

**IRS Preparer-Level Treatment Tests**  
**Results from the First Year of a Multi-Year Study**

December 2014

Karen Masken  
Return Preparer Office  
Internal Revenue Service

## **Introduction**

In 2010, the Internal Revenue Service (IRS) adopted regulations aimed at establishing standards for paid tax return preparers who prepare individual income tax returns (Form 1040). The objective was to improve voluntary compliance through increased oversight of the paid preparer industry with the goal of reducing errors on tax returns. The Return Preparer Office (RPO) was formed to meet this objective. RPO does not have any enforcement authority and so is focused primarily on education, outreach, and partnering with the paid preparer community. Specifically, the three primary strategic goals of RPO are:

1. Register and promote a qualified tax professional community.
2. Improve the compliance and accuracy of returns prepared by tax professionals.
3. Support a stakeholder-focused culture that encourages voluntary compliance and continuous improvement.

One effort to meet these strategic goals was a new requirement that began January 1, 2011, for all paid preparers who prepare Form 1040 returns to register with the IRS, obtain a preparer tax identification number (PTIN), and enter it exclusively as the preparer identifying number on the returns they completed. Previously, PTINs had been optional.

Another effort involved a multi-year study that started in 2012 to determine the effect on tax return preparation accuracy of various treatments applied to paid tax return preparers. Because IRS has historically focused on taxpayer-level treatments, there is currently only a limited understanding of how preparer-level treatments effect change in preparer and client tax compliance. The goal of the multi-year study is to understand what treatments are effective on different segments of the non-compliant preparer population in an effort to determine the cost-effectiveness of the treatments. Details on the development and design of the study and results from the first year are presented in this paper.

## **Motivation**

IRS resources have become increasingly scarce in the recent budget environment. The driver of the preparer-level treatment approach is that treatment of a single preparer is likely to improve the compliance of many tax returns, increasing the expected Return on Investment (ROI) of treatment resources. This is similar to intervening at a wholesale level, rather than a retail level.

As illustrated in **Figure 1**, preparer compliance is not a binomial variable but rather a continuous spectrum. On one end of the spectrum are compliant, well informed preparers while on the other end of the spectrum are preparers who willfully perpetrate fraud. In the middle of the spectrum are those who might be unintentionally making errors due to lack of knowledge or those who are willfully noncompliant in their tax preparation but could be potentially moved towards voluntary compliance with a light touch. Moving preparers towards greater voluntary compliance is the most cost effective action in that it protects revenue by having tax and credits reported correctly on the tax return – rather

than IRS trying to recover revenue through examination and collection after the return has been filed, and in many cases, after a refund has been paid.

**Figure 1.**



Traditionally, IRS has focused its enforcement resources toward the noncompliant and fraudulent end of the spectrum on what are predicted to be the largest, most substantial problems. For less substantial cases and for preparers in the middle of the spectrum, the tradition has been to rely on non-targeted services, such as tax forums and webinars. Preparers subject to enforcement are generally subject to examination and in some cases, criminal investigation and prosecution. While these actions are necessary in some circumstances, they are very resource intensive and costly. The audit of a preparer typically involves auditing approximately 30 of their clients (for individual taxpayers, the current audit rate is less than two percent overall). If the case goes to litigation, then the Justice Department becomes involved and a very costly case has to be put together. Typically, the Justice Department handles less than fifty such cases each year. Since there are currently over 700,000 registered preparers, traditional enforcement efforts can only reach a small fraction of preparers.

A preparer, like taxpayers, has certain rights. While there is a desire by some to take draconian actions against a preparer thought to be non-compliant (“Put them out of business!”, “Shame them publicly!”, etc...), many of these proposed ideas fail to recognize that, in addition to enforcement resource constraints, a preparer is entitled to many of the same rights afforded to individual taxpayers. By law, the IRS cannot share the identity of a preparer under suspicion of filing non-compliant returns with the general public, or even with the preparer’s clients, as it would violate disclosure statutes and would not afford the preparer due process. The only time a preparer’s identity is disclosed is when a criminal or civil suit is brought against the preparer, at which point the case becomes public. Unfortunately, since these are the cases that make headlines, it leaves many with the impression that ALL preparers are nefarious. While there are unequivocally bad actors in the community, RPO’s view is that these are the exception, not the rule, and many non-compliant preparers are in the middle of the compliance spectrum.

While preparers towards the noncompliant/fraudulent end of the spectrum may require more expensive and intrusive treatments (e.g. audits/injunctions), finding effective, lower cost treatments for preparers in the middle of the spectrum could have a significant impact on revenue collected. Many preparers in the middle currently go untouched as traditional IRS exam approaches would not be cost effective and the preparer may not partake in IRS services. Many of these preparers simply may not fully understand all the rules and may benefit from education as opposed to enforcement. Others might knowingly make errors believing they go unnoticed, but could be moved toward

voluntary compliance by a touch lighter than enforcement. If less expensive treatments are found to be effective, they would allow for a larger number of preparers to be treated and moved toward voluntary compliance by the same finite set of resources, thus protecting more revenue. Therefore, the focus of the treatments being tested is on preparers in the middle of the compliance spectrum. It should also be noted that a preparer is defined as an individual, not the firm or business they work for (e.g. many well-known tax preparation businesses employ many individual preparers).

## **Methodology**

### *Issue Selection*

The first step in the treatment development process was to identify which tax compliance issue to address. The National Research Program (NRP) is an initiative at IRS that conducts audits on a random sample of individual taxpayers, thereby providing unbiased estimates of compliance for most line items on Form 1040 (individual income tax return). RPO analyzed the NRP results from tax years 2006 and 2007 and looked at the Tax Gap report based on the 2001 NRP study (the most current at the time) to inform the decision as to which issues to address with the treatments. Not only was the overall magnitude of the compliance problem considered, but also to what extent returns by paid preparers contributed to non-compliance. The issue chosen for treatment was Schedule C net income. RPO found that approximately 75 percent of Schedule C returns were completed by paid preparers, and about 75 percent of those returns had errors. The overall contribution of misreported Schedule C income to the tax gap is approximately \$68 billion per year, which is almost 30 percent of the tax gap associated with individual income taxpayers.

### *Model Development*

RPO inherited an established examination plan beginning in 2010 that identified preparers for treatment and either subjected them to an Educational Visit or sent them a letter stating there were errors on their returns. The filters used to identify these preparers were developed by subject matter experts based on their experience. In an audit<sup>1</sup> by the Treasury Inspector General for Tax Administration (TIGTA), the filters were critiqued for not being data driven and identified as a management challenge. TIGTA recommended that a data driven selection criteria be developed. RPO agreed with this recommendation and error detection models using the National Research Program (NRP) data were developed.

Because the NRP sampling rate overall is less than 0.01 percent, it is rare for more than one client per preparer to be selected for the random NRP audits. Therefore, the model was initially developed at the taxpayer-level and then rolled up to the preparer-level selection criterion. The model was developed using data from the NRP Tax Years 2006 and 2007 studies and tested on the Tax Year 2008 study. In part, this was out of necessity as the Tax Year 2008 data was not yet available when the model was being

---

1

*Implementation of the Return Preparer Visitation Project Was Successful, but Improvements Are Needed to Increase Its Effectiveness*, June 29, 2012 Audit # 2012-30-068

developed. However, this approach does provide the benefit of assessing how robust the model is to choice of tax year. The model performed as expected on the test data.

After the model was developed, an outside expert was brought in to perform an independent evaluation of the model and its development. The expert found the methodology was appropriate and that the model was effective. One technique employed to evaluate the model was a confusion matrix wherein the model was compared to a random draw. Based on the confusion matrix, the model was found to be more effective than a random draw by seven standard deviations.

### *Eligibility*

To be eligible for the test, a preparer had to have an active PTIN, had to prepare at least 20 Schedule C returns with at least fifteen percent of all their returns containing a Schedule C, and the majority (51 percent or more) of their Schedule C returns had to be flagged by the error detection model. The requirement that the preparer have at least 20 returns and fifteen percent of their total returns with a Schedule C was simply to ensure that Schedule C was prevalent enough in the preparer's business to warrant treatment. Each of a preparer's Schedule C returns was scored using the model and those considered high risk were flagged. The Schedule C returns were then aggregated by preparer and the percent of the preparer's Schedule C returns that were flagged was calculated. While it is virtually impossible to solve the endogeneity issue of whether it is the preparer or their client driving the noncompliance, RPO believed that if the majority of a preparer's Schedule C returns were flagged, then the preparer was likely to be at least complicit in the noncompliance. Therefore, if at least 51 percent of a preparer's Schedule C returns were flagged, then the preparer was placed in the treatment pool. There were preparers who met all the criteria except that they did not have an active PTIN. These preparers are both programmatically noncompliant (meaning they did not adhere to the basic requirement that they obtain a PTIN and enter it exclusively as the preparer identifying number on the returns they completed) and are at risk for tax non-compliance as well. RPO believed these preparers needed a different treatment to address both issues so they were excluded from this set of tests.

Once the pool of preparers was identified, RPO worked with various other offices to remove preparers who had been selected for enforcement efforts (such as criminal investigations). In the end, there were approximately 9,600 preparers eligible for treatment. On average, each preparer prepared approximately 280 Form 1040 returns and 80 Schedule C returns (the median was 180 and 50, respectively). The average percent of their Schedule C returns flagged by the error detection model was 60 percent.

In comparing the mean number of returns (280) and the median number of returns (180), it is clear that the volume of returns is highly skewed. The volume range was approximately 5300 returns. It is important to point out that in some instances, a preparer is under the supervision of another preparer who ultimately signs off on the return. Thus while all the returns of the supervised preparer will bear that supervisor's PTIN, the supervisor is not necessarily the one who did the actual preparation. It is not possible to

tell from the data whether a return was prepared by an unsupervised or supervised preparer.

### *Treatments and Test Design*

RPO was interested in comparing three different types of treatments. One treatment was an Educational Visit to the preparer by a Revenue Agent to discuss Schedule C issues found on returns that they had prepared. The second was a letter reminding the preparer of their due diligence requirements when preparing returns and warning that they and their clients might be subject to audit (see Appendix A). The third treatment was a letter with the same message regarding due diligence, but also recommending that, as part of the continuing education required at that time<sup>2</sup>, the preparer take a minimum of four hours of continuing education regarding Schedule C (see Appendix B). An outside expert (different from the one employed to evaluate the model) provided consultation in the development of the test design to ensure the tests would produce the desired information.

RPO was not only interested in learning which of these treatments were effective, but also learning how to predict how well a particular preparer would respond to each treatment. The ability to predict where a preparer lies on the compliance spectrum and the least costly treatment necessary to move that preparer to voluntary compliance will result in the most cost effective approach. To achieve these goals, a controlled experiment design was implemented. Preparers in the treatment pool were randomly assigned to either a treatment or a control group. All treatments were applied prior to the start of the 2013 filing season and the determination of the effectiveness of each treatment was based on individual income tax returns filed in 2013.

The determination of effectiveness was based on the percent of clients who were flagged by the error detection model in the filing season after the treatment was applied. At the design phase, it was decided that a treatment would be deemed successful for a particular preparer if it lowered the percent of clients flagged for that preparer by five percentage points. While the definition of success will always be somewhat arbitrary, RPO thought it was important to define success during the design phase in order to avoid the perception that success was defined based on preliminary results (and the natural inclination to do so). The overall effectiveness for each treatment was determined by counting the number of successful preparers for each treatment and comparing it to the respective control group.

Generally speaking, there are five statistical parameters that go into test size calculations: precision, power, confidence level, the test statistic, and the critical value. For the treatment tests specifically, the precision, or the probability of a false-positive (declaring success when it actually failed), was set at the standard five percent level. The power of the test, or the probability of a false-negative (declaring failure when it actually succeeded), was set at the standard level of eighty percent. The confidence level was set

---

<sup>2</sup> At the time of the first year of the test, preparers were required to take continuing education. It was later determined by the courts that IRS did not have the statutory authority to impose this requirement. (Loving et al. v. IRS)

at ninety-five percent. The test statistic was the number of successes as defined above. The critical value was five percentage points, meaning that there had to be at least a five percentage point difference between the test and control to declare the difference statistically significant. The test size given these parameters was calculated to be approximately 1,250 preparers for each of the treatment and control groups.

While preparers were randomly assigned to a treatment or control group, it is important to note that each treatment had a different set of constraints resulting in a different composition of preparers for each treatment and its respective control group. For the Educational Visits, the allocation of resources, namely Revenue Agents, had to be taken into consideration. Rather than drawing a simple random sample from the pool of preparers, a random sample proportionate to the resources available in each of seven IRS-defined areas was drawn. For the Continuing Education Letter, preparers who held a credential (e.g., Certified Public Accountants and Attorneys) were exempt from the IRS continuing education requirement in place at the time as their own credential held them to a higher standard. They were, therefore, excluded from this treatment. The Due Diligence Letter had no constraints. Each of the treatments has its own control group with corresponding constraints, however the control groups are not mutually exclusive of one another. Finally, each control group incorporates controls for possession of a credential and IRS area.

At the time the treatments were being implemented, Hurricane Sandy struck land and IRS generally suspends enforcement actions during a natural disaster. The visits were already underway so preparers in the disaster area states (New Jersey, New York, Rhode Island, and Connecticut) were dropped from the test. Neither of the letters had been sent yet, so the test and control groups were redrawn excluding preparers in the disaster area. As a result, the size of the Educational Visit test is slightly smaller than the two letter tests.

## **Results**

This section presents general results first, more detailed results by preparer characteristics next, followed by client-level information.

### *General Results*

The number of preparers in each of the test and control groups and the number and percent who prepared and signed returns the subsequent year are shown in **Table 1**. As noted in the previous section, the Educational Visit treatment test and control sizes were reduced as a result of the Hurricane Sandy disaster. While the Educational Visit test group has a slightly higher attrition rate (not preparing returns in the subsequent year) than the control, three percent compared to two percent, respectively, this difference is not statistically significant. The other two treatments had no differences between the test and control groups. As previously discussed, the Continuing Education test excluded CPAs and Attorneys, which may explain the slightly higher attrition rate for this group compared to the Due Diligence treatment group (three percent and two percent, respectively) since non-CPAs are less invested in tax preparation.

One consequence of the treatments could be to effectively move a preparer to stop preparing returns all together. To the extent it put a non-compliant preparer out of business, this would be viewed as a positive result. On the other hand, it could also have an unintended negative consequence of moving the preparer to prepare but not sign returns, making them both programmatically and tax reporting noncompliant. However, neither of these appears to have occurred to a significant degree.

**Table 1. Preparer Attrition Comparison between Treatment Groups and Control Groups**

	<b>Treatment</b>	<b>Test</b>	<b>Control</b>
Number of Preparers in Test	Educational Visit	1,113	1,047
	Due Diligence Letter	1,250	1,250
	Continuing Education Letter	1,250	1,250
Number of Preparers who Continued Preparing Returns in Subsequent Year (post treatment)	Educational Visit	1,080	1,026
	Due Diligence Letter	1,225	1,227
	Continuing Education Letter	1,210	1,216
Attrition Rate	Educational Visit	3%	2%
	Due Diligence Letter	2%	2%
	Continuing Education Letter	3%	3%

The number of successes and the success rates are shown in **Table 2**. As previously discussed, each of the treatment groups have a different composition and the test size for the Educational Visit treatment was smaller than the other two, hence the counts of success are not directly comparable to one another. While comparing the success rate is more accurate, it is still somewhat problematic in that each treatment had different constraints. However, if the treatments were to become operational, the constraints would remain the same. It is therefore beneficial to RPO to make the comparisons. The difference in success rates between the test and control groups were highest for the Educational Visit treatment with a twelve percentage point difference between the test and control groups. The Due Diligence Letter treatment had a success rate of seven percentage points and the Continuing Education Letter treatment had a success rate of eight percentage points. While these two are not statistically significantly different from one another, they are significantly lower than the Educational Visit success rate.

While not shown in the table, the comparison of preparers considered successful who took continuing education was looked at for the Continuing Education test and control groups. Currently, data available to RPO does not show the topic or the actual date of the

continuing education, but there was a five percentage point difference in the percent of preparers who took continuing education between the test group (31 percent) and control group (26 percent).

While all three treatments had a significantly higher success rate than their respective control groups, the success rate of the control groups is of interest and is currently being analyzed by RPO. It could be a function of the underlying selection model, the (arbitrary) definition of success, or likely, both. Prior to the treatments, RPO did communicate with various professional organizations (e.g., American Institute of Certified Public Accountants (AICPA), National Association of Enrolled Agents (NAEA), and others) about RPO’s planned efforts. Hence the control group success rate could also, in part, be due to the Hawthorne effect, whereby individuals improve their behavior in response to knowing they are being observed.

**Table 2. Comparison of Success between the Treatment Groups and Control Groups**

<b>Treatment</b>	<b>Test</b>	<b>Control</b>	<b>Difference</b>
Number of Successes			
Educational Visit	575	424	151
Due Diligence Letter	573	490	83
Continuing Education Letter	572	475	97
Success Rate			
Educational Visit	53%	41%	12%
Due Diligence Letter	47%	40%	7%
Continuing Education Letter	47%	39%	8%

While it has already been demonstrated that none of the treatments caused a significant number of preparers to stop preparing returns, it is of interest to see if the treatments had an impact on the overall volume of returns prepared and, more specifically, the number of Schedule C returns. Included in the next set of tables is the change in the percent of Schedule C returns flagged by the error detection model. Overall results are shown, as well as results broken down by the success of the treatment.

The return volume information for the Educational Visit treatments is shown in **Table 3a**. Preparers in the cases where the treatment was considered a success tended to have a lower average number of returns than in the cases where the treatment was not considered to be successful (an average of 250 and 190, respectively), but the treatment had no statistically significant impact on the mean number of returns prepared before and after the treatment was applied.

The average number of Schedule C returns (70) for preparers where the treatment was a success was also lower than the average for preparers where it was not successful (100). For the successful cases, there was a significant drop (six percent) in the average number of Schedule C returns after the treatment, whereas the control group had a four percent increase.

The overall drop in the percent of returns flagged by the error detection model was nine percentage points for the test group compared to 5 percentage points for the control group. In cases where the treatment was successful, the drop was 19 percentage points for the test and 17 percentage points for the control. While these are not very different, it is important to recall that there were a larger number of successes in the test group. So there appears to be more success in the test but when it was a success it looked the same.

**Table 3a. Educational Visit: Return Volumes Before and After Treatment by Success of Treatment (excludes preparers who did not prepare returns both years)**

	<b>Treatment Outcome</b>	<b>Mean</b>	
		<b>Test</b>	<b>Control</b>
<b>Number of Returns Before Treatment</b>	<b>Overall</b>	<b>290</b>	<b>270</b>
	<i>Success</i>	250	190
	<i>Non-success</i>	340	330
<b>Percent Change After Treatment</b>	<b>Overall</b>	<b>-1%</b>	<b>3%</b>
	<i>Success</i>	0%	2%
	<i>Non-success</i>	-1%	3%
<b>Number of Sch. C Before Treatment</b>	<b>Overall</b>	<b>80</b>	<b>80</b>
	<i>Success</i>	70	60
	<i>Non-success</i>	100	100
<b>Percent Change After Treatment</b>	<b>Overall</b>	<b>-5%</b>	<b>4%</b>
	<i>Success</i>	-6%	4%
	<i>Non-success</i>	-3%	4%
<b>Percentage Point Change in Percent of Sch C Returns Flagged Before and After Treatment</b>	<b>Overall</b>	<b>9%</b>	<b>5%</b>
	<i>Success</i>	19%	17%
	<i>Non-success</i>	-3%	-4%

The same general pattern was found with the Due Diligence Letter treatment as shown in **Table 3b**, and the Continuing Education Letter, shown in **Table 3c**. There was no significant drop in the average number of returns but the preparers deemed successful did have a significant drop in the average number of Schedule C returns they prepared. While the pre-determined percentage point change in the percent of returns flagged by the error detection model was set at five percentage points, the actual average drop was more than three times that with 19, 17, and 18 percentage point decline for successful

preparers in the Educational Visit, Due Diligence Letter, and Continuing Education Letter treatments, respectively.

**Table 3b. Due Diligence Letter: Return Volumes Before and After Treatment by Success of Treatment (excludes preparers who did not prepare returns both years)**

	Treatment Outcome	Mean	
		Test	Control
<b>Number of Returns Before Treatment</b>	<b>Overall</b>	<b>280</b>	<b>280</b>
	<i>Success</i>	210	200
	<i>Non-success</i>	340	330
<b>Percent Change After Treatment</b>	<b>Overall</b>	<b>0%</b>	<b>1%</b>
	<i>Success</i>	-4%	-2%
	<i>Non-success</i>	2%	2%
<b>Number of Sch. C Before Treatment</b>	<b>Overall</b>	<b>80</b>	<b>80</b>
	<i>Success</i>	60	60
	<i>Non-success</i>	100	90
<b>Percent Change After Treatment</b>	<b>Overall</b>	<b>-4%</b>	<b>3%</b>
	<i>Success</i>	-11%	0%
	<i>Non-success</i>	-1%	4%
<b>Average Percentage Point Change in Percent of Sch C Returns Flagged Before and After Treatment</b>	<b>Overall</b>	<b>6%</b>	<b>4%</b>
	<i>Success</i>	17%	17%
	<i>Non-success</i>	-4%	-3%

**Table 3c. Continuing Education Letter: Return Volumes Before and After Treatment by Success of Treatment (excludes preparers who did not prepare returns both years)**

	Treatment Outcome	Mean	
		Test	Control
Number of Returns Before Treatment	<b>Overall</b>	<b>280</b>	<b>280</b>
	<i>Success</i>	220	200
	<i>Non-success</i>	340	330
Percent Change After Treatment	<b>Overall</b>	<b>-1%</b>	<b>0%</b>
	<i>Success</i>	-3%	-4%
	<i>Non-success</i>	0%	2%
Number of Sch. C Before Treatment	<b>Overall</b>	<b>80</b>	<b>80</b>
	<i>Success</i>	60	60
	<i>Non-success</i>	100	100
Percent Change After Treatment	<b>Overall</b>	<b>-5%</b>	<b>2%</b>
	<i>Success</i>	-11%	-5%
	<i>Non-success</i>	-2%	4%
Average Percentage Point Change in Percent of Sch C Returns Flagged Before and After Treatment	<b>Overall</b>	<b>6%</b>	<b>4%</b>
	<i>Success</i>	18%	18%
	<i>Non-success</i>	-4%	-5%

*Success Rates by Preparer Characteristics*

This section looks at success rates within each test by various characteristics of the preparer. The characteristics explored are the volume of returns prepared and signed by the preparer, the volume of Schedule C returns, the self-reported credential of the preparer and the number of years the preparer has been using their PTIN. For each test, the percent of preparers in every category is included for context.

Characteristics of preparers in the Educational Visit test are shown in **Table 4a**. Within the test group, preparers with the largest volume of returns (300 returns or more) had the

lowest rate of success at 40 percent, but they had the largest difference compared to the control group with 19 percentage points more. The next category with the largest difference, 13 percentage points, was the lowest volume preparers (100 returns or less), which also had the highest success rate of 63 percent within the test group. A similar pattern is found when looking at the volume of Schedule C returns. Enrolled Agents responded to the Educational Visit best with a 19 percentage point difference between the test and control groups. Finally, those who had held a PTIN for six years or more did not respond to the Educational Visit as well as those who had held it for five years or less.

**Table 4a. Educational Visit: Success Rates by Various Characteristics of Preparer**

	Percent of Test in Category	Success Rate		
		Test	Control	Difference
Number of Preparers	100%	53%	41%	12%
Volume of Returns Prepared				
100 returns or less	21%	63%	50%	13%
101-200 returns	33%	58%	49%	9%
201-300 returns	22%	51%	40%	11%
301 returns or more	24%	40%	21%	19%
Volume of Returns w/ Sch. C				
30 returns or less	25%	68%	54%	14%
31-50 returns	24%	56%	47%	9%
51-100 returns	29%	50%	38%	12%
101 returns or more	22%	37%	22%	15%
Credential				
CPA / Attorney	17%	56%	48%	8%
Enrolled Agent	11%	54%	36%	19%
No Credential	72%	52%	41%	12%
Age of PTIN				
2 or less years	41%	56%	42%	13%
3-5 years	14%	58%	44%	14%
6 or more years	45%	49%	39%	10%

The same comparison for the Due Diligence Letter test follows in **Table 4b**. Within this test, preparers with a volume of 101–200 returns responded best both in terms of the success rate and the percentage point difference compared to the control group with 56 percent and 10 percentage points, respectively. Preparers with the highest volume of Schedule C returns (101 returns or more) responded the least to the letter with a 28 percent success rate that was 5 percentage points higher than the control group. Preparers with the smallest volume of Schedule C returns saw a 59 percent success rate, but it was

only 6 percentage points higher than the control group. The highest percentage point difference was for those who prepared 51–100 Schedule C returns (9 percentage points). CPAs and preparers who have held a PTIN for 6 or more years responded well to this treatment with 12 and 11 and percentage point differences compared to the control, respectively.

**Table 4b. Due Diligence Letter: Success Rates by Various Characteristics of Preparer**

	Percent of Test in Category	Success Rate		
		Test	Control	Difference
Number of Preparers	100%	47%	40%	7%
Volume of Returns Prepared				
100 returns or less	21%	52%	50%	2%
101-200 returns	36%	56%	46%	10%
201-300 returns	22%	44%	38%	7%
301 returns or more	22%	29%	23%	6%
Volume of Returns w/ Sch. C				
30 returns or less	24%	59%	53%	6%
31-50 returns	28%	51%	43%	8%
51-100 returns	26%	47%	38%	9%
101 returns or more	22%	28%	23%	5%
Credential				
CPA / Attorney	17%	58%	47%	12%
Enrolled Agent	9%	44%	34%	10%
No Credential	74%	44%	39%	5%
Age of PTIN				
2 or less years	43%	46%	42%	3%
3-5 years	17%	47%	42%	5%
6 or more years	41%	48%	37%	11%

Finally, the results for the Continuing Education Letter test are shown in **Table 4c**. While this test had the most success with mid volume level preparers that prepared 101-200 returns (57 percent success rate that was 12 percentage points higher than the control group), those with the smallest number of Schedule C returns (30 returns or less) had the best response to the letter with a success rate of 63 percent that was 11 percentage points higher than the control. As explained earlier, CPAs and Attorneys were excluded from this test. There were no significant differences in response to this letter between Enrolled Agents and non-credentialed preparers. Preparers who had been using a PTIN for two years or less had the lowest response to the letter while those who had held a PTIN for

three to five years responded best with a 52 percent success rate that was 12 percentage points higher than the control.

**Table 4c. Continuing Education Letter: Success Rates by Various Characteristics of Preparer**

	Percent of Test in Category	Success Rate		
		Test	Control	Difference
Number of Preparers	100%	47%	39%	8%
Volume of Returns Prepared				
100 returns or less	22%	54%	48%	6%
101-200 returns	32%	57%	46%	12%
201-300 returns	22%	44%	39%	5%
301 returns or more	24%	31%	20%	11%
Volume of Returns w/ Sch C				
30 returns or less	23%	63%	52%	11%
31-50 returns	27%	54%	44%	10%
51-100 returns	27%	42%	38%	3%
101 returns or more	23%	31%	21%	10%
Credential				
CPA / Attorney*	n.a.	n.a.	n.a.	n.a.
Enrolled Agent	10%	44%	36%	8%
No Credential	90%	48%	39%	8%
Age of PTIN				
2 or less years	42%	46%	40%	5%
3-5 years	17%	52%	40%	12%
6 or more years	41%	47%	37%	10%

\* CPAs and Attorneys were excluded from this test. Line is included to be consistent with previous tables.

Comparing between treatments, low volume preparers responded best to the Educational Visit (13 percentage point difference) but not significantly different from the control group to the Due Diligence Letter (2 percentage point difference). Preparers with 101-200 returns responded to each of the treatments, however those with 201-300 returns responded better to the visit (11 percentage point difference) than to either letter (7 percentage point difference for the Due Diligence Letter and 5 percentage point difference for the Continuing Education Letter). High volume preparers responded well to the Educational Visit and the Continuing Education Letter, but not the Due Diligence Letter (19, 11 and 6 percentage point difference, respectively). The Schedule C volumes roughly followed the same pattern as the overall volume categories. CPAs responded better to the Due Diligence Letter (12 percentage point difference) than to the

Educational Visit (8 percentage point difference) whereas Enrolled Agents and non-credentialed preparers responded better to the visit (19 and 12 percentage point difference, respectively). Relatively new PTIN holders (two years or less) responded well to the Educational Visit (13 percentage point difference) but not to the Due Diligence Letter (3 percentage point difference) or the Continuing Education Letter (5 percentage point difference). Those who have held a PTIN for six years or more responded to each treatment about the same.

The most successful treatment was the Educational Visit for high volume preparers and Enrolled Agents. The least successful groups for this treatment were mid-level volume preparers and CPAs. The Due Diligence Letter treatment was most effective for credentialed preparers who have held a PTIN for a relatively long time. However, it was the least successful treatment for low volume preparers and was also ineffective for relatively new PTIN holders. The Continuing Education Letter was most effective for preparers that prepare 101-200 returns and have held their PTIN 3-5 years, but was ineffective for preparers with 51-100 Schedule C returns.

#### *Client Migration*

While preparers were the primary focus of the treatments, RPO was also interested in what happened with their clients. As mentioned earlier, there is an issue of endogeneity when trying to determine if the taxpayer or the preparer is driving the non-compliance. To the extent it is the taxpayer driving the non-compliance, then for successfully treated preparers, one might expect these clients to discontinue using that preparer and either go to a new preparer or prepare the return themselves. Clients of successfully treated preparers who were flagged by the error detection model in both years actually had the same or lower migration rates in the test group than the control group for all three treatments. For the Educational Visit the test group rate was 22 percent compared to the control group rate of 26 percent. For the Due Diligence Letter, the rate was 27 percent for both the test and control groups. Finally, the rates for the Continuing Education Letter were 30 percent and 33 percent for the test and control groups, respectively. Thus, it appears that the model was effective in identifying preparers who were driving the non-compliance.

#### **Summary**

The results of the tests indicate all three treatments were effective. However, while the Educational Visit treatment was more effective than either letter, it is not as cost effective. While it is difficult to get an exact cost for each treatment, all that is really needed to determine the cost effectiveness is the order of magnitude of the costs. Roughly speaking, each letter costs less than \$1, while each visit costs more than \$100. Since the Educational Visit was not 100-fold more effective, it is not as cost effective as the letters.

The model appears to have worked as expected in identifying preparers who were the driver of the Schedule C errors, but refinements to the model will be made based on results and lessons learned from the study.

The results presented here were from the first year of the study. In the second year of the study, an additional letter was tested that was purely educational, and since results from the first year were not yet available, the original three treatments were reemployed in the second year as well. Results from year two will be reported out once the data become available. In addition, RPO plans to continue to undertake similar studies in future years to confirm and refine the results.

## References

Brown, Robert E. and Mark J. Mazur. "IRS's Comprehensive Approach to Compliance Measurement", Presentation to the National Tax Association Spring Symposium, Washington, DC (May, 2003)

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science+Business Media, LLC, New York, NY, 2009.

Hicks, Charles R. *Fundamental Concepts in the Design of Experiments*. Oxford University Press, New York, NY. 1993.

Internal Revenue Service Data Book, Fiscal Year 2012

Mazur, Mark J. and Alan H. Plumley. "Understanding the Tax Gap", Presentation to the National Tax Association Spring Symposium, Washington, DC (May, 2007)

Treasury Inspector General for Tax Administration. "Implementation of the Return Preparer Visitation Project Was Successful, but Improvements Are Needed to Increase Its Effectiveness", June, 2012

# **Appendix A**

## **Due Diligence Letter**



Department of Treasury  
Internal Revenue Service  
Return Preparer Office  
1122 Town and Country Commons  
Chesterfield MO 63017

Letter	XXXX
Date	Date
To contact us	Phone 1-636-255-1208 8 a.m.- 5 p.m. CT

Tax Return Preparer  
Address  
City, ST zip

**Subject: Reminders about Schedule C preparation**

Dear Tax Return Preparer,

A review of tax returns you have prepared in the past year shows that many have a high percentage of traits we believe typically indicate errors in preparing Form 1040, Schedule C, Profit or Loss from Business (Sole Proprietorship). This letter is to remind you of your responsibilities in this area, provide educational assistance, and request that you pay special attention to it next filing season.

**Due diligence responsibilities**

A paid tax return preparer is expected to take multiple steps to prepare accurate tax returns on behalf of clients. These include reviewing the applicable tax law, and establishing the relevancy and reasonableness of income, credits, expenses, and deductions to be reported on the return. In general, you may rely in good faith without verification upon information furnished by the client. However, you may not ignore the implications of information furnished to, or actually known by you. You must make reasonable inquiries if the information appears to be incorrect, inconsistent with an important fact or another factual assumption, or incomplete.

**Schedule C reminders**

To prepare accurate Schedules C, you should ask your clients sufficient questions to determine that the expenses claimed are correct and allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they are entitled to claim deductions for non-qualifying expenditures. You should also ask your clients if they have documentation to support the expenses in case receipts are requested by the IRS.

**Helpful resources**

Based on our analysis of the Schedules C you prepared, we encourage you to review the Schedule C instructions and other IRS publications available at [www.irs.gov](http://www.irs.gov), keyword: Recommended Reading for Small Businesses. We also encourage you to review Circular 230, Regulations Governing Practice before the Internal Revenue Service, sections 10.22 and 10.34, titled “Diligence as to accuracy” and “Standards with respect to tax returns and documents, affidavits and other papers”, respectively.

## **Potential consequences**

In the future, both you and your clients may be adversely affected by incorrect returns. Consequences may include any or all of the following:

- If your clients' returns are examined and found to be incorrect, your clients may be liable for additional tax, interest, and penalties.
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to an unreasonable position can be assessed a penalty of at least \$1,000 per return (IRC section 6694(a)).
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to reckless or intentional disregard of rules or regulations by the tax preparer, can be assessed a penalty of at least \$5,000 per return (IRC section 6694(b)).

We hope this letter has heightened your awareness of your responsibilities as a paid tax return preparer and provided you with information on preparing accurate Schedules C for your clients.

Sincerely,

Carol A. Campbell  
Director, Return Preparer Office

# **Appendix B**

## **Continuing Education Letter**



Department of Treasury  
Internal Revenue Service  
Return Preparer Office  
1122 Town and Country Commons  
Chesterfield MO 63017

Letter	XXXX
Date	Date
To contact us	Phone 1-636-255-1208 8 a.m.- 5 p.m. CT

Tax Return Preparer  
Address  
City, ST zip

**Subject: Recommendation to take CE programs about Schedule C**

Dear Tax Return Preparer,

A review of tax returns you have prepared in the past year shows that many have a high percentage of traits we believe typically indicate errors in preparing Form 1040, Schedule C, Profit or Loss from Business (Sole Proprietorship). Therefore you may benefit from continuing education (CE) programs on this topic.

The purpose of this letter is to recommend that as part of your **2013** CE requirement, you take a **minimum of 4 hours** of programs related to business income and/or expenses. Information about CE requirements and a list of IRS approved CE providers is available at [www.irs.gov/taxpros/ce](http://www.irs.gov/taxpros/ce).

In addition, as you prepare returns for the next filing season, please pay special attention to your work on Schedule C returns to ensure they are prepared accurately. In particular, make sure you have familiarized yourself with the following areas:

**Due diligence responsibilities**

A paid tax return preparer is expected to take multiple steps to prepare accurate tax returns on behalf of clients. These include reviewing the applicable tax law, and establishing the relevancy and reasonableness of income, credits, expenses, and deductions to be reported on the return. In general, you may rely in good faith without verification upon information furnished by the client. However, you may not ignore the implications of information furnished to, or actually known by you. You must make reasonable inquiries if the information appears to be incorrect, inconsistent with an important fact or another factual assumption, or incomplete.

**Schedule C reminders**

To prepare accurate Schedules C, you should ask your clients sufficient questions to determine that the expenses claimed are correct and allowable. Taxpayers may not fully understand the tax laws and may incorrectly believe they are entitled to claim deductions for non-qualifying expenditures. You should also ask your clients if they have documentation to support the expenses in case receipts are requested by the IRS.

## **Helpful resources**

Based on our analysis of the Schedules C you prepared, in addition to considering CE programs, we encourage you to review the Schedule C instructions and other IRS publications available at [www.irs.gov](http://www.irs.gov), keyword: Recommended Reading for Small Businesses. We also encourage you to review Circular 230, Regulations Governing Practice before the Internal Revenue Service, sections 10.22 and 10.34, titled “Diligence as to accuracy” and “Standards with respect to tax returns and documents, affidavits and other papers”, respectively.

## **Potential consequences**

In the future, both you and your clients may be adversely affected by incorrect returns. We will check whether you complete the continuing education on business income and/or expenses as recommended and we will be looking for improvements in returns you prepare. Incorrect returns may cause any of the following consequences:

- If your clients’ returns are examined and found to be incorrect, your clients may be liable for additional tax, interest, additions to tax and penalties.
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to an unreasonable position can be assessed a penalty of at least \$1,000 per return (IRC section 6694(a)).
- Tax return preparers who prepare a client return for which any part of an understatement of tax liability is due to reckless or intentional disregard of rules or regulations by the tax preparer, can be assessed a penalty of at least \$5,000 per return (IRC section 6694(b)).

We hope this letter and your continuing education focus on Schedule C preparation will heighten your awareness of your responsibilities and help ensure you prepare accurate Schedules C for your clients.

Sincerely,

Carol A. Campbell  
Director, Return Preparer Office