TAS Research and Related Studies 3

Audit Impact Study: The Specific Deterrence Implications of Increased Reliance on Correspondence Audits[1]

1 This research was conducted for the National Taxpayer Advocate by B. Erard and Associates, LLC under contract TIRNO-17-E-00019. This study was conducted by Brian Erard, Erich Kirchler, and Jerome Olsen. Any opinions expressed in this report are those of the authors and do not necessarily reflect the views of the National Taxpayer Advocate. We thank the National Taxpayer Advocate and the Taxpayer Advocate Service (TAS) for their sponsorship and support of this research study, and, in particular, Jeff Wilson for his substantial assistance and helpful advice. The statistical information in this research study was not provided or reviewed by the Secretary under IRC § 6108(d). *See* IRC § 7803(c)(2)(B)(ii)(XII).

INTRODUCTION

Tax administrations rely on audits as a key tool for promoting and enforcing tax compliance. Since audit resources are costly and scarce, however, they are largely reserved for cases with substantive compliance risks. The overall audit rate for U.S. federal individual income tax returns has decreased over time, from one percent of returns filed in 1990 to six-tenths of one percent of returns filed in 2017. There also has been a substantial change in the composition of audits over this period. Whereas face-to-face audits accounted for the majority (62 percent) of all examinations of returns filed in 1990, the lion's share (81 percent) of all audits of returns filed in 2017 were conducted through correspondence.[2]

An important objective of tax audits is specific deterrence: improving the future compliance behavior of those taxpayers who have been targeted for an audit. Past research on the specific deterrent effect of an audit has largely focused on the impact of random audits, either in a laboratory or a field setting.[3] In contrast, most real-world tax audits are targeted towards returns that are considered to be at substantive risk for noncompliance. To better understand how the population of taxpayers who are targeted for risk-based audits responds to examinations, we focus in this study on the role of operational rather than random audits.

A second limitation of existing studies is that they do not distinguish between audit approaches. In comparison with face-to-face audits, correspondence examinations tend to be more narrowly focused and less costly to undertake. At the same time, they are more impersonal. In fact, a recent survey study commissioned by TAS (Erard et al., 2018) indicates that, while most taxpayers who have received a face-to-face examination are able to recall their audit experience, the majority of those who have received a correspondence examination report that they have not been audited. This suggests that many taxpayers do not perceive a correspondence examination as a genuine audit. In this study, we investigate whether face-to-face audits impact future taxpayer reporting behavior differently than correspondence audits. Given the increasing reliance of the Internal Revenue Service (IRS) on correspondence examinations in response to budgetary pressures, this is a question of significant practical importance.

This study relies on a large and unique data base that includes audit and comparison samples covering two different tax years: 2010 and 2014. The tax year 2010 sample includes nearly 53,000 self-employed taxpayers (Schedule C filers) who experienced either a face-to-face or correspondence audit of their tax year 2010 returns as well as a comparison group of approximately 421,000 unaudited Schedule C filers.[4] The sample for tax year 2014 includes about 17,000 audited self-employed taxpayers as well as a comparison group of 377,000 Schedule C filers who were not audited in that year. To estimate the impacts of face-to-face and correspondence examinations on future reporting behavior, we apply an inverse probability weighting methodology. This methodology produces two separate sets of weights for the subsample of unaudited taxpayers. The first set of weights is used to make this subsample representative of taxpayers who received a correspondence audit, while the second set makes it representative of taxpayers who received a face-to-face audit. In this way, the weighted subsample is able to serve as a counterfactual for how the respective groups of audited taxpayers would have behaved in the absence of their audits.

2 Authors' calculations based on Internal Revenue Service (1991, Table 11, p. 24) and Internal Revenue Service (2018, Table 9a, p. 23).

3 Two exceptions are Erard (1991) and Beer et al. (2015).

4 Schedule C is a form used by U.S. federal individual income taxpayers to report income and losses from a non-farm sole proprietorship.

The results indicate that face-to-face audits are consistently effective in promoting future reporting compliance. For tax year 2010, a face-to-face examination is predicted to result in more than a 40 percent increase in reported taxes for the first tax year following the initiation of the audit and a 27 percent increase for the subsequent tax year. For tax year 2014, the estimated pro-deterrent effect is even larger, ranging from 62 to 97 percent, depending on whether the audit takes place later or earlier in the examination cycle. On the other hand, the impact of correspondence audits on self-employed taxpayers is more nuanced. Correspondence audits that are undertaken shortly after filing (prior to the filing for the subsequent tax year) tend to have a counter-deterrent effect, reducing reported taxes by 6 to 15 percent over the two years following the examination. In contrast, correspondence audits that take place later (after the next year's tax return has been filed) have a pro-deterrent effect, similar in size to that observed for face-to-face examinations. We suspect that these contrary outcomes may reflect differences in the types of issues or taxpayers that are addressed over the correspondence audit cycle. However, more research is needed to understand the reasons underlying this result. More generally, the findings point to a need for further investigation into the proper balance between face-to-face and correspondence examinations.

The remainder of this paper is organized as follows. An overview of the theoretical insights on the specific deterrent effect of an audit is provided in Section 2, while Section 3 describes our estimation methodology. The data are summarized in Section 4, and Section 5 presents the estimation results. Section 6 concludes.

Theoretical Insights

There is no consensus from theoretical models regarding the impact of an audit on a taxpayer's subsequent reporting decisions. In the standard economic model of tax compliance behavior (Allingham and Sandmo, 1972), all relevant parameters influencing one's reporting decision are fixed and known with certainty, including the true level of taxable income, the tax rate, the audit rate, and the penalty rate on unreported income. Consequently, no useful information is learned from an audit that would influence future tax reporting behavior.

To allow for a specific deterrent effect of audits, new sources of uncertainty or taxpayer motivation need to be introduced. For instance, it may be the case that future perceptions of the risk of audit and punishment are influenced by an audit experience. If an audit leads to a higher perceived risk of future examinations, this should make one relatively more compliant. Several random audit studies (Kleven et al., 2011; Gemmell and Ratto, 2012; DeBacker et al., 2015; and Advani, Elming, and Shaw, 2017) have attributed their findings of a specific pro-deterrent effect to this cause. On the other hand, some laboratory experiments (Guala and Mittone, 2005; Mittone, 2006) have produced evidence of "bomb-crater effects" — instances of increased cheating following an audit in an earlier round of a tax compliance game. Kastlunger et al. (2009) attribute such behavior to a misperception that the likelihood of receiving a second consecutive audit is extremely low.

In addition to perceptions of audit risk, researchers have proposed a variety of other pathways for audits to impact future taxpayer reporting behavior, including loss repair motivations (Maciejovski, Kirchler, & Schwarzenberger, 2007), dynamic reporting considerations (Engel & Hines, 1999), uncertainty about either one's tax liability (Scotchmer & Slemrod, 1989; Beck & Jung, 1987) or the tax agency's capacity to detect tax underreporting (Beer et al., 2015), and tax morale (Feld & Frey, 2003; Frey, 2011; Frey, Benz, & Stutzer, 2004).

Correspondence audits tend to be more narrow in scope and less personal than face-to-face examinations. From a theoretical perspective, these differences in audit characteristics may have implications for perceptions of future audit risk, beliefs about the capacity of the tax administration to uncover evasion when it is present, the level of certainty about true tax liability, and tax morale. In turn, these factors may influence future reporting behavior. To address this possibility, we develop separate estimates of the specific deterrent effect for these two audit types.

Estimation Methodology

To control for differences in characteristics among taxpayers who have experienced a correspondence audit, a face-to-face audit, or no audit when estimating specific deterrent effects, we rely on the inverse probability of treatment weighting (IPTW) methodology. This methodology requires no assumptions about the functional relationship between the determinants of audit selection and taxpayer reporting behavior. Under this approach, one begins by estimating the propensity scores (predicted probabilities) π_i^c, π_i^f, and π_i^{na}, associated with a correspondence audit, a face-to-face audit, and no audit, respectively. Define the indicator variable for taxpayers in the sample who did not experience an audit as I^{na}, and denote the outcome variable as y. The Horvitz-Thompson estimator of the expected counterfactual outcome among taxpayers receiving a correspondence audit had they not been audited is then defined as:

$$\left(\frac{1}{\sum I_i^{na}}\right)\sum\left[I_i^{na}\ y_i\left(\frac{\pi_i^c}{\pi_i^{na}}\right)\right].$$

Similarly, the Horvitz-Thompson estimator of the expected counterfactual outcome among taxpayers receiving a face-to-face audit had they not received an examination is defined as:

$$\left(\frac{1}{\sum I_i^{na}}\right)\sum\left[I_i^{na}\ y_i\left(\frac{\pi_i^f}{\pi_i^{na}}\right)\right].$$

These counterfactual outcome estimates are thus computed as a weighted average of the outcomes observed for the unaudited taxpayers in the sample, where the weights are computed as the ratio of the relevant propensity scores.[5] Intuitively, greater weight is applied to unaudited taxpayers with a relatively high predicted probability of selection for the specified type of audit, as taxpayers with their characteristics will tend to have greater representation among the sample of filers who received that type of audit. The estimated specific-deterrent effect is computed as the difference between the actual mean outcome for taxpayers who received the specified type of audit and the estimated counterfactual outcome for these taxpayers.

In our analysis, we rely on propensity scores derived from a multinomial logit model of audit selection.[6] This analysis allows for three possible audit selection outcomes: (1) no audit, (2) correspondence audit; or (3) face-to-face audit. We have constructed a set of over 60 candidate explanatory variables for the audit selection process. Included among these covariates are measures of the current and prior year DIF-scores that are relied upon by the IRS to help identify high-risk returns for examination. A sequential selection process is employed to choose the final set of covariates.

Some taxpayers in our sample were audited prior to filing the next year's tax return, and others were audited after doing so. To account for differences in the audit selection process for these two groups, a

5 In our application, we follow the conventional approach to stabilizing the weights used in our analysis.

6 We employ sample weights in estimation to account for the choice-based nature of our data sample.

separate multinomial logit analysis is performed for each group. The estimation results are employed to predict the odds of a correspondence audit and the odds of a face-to-face audit (relative to no audit) for each taxpayer in the estimation sample.

For taxpayers who were audited prior to filing the next year's tax return, our outcome variable is the difference between the natural log of reported tax for a subsequent tax year (either of the next two filed tax returns) and the natural log of reported tax on the audited return. Effectively, then, this approach produces "difference-in-differences" estimates to account for unobserved time-invariant differences between the audited and unaudited taxpayers in our sample. A one-year ahead impact estimate is derived using the very next year as the subsequent tax year, while a two-year ahead impact estimate is obtained using the following year as the subsequent tax year.

In the case of taxpayers who were audited only after filing their tax return for the following year, we rely on our one-period ahead impact estimate as a "placebo test". Since these taxpayers were not aware of the audit until after they had filed a return for the following year, the one-period ahead impact estimate has an expected value of zero. Therefore, this placebo test provides a useful check on the quality of the matching process. The two-period ahead impact estimate calculated as described above is effectively a one-period ahead estimate for this group, since the return filed two years later was the first return that was filed subsequent to the initiation of the audit.

DATA

The data for this study includes detailed line-item information from returns filed by audited and unaudited self-employed taxpayers. The audit sample consists of all Schedule C filers who experienced an audit of the return they filed for the relevant tax year (2010 or 2014), excepting those who failed to satisfy one or more of the following eligibility criteria:

- The taxpayer filed Schedule C for at least three years, including on the return that was audited, the return filed for the previous tax year, and the return filed for the subsequent tax year.[7]
- The taxpayer filed at least five tax returns in chronological order, including the return that was audited, the returns for the previous two tax years, and the returns for the subsequent two tax years.
- The audit was not focused on certain specialized issues that would make it especially difficult to identify a matched unaudited taxpayer.[8]
- The audit was initiated prior to the date when the taxpayer filed the return for the second subsequent tax year.[9]
- No earlier audits were initiated or closed within two years of the date that the audited return was filed.
- If the taxpayer was subsequently audited for another tax year, this subsequent audit was not initiated until at least one year had passed from the start of the earlier audit.
- The taxpayer did not reside in a U.S. territory or outside of the U.S.

7 If returns for additional tax years were filed, these were also required to be filed in chronological order.

8 Certain types of audits were excluded on the basis of source codes that an IRS examination expert deemed to fit this criterion.

9 The records concerning the dates that a given return was filed and the taxpayer was notified of an audit are imprecise, so we excluded ambiguous cases where the return filing date and audit notification date were within forty-five days of each other.

A stratified random sample of taxpayers who did not experience an audit was also drawn. To be eligible for this sample, the following requirements had to be met:

- The taxpayer filed Schedule C at least on the return filed for the reference audit year (tax year 2010 or tax year 2014), the return for the preceding year, and the return for the subsequent year.
- The taxpayer filed returns in chronological order for at least five years that bracket the reference audit year.[10]
- The return for the reference audit year was not audited, and no other audits were initiated or closed in the two years preceding or the two years subsequent to the date that return was filed.
- The taxpayer did not reside in a U.S. territory or outside of the U.S.

The members of the unaudited taxpayer sample were selected to provide a means for developing counterfactual estimates of behavior for the members of the audit sample. Therefore, an important sampling objective was to include a substantial number of unaudited taxpayers who had characteristics similar to those of audited taxpayers. To achieve this goal, a 1 percent probability sample was drawn from the members of the eligible population with DIF-scores that were in the range commonly observed for audited taxpayers in the same examination class. As discussed in Section 3, the DIF-score is a key risk assessment measure employed by the IRS when selecting individual income tax returns for examination. To ensure that all eligible members of the overall population of unaudited taxpayers had at least some chance of being included in the sample, a 0.3 percent probability sample was drawn from this general population. To account for the stratified nature of the sampling process, a set of sample weights was constructed so that the sample can be made representative of all eligible filers in the population.

For each sampled taxpayer, the data include detailed line item information from each tax return filed for the reference audit year, the two prior years, and the two subsequent years. Figure 4.3.1 presents the numbers of audited and unaudited taxpayers that were sampled.

FIGURE 4.3.1, Sample Count of Taxpayers by Audit Type

Reference Audit Year	Correspondence	Face-to-Face	Unaudited
Tax Year 2010	40,359	12,541	421,309
Tax Year 2014	13,629	3,274	377,168

RESULTS

The estimated one-period and two-period ahead specific deterrent effects for correspondence and face-to-face audits of self-employed taxpayers are presented in Figure 4.3.2. These estimates represent the predicted change in the natural log of reported tax liability associated with the specified tax year and examination type. In Figure 4.3.3, these estimates have been translated into measures of the predicted percentage change in reported tax liability using the formula:

$$\%chg = \left(e^{estimate} - 1\right).$$

10 If returns for additional tax years were filed, these were also required to be filed in chronological order.

For audits that were initiated after the return for the following year was filed (but before the 2nd subsequent return was filed), a placebo impact estimate is provided for the next year's return. The expected audit impact is equal to zero in this year, since the taxpayer would not have been aware of the audit when that return was filed. The estimated impact for each audit type is in fact small and statistically insignificant, consistent with expectations. This finding helps to substantiate the validity of the estimation methodology.

The estimation results indicate that face-to-face audits have a large specific deterrent effect. For tax year 2010 audits that began prior to the filing of the tax year 2011 return, reported tax liability is estimated to have increased by 40.8 percent ($e^{0.3423}$-1) for tax year 2011 and 27.3 percent for tax year 2012 as a result of the examination. For audits that began after the tax year 2011 return was filed, reported tax liability is estimated to have increased by 37.5 percent in tax year 2012. The estimated impacts are even larger for tax year 2014 audits. For audits that began prior to the filing of the tax year 2015 return, reported tax liability is estimated to increase by more than 95 percent in tax year 2015 and remain around that level the following tax year. For tax year 2014 audits that began after the tax year 2015 return was filed, reported tax liability is estimated to increase by more than 61 percent on the first return filed since the audit was initiated (tax year 2016).

The estimation results for correspondence audits are more nuanced. For audits that began prior to the filing of the tax year 2011 return, there is evidence of a counter-deterrent effect. Reported tax liability is estimated to have declined by 7.3 percent ($e^{-0.076}$-1) in tax year 2011 and 8.3 percent in tax year 2012 as a result of the examination. On the other hand, reported tax liability is estimated to have been 37.5 percent higher in tax year 2012 for taxpayers whose tax year 2010 audits were initiated later in the examination cycle.

FIGURE 4.3.2, Estimated Specific Deterrent Effect on Natural Log of Reported Tax Liability by Audit Type and Tax Year

Audit Type	Audited Before Next Return Filed		Audited After Next Return Filed	
	1st Year Impact	2nd Year Impact	Placebo Impact	1st Year Impact
Tax Year 2010 Audit Results				
Correspondence	-0.0760* (5.26)	-0.0869* (4.83)	-0.0125 (0.62)	0.3187* (14.98)
Face-to-Face	0.3423* (7.92)	0.2414* (5.37)	0.0406 (1.17)	0.3184* (8.88)
Tax Year 2014 Audit Results				
Correspondence	-0.0585* (2.30)	-0.1622* (5.17)	0.0897* (2.58)	0.4766* (12.15)
Face-to-Face	0.6696* (11.87)	0.6796* (11.19)	0.0906 (1.31)	0.4809* (6.57)

(absolute value of t-statistics in parentheses; asterisk indicates significance at the 5 percent level)

FIGURE 4.3.3, Predicted Percentage Change in Reported Tax Liability by Audit Type and Tax Year

Audit Type	Audited Before Next Return Filed		Audited After Next Return Filed	
	1st Year Impact	2nd Year Impact	Placebo Impact	1st Year Impact
Tax Year 2010 Audit Results				
Correspondence	-7.32%	-8.32%	-1.24%	37.53%
Face-to-Face	40.82%	27.30%	4.14%	37.49%
Tax Year 2014 Audit Results				
Correspondence	-5.68%	-14.97%	9.38%	61.06%
Face-to-Face	95.34%	97.31%	9.48%	61.75%

The disparity among the findings within the correspondence audit group may reflect differences in the types of issues or taxpayers that are addressed over the correspondence audit cycle. Based on a preliminary analysis of audit findings, approximately half of all correspondence audits involving self-employed taxpayers are initiated before the taxpayer has filed a return for the following tax year. A very substantial share (over 70 percent) of these early audits involve taxpayers who claim the Earned Income Credit (EIC).[11] In contrast, only about 19 percent of the audits initiated later in the cycle (after the return for the following tax year has been filed) involve EIC claimants.

Among self-employed taxpayers who do not claim the EIC, a disproportionate share of the correspondence audits that are initiated early in the examination cycle involve questionable refunds or claims for certain other tax credits. On the other hand, audits that take place later in the cycle (after the return has been filed for the following tax year) are much more likely to involve issues pertaining to the business, various claims for itemized deductions, or claims for some other types of tax credits.

Since EIC claimants account for such a large share of correspondence examinations overall (approximately 45 percent) and, especially, of audits that take place early in the examination cycle, we have extended our estimation methodology to develop separate audit impact estimates for claimants and non-claimants. The results are summarized in Figure 4.3.4 and 4.3.5. Similar patterns are observed for both groups. Correspondence audits that take place later in the examination cycle (after the next tax return has been filed) have a substantial pro-deterrent effect, while those that take place early in the cycle have a moderate counter-deterrent effect. This heterogeneity in outcomes may be attributable to differences in the characteristics of the taxpayers and the issues facing them at different points in the examination cycle. Alternatively, it could have to do with the amount of time that lapses between filing a return and being notified of an audit.

11 Some of these correspondence audits involve issues beyond the EIC.

FIGURE 4.3.4, Estimated Specific Deterrent Effect of Correspondence Audits on Natural Log of Reported Tax Liability by EIC Claim Status and Tax Year

EIC Claim Status	Audited Before Next Return Filed		Audited After Next Return Filed	
	1st Year Impact	2nd Year Impact	Placebo Impact	1st Year Impact
Tax Year 2010 Audit Results				
Claimant	-0.1455* (2.20)	-0.1162* (2.39)	-0.0175 (0.72)	0.5334* (6.75)
Non-claimant	-0.0503* (2.34)	-0.1306* (5.45)	-0.0343 (1.56)	0.2507* (11.35)
Tax Year 2014 Audit Results				
Claimant	-0.0201 (0.64)	-0.1527* (3.83)	0.1562 (1.81)	0.6459* (5.25)
Non-Claimant	-0.1432* (3.72)	-0.1428* (11.19)	0.0490 (1.14)	0.4434* (6.52)

(absolute value of t-statistics in parentheses; asterisk indicates significance at the 5 percent level)

EIC claimants make up about 19 percent of the self-employed taxpayers who experience an audit later in the examination cycle. Although the pro-deterrent impact of these later audits is substantial for both EIC claimants and non-claimants, it is especially large for EIC claimants (53 percent increase in reported tax liability for tax year 2010 audits and 65 percent for tax year 2014 audits, compared to 25 percent and 44 percent, respectively, for non-claimants).

FIGURE 4.3.5, Predicted Percentage Change in Reported Tax Liability for Correspondence Audits by EIC Claim Status and Tax Year

EIC Claim Status	Audited Before Next Return Filed		Audited After Next Return Filed	
	1st Year Impact	2nd Year Impact	Placebo Impact	1st Year Impact
Tax Year 2010 Audit Results				
Claimant	-13.54%	-10.97%	-1.73%	70.47%
Non-claimant	-4.91%	-12.24%	-3.37%	28.49%
Tax Year 2014 Audit Results				
Claimant	-1.99%	-14.16%	16.91%	90.77%
Non-claimant	-13.34%	-13.26%	5.02%	55.80%

As summarized in the Appendix, we have performed some sensitivity analyses involving alternative estimation methodologies and additional tax years. The results corroborate our main findings.

CONCLUSION

An important purpose of audits beyond immediate revenue generation is to discourage future reporting noncompliance. The existing empirical literature on the specific-deterrent effect of an audit has generally found that audits do improve future reporting behavior, although some laboratory experiments have uncovered "bomb-crater" effects and, in earlier work (Beer et al., 2015), we have found evidence of a counter-deterrent effect in cases where audits fail to uncover any noncompliance. However, much of the empirical literature has focused on random audits, and the role of audit type has not been explored. In practice, audit selection at the IRS is overwhelmingly risk-based rather than random, and there has been a marked shift over time away from face-to-face examinations and towards correspondence audits.

In this paper, we have conducted a preliminary analysis of how operational audits impact future reporting behavior, and we have paid special attention to how correspondence and face-to-face examinations may differ in this regard. Our estimation results indicate that correspondence audits that take place later in the examination cycle (after the subsequent tax return has been filed) are comparable to face-to-face audits in terms of their impact on future reporting behavior. Both types of audits have a substantial pro-deterrent effect when they are initiated after the following year's tax return has been filed. In contrast, however, correspondence audits that take place early in the audit cycle are actually associated with a counter-deterrent effect. Reported tax liability is estimated to fall by 6 to 15 percent in the first two tax years following the initiation of the audit. This is an important finding, because approximately half of all correspondence examinations take place early in the audit cycle.

Overall, then, the results of this study suggest that correspondence audits are not a perfect substitute for face-to-face examinations. Not only do they tend to be more narrowly targeted and impersonal, they also appear to be less consistent in terms improving future taxpayer reporting behavior. This raises concerns about IRS' increasing reliance on this form of enforcement. The disparate findings for correspondence audits that take place at different points in the audit cycle may reflect differences in the types of issues or taxpayers that are addressed over the cycle. Alternatively, the amount of time that lapses between filing a return and notification of an audit may have a direct impact on future reporting behavior. Further research is needed to understand this result. More generally, the findings suggest that further study on the proper balance between face-to-face and correspondence audits is warranted.

REFERENCES

Advani, Arun, W. Elming, and J. Shaw. 2015. How long-lasting are the effects of audits? TARC Discussion Paper 011-15. Exeter, UK: Tax Administration Research Centre, University of Exeter.

Allingham, M. G. and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3-4), 323-338.

Beck, Paul and W. O. Jung. 1987. An economic model of taxpayer compliance under complexity and uncertainty. *Journal of Accounting and Public Policy*, 8, 1-27.

Beer, S., M. Kasper, E. Kirchler, and B. Erard. 2015. Audit impact study. *National Taxpayer Advocate 2015 Annual Report to Congress*, 2, 68-98.

DeBacker, J., B. T. Heim, A. Tran, and A. Yuskavage. 2018. Once bitten, twice shy? The lasting impact of IRS audits on individual tax reporting. *The Journal of Law and Economics*, 61(1), 1-35.

Erard, B. 1992. The influence of tax audits on reporting behavior, in *Why people pay taxes: Compliance and enforcement*, ed. Joel Slemrod, Ann Arbor: University of Michigan Press, 95–114.

Erard, B., M. Kasper, E. Kirchler, and J. Olsen. 2019. What influence do IRS audits have on taxpayer attitudes and perceptions? Evidence from a national survey. *National Taxpayer Advocate 2018 Annual Report to Congress*, 2, 77-130.

Feld L. and B. S. Frey. 2003, Deterrence and tax morale: How tax administrations and taxpayers interact, *European Review*, 11(3), 385-406.

Francesco, G. and L. Mittone. 2005. Experiments in economics: External validity and the robustness of phenomena. *Journal of Economic Methodology*, 12, 495-515.

Frey, B. S., M. Benz, and A. Stutzer. (2004). Introducing procedural utility: Not only what, but also how matters. *Journal of Institutional and Theoretical Economics*, 160, 377-401.

Frey, B. S. (2011). Punishment – and beyond, *Contemporary Economics*, 5 (2), 90-99.

Gemmell, N. and M. Ratto. 2012. Behavioral responses to taxpayer audits: Evidence from random taxpayer inquiries. *National Tax Journal*, 65 (1), 33-58.

Internal Revenue Service. 1991. Annual report. Retrieved from https://www.irs.gov/pub/irs-soi/91dbfullar.pdf.

Internal Revenue Service. 2018. Data book. Retrieved from https://www.irs.gov/pub/irs-soi/18databk.pdf.

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. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. 2011. Unwilling or unable to cheat? Evidence from a randomized tax audit experiment in Denmark. *Econometrica*, 79(3), 651-692.

Maciejovsky, B., E. Kirchler, and H. Schwarzenberger. 2007. Misperceptions of chance and loss repair: On the dynamics of tax compliance. *Journal of Economic Psychology*, 28(6), 678-691.

Mittone, L., F. Panebianco, and A. Santoro. 2017. The bomb-crater effect of tax audits: Beyond the misperception of chance. *Journal of Economic Psychology*, 61, 225-243.

Scotchmer, S. and J. Slemrod. 1989. Randomness in tax enforcement. *Journal of Public Economics*, 38(1), 17-32.

APPENDIX: SENSITIVITY ANALYSES

To investigate the robustness of our findings, we have experimented with two alternative estimation methodologies:

1. **Inverse Probability of Treatment Weighting with Regression Adjustment.** This method extends the methodology discussed in the main text by incorporating a regression specification for the natural log of reported tax liability. The regression equation relies on the same explanatory variables as the multinomial logit specification for audit type. A separate regression is estimated for taxpayers associated with each audit status (face-to-face audit, correspondence audit, or no audit). The regression estimates are used jointly with the predicted audit status probabilities to estimate audit impacts. These estimates are "doubly robust", meaning that they will be consistent so long as either the multinomial logit model is correctly specified or the regression model is correctly specified, even if the other model is incorrectly specified.
2. **Nearest Neighbor Matching on the Vector of Propensity Scores.** As with the IPTW methodology described in the main text, a multinomial logit model is employed to predict the likelihood of a face-to-face audit and the likelihood of a correspondence audit for each taxpayer in the estimation sample. For each audited taxpayer, a match is found to one or more unaudited taxpayers (with replacement) who have similar predicted probabilities of each type of audit. A Mahalanobis distance criterion is used to identify the best match(es).[12] The mean difference between the future reported tax liability of the audited taxpayers and their matched counterparts then serves as the audit impact estimate for a given period. We have experimented with alternately matching 1, 5, and 15 unaudited taxpayers to each audited taxpayer.

We have also experimented with some additional tax years between 2010 and 2014. Overall, the results from our alternative estimation methodologies and estimation years are qualitatively quite similar to those presented in Section 5, which lends credibility to the main findings of our study.[13]

12 No adequate matches can be found for a modest number of audited taxpayers under this approach. These taxpayers are therefore excluded from the audit impact estimation.

13 The one exception is that our nearest neighbor matching methodology produces evidence of a counter-deterrent effect for tax year 2012 correspondence audits of EIC recipients taking place after the tax year 2013 return had been filed. In contrast, both the IPTW methodology and the IPTW with regression adjustment methodology provide evidence of a pro-deterrent effect for this population and tax year.