: individuals with very large exam-detected evasion are allocated too much un-detected evasion. To address this specific concern in a simpler fashion, we apply a multiplier of 1 to large adjustments (i.e. no DCE adjustment) and scale up the multiplier for others to keep total DCE adjustments fixed. Specifically, we rank taxpayers by their exam-detected mis-reporting rate  $E_{iy} \equiv \frac{y_{i,exam} - y_{i,rep}}{y_{i,exam}}$ . We again specify a parameter  $p \in [0, 1]$  where  $p = 1$  will nest the uniform multiplier method. But here we assign a multiplier of 1 to the fraction  $1 - p$  of taxpayers with the largest values of  $E_{iy}$ , and to everyone else we assign a constant multiplier that ensures that total mis-reporting for income type  $y$  in both groups  $g$  satisfies equation (3).<sup>10</sup> Thus a fraction  $1 - p$  of the population with the largest exam-detected rates of evasion for a particular type of income get a multiplier of 1, and the remaining  $p$  fraction get a multiplier larger than the uniform multiplier approach would use. The former group will be allocated much less evasion than the uniform multiplier method would suggest, while the opposite is true for the latter group. We implement this approach line-by-line for the same two groups of taxpayers as above.

The results of this final, simple sensitivity check are in Figure 3. Just like in Panel 2b of Figure 2, these results suggest that the uniform multiplier approach represents a *downward-biased* estimate of top 1% evasion. Even when explicitly targeting the main concern expressed by AS and others for a source of upward bias with the uniform multiplier approach, the fact that individuals with less exam-detected under-reporting must be allocated larger DCE adjustments to keep population totals fixed, i.e. that we must account for sources of downward bias as well, dominates the potential source of upward bias here. Within the top 1%, however, we observe that the estimate for the top 0.01% decreases by a few percentage points for some values of  $p$ , while evasion in the rest of the top 1% increases. Thus in the very top bin there may be scope for some upward bias in the uniform multiplier method through the channel AS describe, but the source of downward bias dominates for the rest of the top 1% in these scenarios.

---

<sup>10</sup>Note that this approach implicitly assumes  $y_{i,exam} > 0$ , so that  $E_{iy} > 0$  for any under-reporting taxpayer. For simplicity, we assign a multiplier of 1 to situations where  $y_{i,exam} < 0$ . The parameter  $p$  represents the fraction of the population with positive  $y_{i,exam}$  who are assigned the larger multiplier. We find that virtually any alternative way of handling the negative values of  $y_{i,exam}$  (e.g. using absolute values for negatives, using the difference in the inverse hyperbolic sine of exam and reported income for  $E_{iy}$ ) leads to very similar results.

FIGURE 3: TWO-STEP GRADIENT-BASED DCE



**Summary.** The results of our sensitivity analysis underscore our central claim here: there is no compelling reason to expect that the uniform multiplier approach must be downward biased with respect to the rate of evasion at the top of the distribution. We reject the assertion by AS that our use of this approach created “inherent upward bias” in our benchmark scenario (or other scenarios incorporating DCE). We conclude that despite significant uncertainty and multiple sources of potential bias here, our use of the uniform multiplier method is a reasonable and simple means of allocating DCE adjustments through the income distribution.

One additional limitation that bears mentioning is that, in all the approaches to distributional DCE applied by ourselves and by AS, DCE adjustments are only allocated to those found under-reporting by examiners. In reality, some un-detected evasion may be more appropriately allocated to those found to over-report income by examiners or those with no adjustments by examiners. How accounting for this possibility would affect the estimated concentration of evasion is ambiguous for similar reasons as those discussed above. We note that at least for a sizable share of total under-reporting, this issue is not quite so restrictive as one might guess. The majority of exam-detected under-reporting is under-reporting of Schedule C income, and a large share of those with Schedule C income (over 65%) are in fact found to under-report that income, suggesting that it is appropriate to allocate a large amount of DCE adjustments to Schedule C income to those found under-reporting in exams. We also note that according to documentation in official

tax gap studies, recent innovations to DCE have attempted to address this issue, and to simulate micro-level DCE in a more systematic fashion (IRS, 2019). At present we have been unable to implement these newer methods with what is available to us, but we are exploring our options and may implement some newer methods in further revisions of our paper.

Finally, we emphasize that neither the uniform multiplier method nor any of the alternatives discussed above are definitive, perfect, or clearly unbiased for distributional questions. We hope that future research will be able to derive from first principles a distributional approach to Detection Controlled Estimation, which accounts for the importance of re-ranking, heterogeneity in exam-detected and undetected evasion and in the composition income, and the plausibly heterogeneous effect of auditor quality through the distribution. This is a challenging and important methodological question.

### 3 Quantifying Pass-Through Evasion

**Review of Findings from Random Audit Data.** An essential motivating fact for our analysis of NRP random audits is that when an auditor conducting a random audit of an individual tax return encounters pass-through business income (from S corporations or partnerships), the auditor only proceeds to audit the pass-through business return (the Form 1120-S or Form 1065) in *less than 4 percent of cases*. Consequently, pass-through business evasion is substantially under-measured in random audit data. Before DCE, just 4.6% of pass-through business income is estimated to be under-reported in the random audit data, of which 2.7% (58% of total estimated pass-through evasion) comes from the 4% of cases where business-level returns were examined. This low rate of under-reporting compares to a 37% rate before DCE for sole proprietorships. In our benchmark scenario, we therefore proposed adjustments to overall estimated under-reporting and the tax gap to account for missing pass-through evasion. Because pass-through income is highly concentrated at the top of the income distribution, most of these adjustments plausibly belong at the top of the income distribution.

Because we have almost no data on pass-through evasion, we make three assumptions to do these adjustments quantitatively. First, we need to determine just how much pass-through evasion to include, i.e. what fraction of true pass-through income is under-reported. Second, we need to distribute these adjustments in some way by rank in the (true) income distribution. Third, we need to adjust for double-counting between random audit estimates and our assumed overall levels of pass-through evasion. AS argue that the assumptions we made each step of the way led us to over-state pass-through evasion in the top 1% of the distribution.

**The Rate of Pass-Through Evasion.** The first question we confront is the rate of under-reporting for pass-through business income. Based on various estimates of rates of evasion for different types of business income, we chose to assume that 20% of pass-through business income was under-reported. We arrived at the 20% figure from the following: the sole prop under-reporting rate before DCE is 37%; IRS (2016) puts the C corp tax gap at 18% of true tax; a 2003-2004 random audit study by Johns (2009) estimates a net mis-reporting rate of 12 to 16% of true income without any DCE-type adjustments. We opted for a rate above that of C corps but well below sole proprietorships. We reduced this rate from 20% to 12% (and increased it to 28%) in sensitivity checks.

AS point out that more recent estimates suggest a slightly lower rate for C corporations of perhaps 15%, and combined with the findings of the S corp study above they suggest that a rate of 15% is more appropriate. We note that this is within the scope of the existing sensitivity analysis, but continue to believe that 20% was already reasonably conservative, for several reasons.

Most importantly, on inspection the figures from Johns (2009) we quoted in the paper from the S corp random audit data correspond to a slightly different quantity than the quantity of interest for our purposes.<sup>11</sup> These figures were estimating:

$$\frac{\text{total under-reporting of net business income (net of over-reporting)}}{\text{total abs(true net business income)}},$$

while the under-reporting rate for our purposes should include true total S corp *net business income* in the denominator,<sup>12</sup> i.e.

$$\frac{\text{total under-reporting of net business income (net of over-reporting)}}{\text{total true net business income}}.$$

Based on our own analysis of the data from this random audit program, we replicated the totals in Johns (2009) precisely, and then estimated the revised parameter of interest for our analysis.<sup>13</sup> Accounting for this issue increases the estimated under-reporting rate to 19% of true net business income.

Two additional factors suggest that our 20% under-reporting rate is quite conservative. First, these esti-

<sup>11</sup>Presentation slides on this work are available at: <https://www.irs.gov/pub/irs-soi/09resconawardscorp.pdf>.

<sup>12</sup>Official tax gap statistics often use absolute values of negative amounts, as is done here. This is a fine way to account for negative dollar amounts, but it was not the right empirical quantity for our purposes here.

<sup>13</sup>One issue noted by Johns (2009) is that there are substantial corrections that increase deductions for owner compensation in the S corp random audits, due to reasonable owner compensation rules to combat SECA tax avoidance. Corrections to the deduction for owner compensation will be offset by increases in wage income on individual owners' tax returns, so including increases to these deductions in an estimate of the mis-reporting rate under-states the true amount of under-reporting of net business income. Our preferred 19% figure excludes corrections to owner compensation deductions, which increases the mis-reporting rate by about 4 percentage points. We agree with Johns (2009) that we should exclude corrections to owner compensation in the numerator of the mis-reporting rate, because these constitute line switching on the owners' individual tax return. Whether we ought to exclude these corrections in the denominator is less clear, but it makes little difference: the estimated mis-reporting rate is 18.5% when excluding these corrections in the denominator and 19.4% when including them.

mates based on the S corp random audits did not include any DCE-type adjustment for undetected evasion. Even a modest adjustment, let alone the multipliers in excess of 2 suggested by individual tax gap estimates, would imply a mis-reporting rate well in excess of 20%. Second, while we do not have similar random audit data on partnerships, suggestive evidence in [Cooper et al. \(2016\)](#) and conversations with a number of practitioners lead us to believe that partnerships are less compliant overall than S corporations.<sup>14</sup> Thus, if 19% is a conservative under-reporting rate for S corporations, which earn about two thirds of S corporation plus partnership business income, then 20% is conservative for both S corporations and partnerships combined.

**The Distribution of Pass-Through Evasion.** Given a total amount of pass-through evasion in dollars of under-reported income, we next turn to how to distribute this total through the income distribution. Our basic approach, detailed in Section 3.2.2 of the paper, is to assume that un-reported pass-through income is distributed like reported pass-through income. For example, if the top 1% by *reported* income earns 70% of *reported* pass-through business income, we assume that the top 1% of the *true* income distribution earns 70% of *unreported* pass-through business income.

Contrary to how AS describe this approach in their critique, none of the above implies that taxpayers with high reported incomes are especially non-compliant. As we note elsewhere in the paper, reported incomes are positively selected on compliance: more compliant taxpayers will have higher reported incomes, holding all else fixed. Due to re-ranking effects, however, which are plausibly substantial given the dollar amounts involved, assuming unreported pass-through income is distributed like reported pass-through income is different from allocating unreported pass-through income to specific individuals in proportion to their reported pass-through income. The latter would be biased, but it is simply not our method. Thus we reject AS' assertion that there is any mechanical bias in our approach from the fact that reported incomes are positively selected on compliance.

**Reported Losses.** Ambiguities with our basic approach to distributing pass-through under-reporting do arise because the bottom 10% of the income distribution has negative total net pass-through income. We opted to assign zero pass-through under-reporting to the bottom 10%, rather than a negative amount of under-reporting (i.e. over-reporting). We allocate the macro total for pass-through under-reporting among the remaining 90% of the distribution, assuming this is distributed like reported net business income, which is positive in aggregate in each bin in the top 90%. Economically, the assumption here is that after correcting under-reporting, under-reporters with reported pass-through losses will have positive corrected true

---

<sup>14</sup>Our conversations with practitioners also led us to believe that there is likely substantial heterogeneity in the character of partnership under-reporting by the type of partnership, and much more ambiguity in the lawfulness of some tax positions adopted via partnership structures. We have limited data to speak to these factors empirically and hope that future research can quantify them.



income, i.e. that re-ranking pulls these taxpayers out of the bottom 10%.

We argue that this is a reasonable means of dealing with negative business incomes. Economically, there are two potential types of pass-through under-reporting for taxpayers with reported pass-through business losses. First, the taxpayer claims a large loss when true business income is zero or positive. For example, auditors often disallow large losses entirely due to one of several anti-avoidance rules. Second, the taxpayer claims slightly larger losses when true business income is still a loss. For example, a taxpayer owning a business with true losses in a given year might claim some non-eligible expenses. If the first type is the predominant form of pass-through under-reporting, simulations suggest our approach is reasonable. We showed that taxpayers with reported business losses are generally re-ranked to the top of the income distribution if those losses are randomly replaced by zero to simulate this form of non-compliance (see Figure A12 of the paper).

The second type, however, we assume to be negligible for the purposes of distributing pass-through non-compliance. The scope for bias from ignoring this type is mechanically very small. To understand why, first consider that aggregate losses for those with negative incomes are about 10% of total net pass-through business income in non-recession years. The total amount of the second type of non-compliance must be mechanically smaller than 10% of total net pass-through income, likely much smaller. For illustration, suppose losses were over-stated by 20% by all tax-payers with reported losses (and the first type of under-reporter is nonexistent). Then about 2% of total pass-through business income is non-compliance attributable to these types of taxpayers, or 5% of our assumed total pass-through under-reporting. In reality, the share of under-reporting that belongs at the bottom by true income is likely even smaller. For instance, in NRP data on sole proprietorships (before DCE), we estimate that 14% of all sole proprietorship under-reporting is attributable to taxpayers with reported losses, but just 1.2% of under-reporting is attributable to those for whom exam-corrected income is a loss. All this suggests that a negligible amount of total under-reporting would be mis-allocated through this channel.<sup>15</sup> We view this as well within our margin for error, given the lack of an obvious better way to handle the negative income issue and the other, bigger sources of uncertainty here.

---

<sup>15</sup>Re-ranking effects, especially for taxpayers with large disallowed losses, also lower the implied marginal tax rate on under-reported income in our results. AS note that the implied marginal tax rate at the very top of the distribution is sometimes low in our Tables. Part of this stems from this re-ranking issue, part from the low rate on long-term capital gains and qualified dividends, and part from a conservative specification of marginal tax rates to map DCE income adjustments into tax liabilities, which we employed in the working paper when doing DCE adjustments. Specifically, we applied the marginal tax rate applicable to exam-corrected income (including preferred rates for long-term capital gains and qualified dividends) to the micro DCE adjustments to income under-reporting. Using a more precise tax calculation for micro-level DCE adjustments would increase the estimated tax gap at the top without affecting estimated income under-reporting, but the increase in the tax gap would be solely due to the re-allocation of DCE adjustments to higher tax brackets. We opted to be conservative on this point in the paper. This choice makes little difference for most of the distribution, but it decreases the average MTR in the very top bins by a few percentage points. We discuss some additional issues related to taxpayers with reported losses in Section 5 below.

**Double Counting.** Even though it came from audits of just 4% of owners, detected business-level under-reporting did account for 58% of the (relatively little) pass-through non-compliance that was detected by examiners in the full population estimates. Including this business-level under-reporting from NRP-based estimates and our pass-through adjustments together would create double counting. Moreover, the magnitude of the potential double counting multiplies three-fold when we also incorporate DCE adjustments. When we combine our pass-through under-reporting adjustments with NRP estimates, we therefore remove 58% of NRP pass-through under-reporting from NRP estimates, with or without DCE. The net effect of the pass-through adjustment plus the double-counting adjustment is to allocate a larger share of under-reporting to the top 1% than if we performed the pass-through adjustments without the double counting adjustment. This fact is the sole cause of the phenomenon documented in Table 3 of AS. AS contend that we should respect the distributional information in the NRP-DCE estimates, and therefore they characterize what we have done here as allocating too much under-reporting to the top 1%. We agree that we discard the distributional information in the NRP-DCE estimates, but in light of everything discussed above, especially the uncertainty around DCE and re-ranking, we view this as a reasonable approach.

We are skeptical of the alternative approach suggested by AS, which is to somehow use the individual NRP data to empirically discipline the distribution of pass-through non-compliance overall. The reason this approach is not credible is the same reason we incorporated pass-through adjustments in the first place: only 4% of returns with pass-through business income are examined at the business level. There is every reason to believe that this 4% is a highly selected sub-sample, where the selection process could involve 1) the simplicity of the business return, 2) whether the audited 1040 taxpayer was the sole owner of the business, 3) whether the auditor has access to the records of the business, 4) anything suspicious or incidental that motivated the auditor to audit the business return, and 5) the thoroughness of the auditor. With all these potential margins for selection and just a 4% sub-sample, and all the uncertainty created by DCE adjustment on top of this, there is no credible method to extract distributional estimates of pass-through under-reporting from these data.

**Summary.** As with DCE adjustments, we acknowledge that there is uncertainty about the extent and distribution of pass-through under-reporting. In the absence of further evidence, the method we used in our benchmark estimates, a conservative rate of under-reporting and an allocation based on the concentration of reported pass-through business income, seems reasonable. But we welcome further research on this question, which could be used to refine our estimates. For our part, we agree with AS that it would be particularly useful to empirically discipline where in the corrected income distribution one ought to distribute pass-through under-reporting. Our ongoing work attempts to speak to this question using data from the

2003-2004 S corp random audit study. We are analyzing these data now, and may use them to refine our estimates in future revisions of the working paper.

## 4 Minor Issues

In this section, we respond to some more minor points raised by Auten and Splinter in their comment. Each of these is unrelated to the others.

**National Income Comparisons.** In their critique of our application of DCE, AS state that our estimates from the NRP with DCE “appears to exceed amounts in national income by about one half.” They suggest that this implies we must be over-estimating evasion.

Most importantly, as we noted in the paper, the estimate of the total tax gap for individual income tax filers that we obtain from the random audit data (before adjustments for sophisticated evasion) matches official estimates, and reported-income aggregates in our data match SOI tables on reported incomes. We sought to match what prior studies of the tax gap had done in our benchmark scenario. Contrary to other work in [Piketty et al. \(2018\)](#) and [Auten and Splinter \(2019\)](#), we are not estimating the distribution of national income and therefore do not calibrate our estimates to national income.

Nevertheless, it is useful to better understand what is going on here. First, NIPA statistics make explicit adjustments for under-reporting on tax forms for just two types of income: wage income and non-farm “proprietor income,” which corresponds to sole proprietor and partnership income. For these two, NIPA includes an adjustment for under-reporting based on IRS tax gap statistics. For proprietor income, in 2012, NIPA uses \$561 billion in under-reporting, while our data suggest \$630 billion (\$592 billion for sole proprietorships plus about \$40 billion for partnerships). For wage income, NIPA uses \$75 billion in under-reporting and we estimate \$53 billion. These totals could be slightly different for a few reasons: the wage difference is almost surely due to additional wages included in NIPA for non-filers, while other differences could arise due to the use of different waves of NRP data or differences in DCE methods. But plainly these two income categories are not the culprit for the large “discrepancy” described by AS.

The difference between our total and the one AS derive from NIPA statistics must therefore be due to other components of income besides proprietor and wage income. Here, however, the comparison to NIPA data is far less straightforward than AS assume. For these other types of income, NIPA does not draw on individual tax return (Form 1040) data, and its definitions of income can be different from what would appear on the 1040.<sup>16</sup> Differences between NIPA totals and 1040 totals here reflect not simply under-

---

<sup>16</sup>To be specific: rental income, farm income, S-corp income, and even partnership income in NIPA are not estimated using 1040 data – they are all estimated based on other data sources (residential fixed assets and market rents, arable land, 1120S + adjustments for

reporting of income on tax forms, but the different data sources and differences between the definition of taxable income and NIPA's definition of economic income. For S corp income, for example, 1040 income is net of Section 179 depreciation, but NIPA income is gross of Section 179 (because Section 179 bonus depreciation is more generous than economic depreciation). Attributing the difference between 1040 totals and NIPA totals for these types of income to tax evasion, which AS do in their comment and in [Auten and Splinter \(2019\)](#), is therefore problematic. It is not clear that the NIPA totals do include under-reported income for some or all of these types of income, nor is it clear what portion of the difference between NIPA and 1040 totals reflect differences in definitions rather than under-reporting.<sup>17</sup>

That NIPA aggregates in national income may imply less under-reporting than official tax gap statistics (and our estimates) raises some interesting questions for future work. However, we see little reason to revise our estimate of total under-reporting dramatically downwards on account of this issue. Doing so would bring our estimates in conflict with official estimates of the individual income tax gap, while a central objective of the paper is to start from official estimates and then propose an adjustment for sophisticated evasion.

**On Offshore Evasion.** We state in the paper that our estimates for offshore wealth reflect a time before the enforcement that started in 2008 and culminated in FATCA (see [Johannesen et al. \(2020\)](#) for a description of this enforcement). Mapping the offshore estimates to today's policy environment would require knowledge of 1) how much enforcement curbed offshore evasion, and 2) the extent to which those who previously evaded via offshore wealth substituted toward other forms of evasion. We do not have good data on either of these questions. As a result, we chose to try to simply give readers an appreciation for how much this might matter in sensitivity analysis (see e.g. Figure 5b/A6, Figure A17, and the surrounding discussions). The changes AS propose to account for this issue are more modest than those we illustrated in the existing sensitivity analysis.<sup>18</sup>

**Re-ranking and Reported Losses.** AS point out that a substantial amount of under-reporting is attributable to taxpayers with reported losses, i.e. negative reported incomes. In our figures on estimated total under-reporting as a share of total corrected income by rank in the corrected income distribution, we omit economic income, 1065S + adjustments for economic income, respectively).

<sup>17</sup>We also note that a non-trivial share of estimated under-reporting belongs to categories of income that do not belong in national income, like net operating loss carry-forwards.

<sup>18</sup>The findings in [Johannesen et al. \(2020\)](#) suggest perhaps 10% of offshore wealth was brought into compliance by early enforcement. Relatedly, AS suggest decreasing the fraction of wealth concealed to 80% of the total rather than 95%, based on some discussion of this issue in [Zucman \(2015\)](#). A change of roughly this magnitude would result in a modest adjustment of the totals, much more modest than the sensitivity analysis already done in the paper. In a footnote, AS also raise the question of whether wealth held by tax-exempt non-profits could contribute to our estimated total offshore wealth in havens. Non-profit organizations with endowments, like universities, do make substantial offshore investments. But because they report on their assets directly to the government, these non-profits do not contribute to the assets-liabilities gap used to estimate total offshore wealth in tax havens in [Zucman \(2013\)](#) and [Alstadaeter et al. \(2018\)](#), so they do not contribute to our estimate.

the bottom 10% because the denominator of the statistic of interest is negative in this part of the distribution, making it difficult to interpret. We should clarify that while much non-compliance is indeed attributable to those with reported losses, omitting the bottom 10% from most of our figures omits a negligible share of total under-reporting due to re-ranking effects. As we alluded to above, the vast majority of under-reporting taxpayers with reported losses end up re-ranked higher in the distribution, i.e. they have positive corrected income (see also Figure A3 of the paper). Such individuals' non-compliance is therefore included in our key figures even excluding the bottom bin. When we report the shares of total under-reporting, total income, or total tax by (corrected) income rank, we do include the bottom 10%, allowing us to see how much this matters. Depending on the specification, we observe that the bottom 10% by corrected income accounts for 0.7 to 1.6 percent of all under-reporting (see Table 2 of the paper), and an even smaller share of taxes paid or unpaid (See Table A7 of the paper).

Having responded to the specific issues raised by AS, we next conclude by revisiting the bigger picture.

## 5 The Bigger Picture.

As discussed in the introduction, measuring the full extent of tax evasion is difficult. On this point seemingly everyone agrees. A lack of perfect data capturing all forms of evasion means that we must confront some uncertainty in estimates of the tax gap. In our view, there are a number of sources of uncertainty when it comes to evasion at the top of the distribution specifically; some are mentioned by AS and some are not. Our objective in constructing our benchmark estimates was to convey how accounting for un-detected evasion would affect macroeconomic statistics. We believe the decisions we made led to a reasonable and conservative set of estimates. We welcome criticism of these assumptions and attempts to refine them, and we wish to be as transparent as possible in order to facilitate this.

Here, in our view, are the main issues creating uncertainty in the benchmark estimates:

1. There may be forms of under-detected sophisticated evasion other than the two we include – i.e. other than pass-through businesses and offshore finance. For example, we do not directly account for a number of known schemes involving charitable contributions.<sup>19</sup>
2. The rate of evasion for pass-through business income is uncertain. (see Section 3 above.)
3. The rate of evasion for financial capital income flowing through pass-throughs is also uncertain. (see Section 3.2 of the paper).

---

<sup>19</sup>We note that some of these issues would only matter for under-payment of tax and not under-reporting of income, as e.g. charitable contribution deductions are not a component of income.

4. The concentration of pass-through evasion at the top is uncertain. Partnership complexity and control over pass-through businesses is rising with income at the very top, which may increase opportunities for evasion.
5. In mapping these estimates to today's environment, we ignore any behavioral response to recent, large drops in audits of wealthy individuals & pass-throughs. The magnitude of such a response is not well-understood but it is probably not negligible.
6. The size of total offshore evasion, and how much recent enforcement has curbed it, are uncertain. This point we tried to address in sensitivity analysis in the paper, see Figure 5b/A6, Figure A17, and the surrounding discussions.
7. The proper allocation of DCE adjustments is uncertain. Here much uncertainty remains, but we argue based on Section 2 that the approach in our working paper was reasonable.

Roughly speaking, on 1 through 5 we take in our view a rather conservative stance. On 6 and 7, we acknowledge more uncertainty, and in our paper we try to be reasonable and transparent about how these things matter. Still, on all of the above, more research is needed. We hope the findings of our paper will motivate research along these lines.

**What about theory?** In the face of such data limitations, perhaps a return to theoretical reasoning would help us make progress. What would a theory of tax evasion predict about the relationship between the share of income under-reported and true income?

In the classic Allingham and Sandmo model, the answer to this question turns on risk preferences, specifically on relative risk aversion. Under constant relative risk aversion, the share of true income under-reported is constant over true income. Evading tax entails a (financial) risk in this model, and the taxpayer's appetite for financial risks determines what fraction of income they would like to put at risk. Under constant relative risk aversion, that fraction does not change as the taxpayer gets richer. If anything, we would expect that relative risk aversion is decreasing at very high incomes, in which case the share of income under-reported would increase with true income. Intuitively, in this case taxpayers' appetite for risk grows as they get richer, so they take more risks. We show in the paper that in this framework, high-income individuals will also tend to adopt concealment strategies to lower their probability of detection, which would enable higher quantities of under-reporting at higher incomes. Naturally, all this presumes that main the down-side risk of evasion entails a financial loss and not prison time. But this is realistic for the vast majority of cases (see [IRS \(2020\)](#)).

One important limitation of the Allingham and Sandmo framework, however, is the absence of an accounting for third-party information (Kleven et al., 2011). In models with third-party information reporting like that of Kleven et al. (2011), the share of income evaded depends not only on risk preferences, but also on the share of income reported by third parties to the IRS. To understand how this might matter, we estimate the share of income subject to information reporting to varying degrees by rank in the income distribution. We use the same categories as in IRS (2019), which documents, somewhat famously, that evasion increases monotonically as we move from "substantial information and withholding" to "little or no information."

We observe in Figure 4 that the extent of information reporting changes sharply at the top percentile of the income distribution.<sup>20</sup> This finding derives from the change in the composition of income away from wage and salary income ("Substantial information and withholding") and sole proprietor income ("Little or no information") and toward financial capital income and pass-through business income (both mostly in "some information"). About 80% of income in the bottom 95% of the distribution is wage and salary income, which entails little scope for under-reporting.<sup>21</sup> At the top of the distribution, third-party information may not be completely absent, but at best it provides an incomplete account of taxpayers' income. This creates opportunities for mis-reporting.

In summary, canonical theory suggests that there is plausibly an appetite for the risky under-reporting of income at the top of the distribution. For types of income that are concentrated at the top, the tax authority has limited capacity to deter under-reporting via third-party information reporting alone. Finally, at least recently, audit rates at the top of the distribution are very low. In short, we find a number of reasons theory would predict that the rate of under-reporting increases with income and we find fewer reasons it might decrease, especially in recent years.

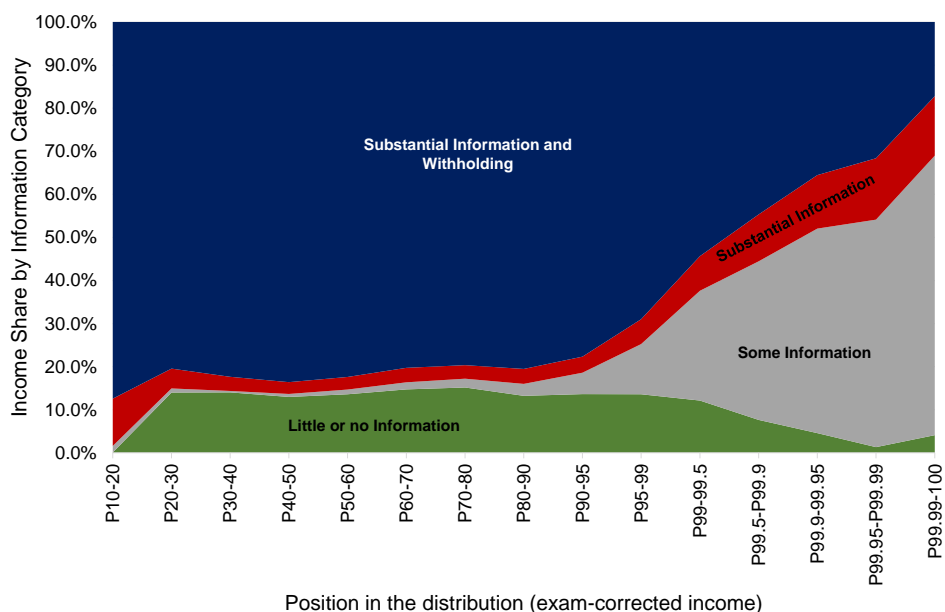
**The Measurement of Top Income Shares.** Both our paper and the comment by Auten and Splinter discuss the implications of these results for top fiscal income shares. If under-reporting as a share of income is larger at the top, accounting for under-reporting will tend to increase top 1% income shares. We show in the paper that in our preferred estimates, accounting for under-reporting increases top 1% fiscal shares by perhaps 0.9 p.p. (with DCE) to 1.2 p.p. (without DCE). AS point out that with much more conservative assumptions about the concentration of evasion imply a smaller increase in top 1% fiscal income shares. We have difficulty rationalizing a decrease. We have discussed above why we disagree with some of these assumptions, but from a bigger-picture perspective, the debate around this question concerns not the level

---

<sup>20</sup>We plot this figure by exam-corrected income in the NRP data for simplicity, but the figure is nearly identical if we plot these quantities conditional on reported or even DCE-adjusted income.

<sup>21</sup>This is closely related to the strength of re-ranking effects discussed above. With a large share of individuals in the bottom 95% having near-zero capacity for evasion because their income is third-party reported, evasion will be concentrated among a relatively small share of individuals. Re-ranking then tends to move this evasion upward significantly in the distribution, especially in DCE-inclusive estimates where the extent of overall under-reporting is large.

FIGURE 4: INCOME SHARES BY INFORMATION TYPE



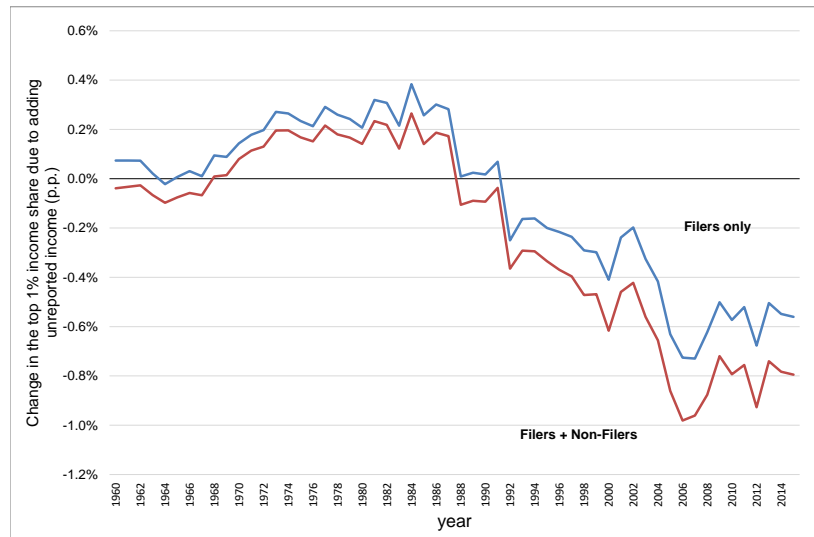
but the dynamics of top income shares. Piketty et al. (2018) and Auten and Splinter (2019) arrive at divergent trends in estimated top 1% income shares over time, with top 1% shares increasing by about 18 pp from 1980 to 2014 in Piketty et al. (2018) and only increasing by 5 pp in Auten and Splinter (2019). Part of the divergence is driven by the decision by Auten and Splinter (2019) to distribute under-reported income using an allocation based on random audit data.

Figure 5 shows how accounting for under-reporting in Auten and Splinter (2019) affects the level of the top 1% share in our sample period of 2006-2013, and the trend in top 1% shares over time. We observe that this method of accounting for under-reporting causes top 1% income shares to decrease a full percentage point from the 1980s to 2014. It is unclear what economic forces are behind this pattern. If these numbers are correct, something must have caused the 1% to become more compliant, or the bottom 99% to become less compliant over time. A number of forces would seem to work in the opposite direction. For example, the main change to the composition of income in the top 1% has been the monumental increase in the importance of pass-through income over time (Smith et al., 2019). Given that these estimates rely on random audit data whose limitations we have established, it is natural to wonder if some measurement issue might be driving these dynamics as well as the negative changes in top 1% shares in the later years.<sup>22</sup> But here, as

<sup>22</sup>One potential measurement issue involves the transition from early, line-by-line random audits done under the Taxpayer Compli-



FIGURE 5: IMPACT OF UNREPORTED INCOME ON TOP 1% INCOME SHARE



elsewhere, much remains to be studied.

**Conclusion.** We thank AS for their critical engagement with our paper, which pushed us to think harder about a number of difficult questions. We welcome further comments on the paper, and we very much hope that our discussions will motivate further research with new theory or data. On balance, while we disagree with the contention that our benchmark estimates suffered from an “inherent upward bias,” we agree with the broader idea that the measurement of the extent of tax evasion is very challenging. We hope that more researchers in academia and in government will embrace the challenge. And consumers of research on tax evasion should bear in mind that the overall extent of tax evasion is never estimated with total certainty. This uncertainty creates some unavoidable difficulties, especially for policymakers who must make do with the best available evidence. We have done our very best to be transparent about major sources of uncertainty and to provide reasonable and informative analysis despite the underlying challenges.

We leave the readers who were patient enough to read this far with one final thought. What data are available to study tax evasion is to some extent a *policy choice*. The NRP and its predecessors exist because policymakers believed it would be valuable to conduct periodic estimates of the tax gap. We agree. Random audit data have facilitated monumental improvements in our understanding of tax evasion. But they are currently limited in their coverage of the wealthiest and highest-income Americans. We therefore enhance Measurement Program (TCMP) to NRP random audits. The last TCMP study examined tax year 1988 and the first NRP random audit study examined tax year 2001. See [Brown and Mazur \(2003\)](#) for a discussion of the change.

courage readers to ponder the policy implications of *uncertainty* about the tax gap at the top of the income distribution. How could policy help us to understand this issue better?

## References

- Alstadsaeter, A., Johannesen, N., and Zucman, G. (2018). Who owns the wealth in tax havens? Macro evidence and implications for global inequality. *Journal of Public Economics*, 162:89 – 100.
- Alstadsaeter, A., Johannesen, N., and Zucman, G. (2019). Tax Evasion and Inequality. *American Economic Review*, 109(6):2073–2103.
- Auten, G. and Langetieg, P. (2020). The distribution of underreported income: What we can learn from the NRP? Presentation at the national tax association spring symposium on 14 may 2020.
- Auten, G. and Splinter, D. (2019). Income inequality in the united states: Using tax data to measure long-term trends. Working paper.
- Auten, G. and Splinter, D. (2021). Comment: Tax evasion at the top of the income distribution: Theory and evidence. Comment, available online at <http://www.davidsplinter.com/AutenSplinter-TaxEvasion.pdf>.
- Brown, R. E. and Mazur, M. J. (2003). IRS’s comprehensive approach to compliance measurement. *National Tax Journal*, 56(3).
- Cooper, M., McClelland, J., Pearce, J., Prinsziano, R., Sullivan, J., Yagan, D., Zidar, O., and Zwick, E. (2016). Business in the United States: Who Owns it, and How Much Tax Do They Pay? *Tax Policy and the Economy*, 30(1):91–128.
- DeBacker, J., Heim, B., Tran, A., and Yuskavage, A. (2020). Tax noncompliance and measures of income inequality. *Tax Notes*, February 17, 2020.
- Efron, B. and Morris, C. (1975). Data analysis using stein’s estimator and its generalizations. *Journal of the American Statistical Association*, 70(350):311–319.
- Feinstein, J. S. (1991). An Econometric Analysis of Income Tax Evasion and Its Detection. *The RAND Journal of Economics*, 22(1):14–35.
- IRS (2007). Tax year 2001 federal tax gap overview. <https://www.irs.gov/pub/irs-soi/01rastg07map.pdf>.
- IRS (2016). Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008-2010. Publication 1415 (rev. 5-2016), Washington, DC.
- IRS (2019). Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2011-2013. Publication 1415 (rev. 9-2019), Washington, DC.

- IRS (2020). Internal revenue service data book 2019. Publication 55-b (rev. 06-2020), Washington, DC.
- Johannesen, N., Langetieg, P., Reck, D., Risch, M., and Slemrod, J. (2020). Taxing hidden wealth: The consequences of us enforcement initiatives on evasive foreign accounts. *American Economic Journal: Economic Policy*, 12(3):312–46.
- Johns, A. and Slemrod, J. (2010). The Distribution of Income Tax Noncompliance. *National Tax Journal*, 63(3):397–418.
- Johns, D. (2009). Preliminary results of the 2003/2004 national research program s corporation underreporting study. presentation, 2009 IRS research conference. Technical report.
- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, 79(3):651–692.
- Piketty, T., Saez, E., and Zucman, G. (2018). Distributional national accounts: methods and estimates for the united states. *The Quarterly Journal of Economics*, 133(2):553–609.
- Smith, M., Zidar, O., and Zwick, E. (2019). Top wealth in the united states: New estimates and implications for taxing the rich. *Unpublished*.
- Zucman, G. (2013). The Missing Wealth of Nations: Are Europe and the U.S. net Debtors or net Creditors? *The Quarterly Journal of Economics*, 128(3):1321–1364.
- Zucman, G. (2015). Factual mistakes in “robin hood’s very merry man”, *Wall Street Journal*. Available online at <https://gabriel-zucman.eu/files/press/201510WSJAnswer.pdf>.