Tax Administration at the Centennial: An IRS-TPC Research Conference

Papers Given at the
2013 IRS-Tax Policy Center Research Conference

Held at the Urban Institute
Washington, DC
June 20, 2013

Compiled and edited by Alan Plumley*
Research, Analysis, and Statistics, Internal Revenue Service

*Prepared under the direction of Barry W. Johnson, Acting Director, Statistics of Income Division

Foreword

This edition of the IRS Research Bulletin (Publication 1500) features selected papers from the IRS-Tax Policy Center (TPC) Research Conference "Tax Administration at the Centennial," held at the Urban Institute in Washington, DC, on June 20, 2013. Conference presenters and attendees included researchers from all areas of the IRS, officials from other government agencies, and academic and private sector experts on tax policy, tax administration, and tax compliance. In addition to those who attended in person, many participated live online, as the TPC broadcast video of the proceedings over the Internet. The videos are archived on their website to enable additional participation. Online viewers participated in the discussions by submitting questions via e-mail as the sessions proceeded.

The conference began with welcoming remarks by Eric Toder, Co-Director of the Tax Policy Center and by Pat McGuire, the IRS Deputy Director of Research, Analysis, and Statistics. The remainder of the conference included sessions on individual income tax dynamics, business compliance behavior, corporation income tax enforcement, and lessons from other tax administrations. Eric Toder also provided insights on tax reform.

We trust that this volume will enable IRS executives, managers, employees, stakeholders, and tax administrators elsewhere to stay abreast of the latest trends and research findings affecting tax administration. We anticipate that the research featured here will stimulate improved tax administration, additional helpful research, and even greater cooperation among tax administration researchers worldwide.

Acknowledgments

The IRS Research Conference was the result of substantial effort and preparation over a number of months by many people. The conference program was assembled by a committee representing research organizations throughout the IRS. Members of the program committee included: Alan Plumley, John Guyton, Drew Johns, Rahul Tikekar, Hellon Moore, and Leann Weyl (National Headquarters Office of Research); Zhicheng Swift (Office of Program Evaluation and Risk Analysis); Jael Jackson (Statistics of Income); Mike Gilvary (Small Business and Self-Employed); Javier Framiñan (Wage and Investment); Fran Cappelletti (Taxpayer Advocate); Christina Hui (Criminal Investigation); Charles Boynton (Large and Midsize Business); and Elaine Maag (Tax Policy Center). In addition, Blake Greene from the Tax Policy Center oversaw numerous details to ensure that the conference ran smoothly

This volume was prepared by Paul Bastuscheck, Lisa Smith, and Camille Swick (layout and graphics) and Beth Kilss and Georgette Walsh (editors), all of the Statistics of Income Division. The authors of the papers are responsible for their content, and views expressed in these papers do not necessarily represent the views of the Department of the Treasury or the Internal Revenue Service.

We appreciate the contributions of everyone who helped make this conference a success.

Janice M. Hedemann
Director, National Headquarters Office of Research
Chair, 2013 IRS-TPC Research Conference

Tax Administration at the Centennial: An IRS-TPC Research Conference

Contents

4. Lessons From Other Tax Administrations

5. Appendix

1

Individual Income Tax Dynamics

Burman ◆ Coe ◆ Pierce ◆ Tian

Langetieg ◆ Payne ◆ Vigil

Jones

Older Taxpayers' Response to Taxation of Social Security Benefits

Leonard Burman, Syracuse University and Tax Policy Center; Norma B. Coe, University of Washington and NBER; Kevin Pierce, Internal Revenue Service; and Liu Tian, Syracuse University[1]

I. Introduction

Social Security benefits are taxed under a complex regime that raises marginal effective tax rates by up to 85 percent. Over a range of Modified Adjusted Gross Income (MAGI),[2] affected taxpayers must include in their taxable income $0.50 of Social Security benefit for every additional dollar of other taxable income;[3] at higher income levels, $0.85 of benefits must be added, until 85 percent of Social Security benefits are included. In these income ranges, an additional dollar of other taxable income increases total taxable income by $1.50 or $1.85. At the highest income levels, this can convert a modest 25 percent statutory tax rate into a 46.25 percent marginal rate. This is much higher than the top income tax bracket,[4] but it applies to older households with relatively modest incomes.

The tax on benefits is in some ways similar to the Social Security earnings test (SSET), which reduces Social Security benefits by 50 cents for every dollar earned above an exempt amount for those younger than the Full Retirement Age (FRA, currently 66).[5] However, the taxation of benefits applies at all ages while the SSET applies only to Social Security recipients who claim benefits before reaching FRA. Moreover, unlike the benefit tax, the SSET is not a pure tax since the reduced current benefits translate into higher benefits once FRA is reached. In contrast, the tax on benefits has no actuarial adjustment.

While the tax on benefits could have significant effects on behavior, it has been thus far largely ignored in the literature. This is a potentially important oversight. If taxpayers understand the rules, one would expect them to be even more sensitive to this work disincentive than to the SSET, which most research has found to significantly affect labor supply. Moreover, this tax not only affects earnings but also nonlabor income, so it can influence nonlabor decisions, such as when to realize capital gains. Early retirees may be subject to both the SSET and Social Security benefit taxation, so the effective combined work disincentive may be quite large. Further, if the tax is inefficient, reform options might exist that could bolster the trust fund, extend older people's attachment to the labor force, significantly reduce tax compliance costs for older workers, and raise overall economic welfare.

This paper investigates older taxpayers' response to the taxation of Social Security benefits by looking for evidence of bunching at the kink points created by the taxation of benefits. In theory, some individuals with incomes above the taxation thresholds have an incentive to reduce their incomes to the threshold—by working less, delaying realization of capital gains, or using other techniques to reduce reported income. We test this hypothesis using a panel of data from individual income tax and information returns.

[1] Some of the research reported herein was pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium (RRC) through the Boston College Center for Retirement Research (CRR). The findings and conclusions expressed are solely those of the authors and do not represent the views of the SSA, any agency of the Federal Government, the RRC, the Urban Institute, Syracuse University or Boston College. The authors would like to thank: Zhenya Karamcheva for programming assistance; Rachel Johnson for supplying estimates from the IRS public-use file; Dan Feenberg for help with TAXSIM; Jan Ondrich, John Sabelhaus, seminar participants at the Center for Policy Research at the Maxwell School of Syracuse University and at the IRS—Tax Policy Center 2013 Joint Research Conference for helpful comments.

[2] MAGI includes most of the income and adjustments reflected in adjusted gross income (AGI), but it includes one-half of Social Security benefits, rather than the taxable portion. It also includes tax-exempt interest.

[3] That is, any taxable income included in MAGI other than Social Security benefits.

[4] In 2013, the top income tax bracket is 39.6 percent and applies to households with taxable incomes over $450,000 (married) and $400,000 (single).

[5] A 33-percent reduction and a higher exemption apply to workers in the year in which they reach FRA.

We find no evidence of bunching at or around the thresholds for the population as a whole, and only a very small response for single self-employed taxpayers who have previously been found to be more sensitive to changes in tax rates (Saez 2010; Chetty *et al*. 2011). This implies that this complicated tax does not lead to any important behavioral response and therefore imposes little or no deadweight loss.

The paper continues as follows. Section II describes the taxation on Social Security Benefits. Section III surveys the relevant literature. Section IV develops a simple theoretical model. Section V discusses the data and Section VI presents the empirical results. Section VII summarizes our findings and discusses planned future work.

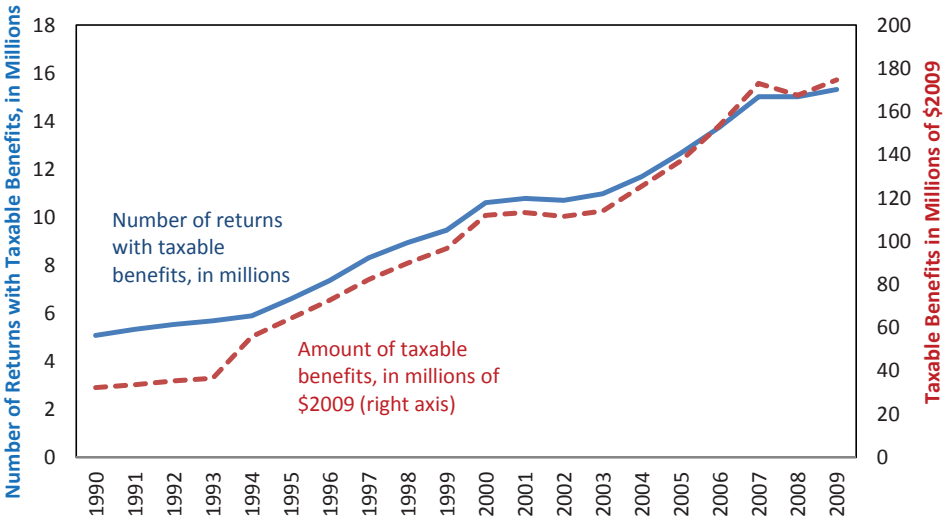
II. Taxation of Social Security Benefits

Prior to 1983, Social Security benefits were not subject to income tax. In 1983, the Greenspan Commission recommended that a portion of benefits be subject to income taxation, with the resulting additional tax revenue allocated to the OASDI (Old Age Survivors and Disability Insurance, or Social Security) trust fund. Legislation enacted in 1993 increased the amount of benefits included in taxable income for higher-income taxpayers, with the additional revenues allocated to the HI (Medicare) trust fund.

The formula for taxation is complex. OASDI benefits become subject to income taxation when MAGI exceeds $25,000 for single ($32,000 for married) taxpayers. Above those thresholds, the taxable portion of benefits phases in starting at a 50-percent rate. Fifty cents of benefits are included in taxable income for every additional dollar of MAGI. After a second threshold ($34,000 for singles and $44,000 for married households), the phase-in rate increases to 85 percent. The phase-in continues until 85 percent of Social Security benefits are included in taxable income.

The thresholds for taxation have been fixed in nominal terms since their inception. Since the thresholds are not adjusted for inflation, they decrease in real terms over time, unlike Federal income tax brackets and many other income tax parameters. As a result, taxation of Social Security affects an increasing proportion of beneficiaries over time, pushing people into higher tax brackets. The number of returns with taxable Social Security benefits nearly tripled—from 5.3 million to 15.3 million—between 1990 and 2009 (see Figure 1). The dollar amount of Social Security benefits subject to taxation increased even more, from $33.6 million in 1990 to $174.6 million in 2009, in part because of the 1993 legislation and partly because of increases in nominal income of the elderly.

FIGURE 1. Number of Returns with Taxable Social Security Benefits, and in Millions of $2009, in Millions, 1990–2009



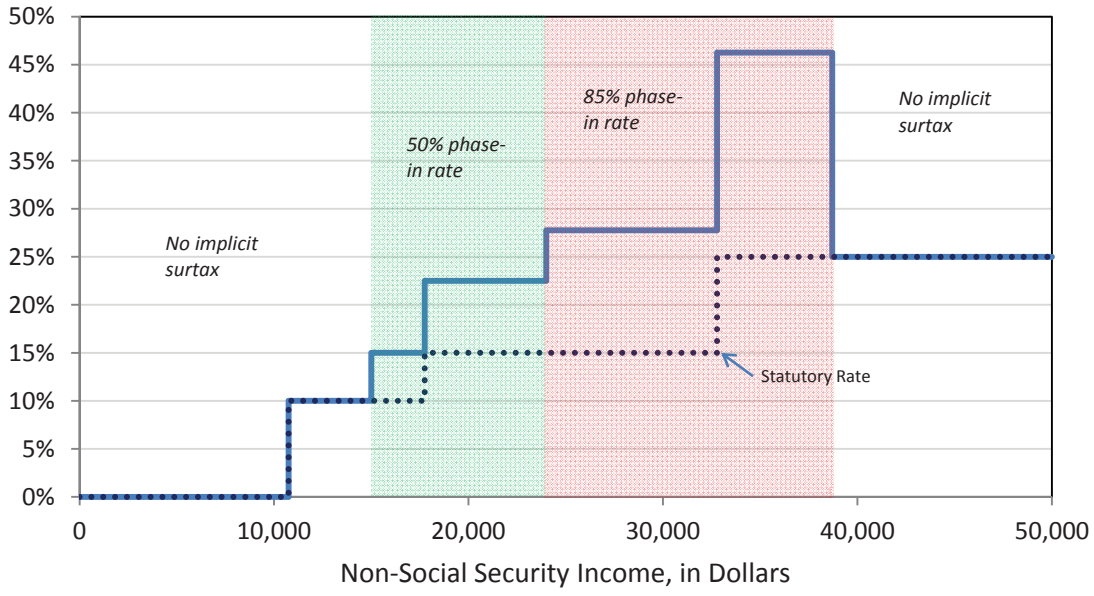
Source: IRS, "Selected Income and Tax Items for Selected Years (in Current and Constant Dollars)," http://www.irs.gov/file_source/pub/irs-soi/09intba.xls.

Although the levy may seem to be a tax on Social Security benefits, it is actually a large implicit surtax on all income included in MAGI.[6] Taxpayers with low Social Security benefits or modest amounts of other income have MAGI below the threshold for taxation and are not affected. However, as either benefits or other income increase, effective marginal tax rates may increase quite dramatically. For example, a single person with $15,000 of non-Social Security income and $19,900 of Social Security benefits has none of her Social Security included in taxable income; her marginal income tax rate equals the statutory rate of 10 percent. If either her Social Security benefit or income increases by $100, her marginal tax rate would increase to 15 percent.

The taxation of Social Security benefits increases *effective* marginal tax rates by 50 percent in the first phase-in range and by 85 percent in the second. This is because an additional dollar of AGI (earnings or non-labor income) increases MAGI by $1.50 in the 50-percent phase-in range and by $1.85 in the higher interval, until 85 percent of Social Security benefits are included in taxable income. Figure 2 illustrates how the taxation of benefits distorts effective tax rates for a taxpayer with $20,000 in Social Security benefits in 2010. The effective tax rate schedule is marked by significant discontinuities—much larger than under the regular income tax. Over the phase-in range of income, a taxpayer would ordinarily face three marginal rates—10, 15, and 25 percent. However, because of the partial inclusion of Social Security benefits, three additional effective rates are created—22.5, 27.75, and 46.25 percent. The top effective rate, which applies to seniors with relatively modest incomes ($33,000-$39,000 in Figure 2), is actually higher than the top statutory income tax rate of 35 percent that applied to households with taxable income over $373,650 in 2010.

As shown in Figure 2, taxpayers with income just beyond the phase-in region face a marginal rate of 25 percent, which is more than 20 percentage points lower than those with lower incomes. Taxation of benefits reduces their after-tax income, but there is no implicit surtax or marginal disincentive to work or earn other income.

FIGURE 2. Effective Marginal Tax Rates for Single Non-Itemizer, Age 66 or Older, With $20,000 of Social Security, by Non-Social Security Income, 2010



[6] Note that the tax potentially applies to taxpayers collecting disability and survivor benefits under the OASDI program, but our analysis will focus on Social Security beneficiaries.

The implicit tax affects not only earnings but also nonlabor income. Burman (1999) points out that the taxation of Social Security can have disproportionate effects on effective long-term capital gains tax rates; it can add up to 21.25 percentage points (85 percent of 25 percent) to the statutory capital gains tax rate of 15 percent that applies to taxpayers in that income range.

If there is a behavioral response to the taxation of benefits, the substantial kinks in the tax schedule could create clustering of households at the kink points, and potentially discourage labor supply at both the extensive and intensive margins. Although taxpayers with very low and very high nonlabor income are likely to be unaffected, taxpayers whose earnings would be subject to partial taxation might be less likely to work than other similar taxpayers. Secondary earners may face especially strong disincentives if the primary earner's income puts the second earner in the phase-in range.

The tax treatment of benefits could also affect decisions about when to begin claiming Social Security. The steeply rising marginal tax rate schedule creates an incentive for many people to claim benefits early, getting a reduced benefit over more years. Individuals born after 1942 can reduce their annual benefit by 25 percent or more by claiming at age 62 rather than the full retirement age and fully or partially avoid taxation of Social Security benefits. As a result, the adjustment for delayed retirement may no longer be actuarially fair when taxes are considered. On the other hand, some taxpayers may have an incentive to delay claiming Social Security benefits. If a worker reaches the full retirement age and expects to keep working for a few more years after which his non-Social Security income would drop significantly, he may elect to delay claiming Social Security benefits if the future drop in income means that much less of his benefits would be subject to tax. In this case, the after-tax value of delaying retirement is better than actuarially fair, even if before tax, the tradeoff is neutral.[7]

Finally, it should be noted that the very complicated taxation of Social Security benefits might affect behavior much differently than predicted by a pure optimizing model. It is possible that people do not understand how the tax affects marginal tax rates, the incentives on labor supply, or the timing of benefits. If people ignore these incentives, then the tax may be a type of optimal tax—raising revenue with little or no effect on behavior. On the other hand, taxpayers may overreact to misunderstood incentives—magnifying the economic distortion.

III. Previous Literature

While Social Security has been extensively studied, very little attention has been paid to the taxation of benefits. The closest analogue is the SSET, which reduces Social Security benefits for individuals who have not reached the full retirement age and whose earnings exceed a threshold.[8] The SSET is different in several key ways. For one thing, it is much easier for individuals to determine if they are affected since it depends only on individual earnings and age. In another sense, though, it is more complicated because there is an actuarial adjustment. The reduced Social Security benefits translate into higher future benefits (assuming the individual lives long enough to claim them) making labor supply decisions a function not only of the tax rate, but life expectancy and discount rates. Evidence, however, suggests that older workers view the SSET as a tax with little or no awareness of the actuarial adjustment. Several studies find evidence that the SSET discouraged work among older Americans.[9] Also, eliminating the earnings test for beneficiaries who had reached the full retirement age increased the likelihood that workers would claim Social Security benefits before age 70 (Song and Manchester 2007; Friedberg and Webb 2009).

The Social Security benefit formula itself impacts the implicit taxes on work. The formula is progressive, so those with high earnings get much less in additional benefits per dollar of payroll tax than those with lower

[7] Coile et al. (2002) model the timing of claiming Social Security. Even ignoring the taxation of Social Security benefits as they do, the decision is very complicated. They present nonlinear simulations for the case of a single earner, leaving the more complex case of dual earners to later research. They find that men generally claim benefits too early compared with the optimal choice.

[8] Prior to 2001, there was also a SSET at a reduced rate for individuals between the full retirement age and 69.

[9] Friedberg (2000), Gustman and Steinmeier (2005), Benitez-Silva and Heiland (2007), Song and Manchester (2007), Heider and Loughran (2008), Engelhardt and Kumar (2009), and Friedberg and Webb (2009). Burtless and Moffitt (1985), Gruber and Orszag (2003), Gustman and Steinmeier (1985) and Song and Manchester (2007) find small effects.

incomes. For some workers, including those who expect to have fewer than 40 covered quarters of work—and are thus ineligible for benefits—or who will receive benefits based on their spouse's earnings, the payroll tax is a pure tax. Liebman, Luttmer, and Seif (2009) find labor supply and retirement decisions of older workers to be sensitive to the variation in the effective tax rate on earnings. Of particular relevance, this research suggests a surprisingly sophisticated understanding of complex rules. A survey by Leibman and Luttmer (2012) finds a fair amount of knowledge of some Social Security provisions and relatively less about others (including the earnings test).

We know of only three previous studies that have examined the taxation of Social Security benefits. Liebman and Goodman (2008) look at the taxation of benefits as a form of means-testing and conclude that it is sub-optimal. They do not explicitly consider the effect of taxation of benefits on economic incentives, but, citing behavioral economics research, they question whether and how individuals might respond to the tax incentives:

While this analysis shows that the taxation of Social Security benefits raises marginal tax rates for a sizable minority of Social Security beneficiaries, the complexity of these provisions raises questions about how future and current beneficiaries perceive these incentives and whether their behavior responds to them. (Liebman and Goodman 2008, pp. 17–18)

One possibility is that, overwhelmed by the complexity of the incentives, taxpayers might simply ignore the tax. Alternatively, they might apply a simple rule of thumb—e.g., on average, 4 percent of Social Security benefits are included in income—that could similarly result in little distortion. Or, Liebman and Goodman (2008) conjecture, taxpayers may misperceive the tax as applying to 85 percent of Social Security benefits. This could create a quite large income effect—even for taxpayers with incomes so low that little or none of their benefits are taxable—although presumably it would have no effect on the perceived after-tax return to working or earning other income.

Page and Conway (2011) measure the income effect of taxation of benefits directly by exploiting the natural experiment of introduction of the taxation in 1983, using difference-in-differences methodology with data from the Current Population Survey (CPS). They estimate that a 20-percent reduction in after-tax Social Security benefits boosts labor force participation among high-income elderly by 2 to 5 percentage points. They argue that taxation of Social Security benefits increases labor supply through the income effect: people above the threshold where 85 percent of benefits are subject to tax, even before including OASDI benefits, have less after-tax income, which increases hours of work. They do not attempt to measure the marginal effect of reduced after-tax income within the phase-in range.

Burman, Coe, and Tian (2011) attempt to measure directly the effect of taxing Social Security benefits on labor force participation and earnings using data from the Health and Retirement Study (HRS). They do not find evidence that taxation of benefits significantly affects labor market behavior, but they raise the major caveat that their estimates may be unreliable because of errors in variables and small sample size. Survey estimates of tax information are notoriously imprecise and the HRS lacks key components of taxable income, such as capital gains.

IV. Effects of Taxing Social Security Benefits

If taxpayers understand how the taxation of Social Security benefits affects their budget, then we should observe bunching of MAGI near the thresholds. Taxing Social Security benefits generates convex kinks in the budget constraint at the thresholds for the 50-percent and 85-percent phase-in rates (corresponding to MAGI of $25,000 and $34,000 for single filers). In a simple model of utility maximization, taxpayers with incomes only slightly greater than the threshold will reduce their incomes to the threshold.

To see this, consider a simplified example in which there is a flat-rate income tax and only one rate of taxation of Social Security (as was the case between 1983 and 1993), which increases tax rates by 50 percent. The optimal level of MAGI will maximize utility subject to the kinked budget constraint (Figure 3). Assuming that individuals are averse to work and other activities that increase MAGI and that they value consumption (after-tax income), higher utility corresponds to indifference curves that move in a northwesterly direction on the figure.

Figure 3 illustrates three categories of taxpayers who will be affected differently by the introduction of taxation of benefits. In Panel A, MAGI in the absence of taxation of Social Security would fall below the threshold. That individual is unaffected by benefit taxation. Panel B shows a taxpayer who before the tax change would have MAGI of $z^*+\Delta z^*$, but after introduction of the taxation regime chooses MAGI of z^*. Saez (2010) shows that in the case where individuals have identical preferences but differ in their ability to earn income (e.g., their hourly wage rate differs), all individuals with initial incomes between z^* and $z^*+\Delta z^*$ would bunch at the kink. Taxpayers who initially have higher incomes than $z^*+\Delta z^*$ may also reduce their incomes, but their new incomes would be tangent to the new budget constraint to the right of z^*. Finally, Panel C depicts high-income taxpayers for whom the tax produces only an income effect.

With perfect information and complete ability to choose MAGI, this framework would produce bunching at the threshold z^* (see Figure 4). The kink has no effect on taxpayers with initial incomes below z^*, but it produces a leftward shift in the distribution of income among those with initial incomes above z^*. Saez (2010) extends this analysis to allow for adjustment frictions (e.g., people can only imperfectly adjust income or they have imperfect information about the location of the threshold) and shows that under certain simplifying assumptions, the amount of bunching near z^* provides a measure of the compensated elasticity of taxable income. If individuals are very sensitive to taxation (high elasticity), then there will be an unusually large mass of tax returns near the threshold.

There are many contexts in which such bunching may be observed. Saez (2010) shows that self-employed individuals' incomes tended to bunch at the level where the earned income tax credit starts to phase out. Wage earners showed no such response, which is consistent with the notion that the self-employed have more control over hours worked and taxable earnings, and self-employment income is not subject to third-party information reporting, making it easier to misreport on a tax return. Friedberg (2000), Song and Manchester (2007), Engelhardt and Kumar (2009), etc. observe that older workers clustered to the left of the SSET exempt threshold. Chetty, et al. (2011) examine bunching around large jumps in tax brackets in Denmark to measure elasticity of taxable income in the context of search costs.

Our hypothesis is that if taxpayers are aware of the incentives created by the taxation of Social Security benefits, there should be a bump in the empirical distribution of tax returns near the two thresholds for taxation. We would expect the bump to be more pronounced for those with income from self-employment.

V. Data

To look for evidence of bunching, we use administrative data—the 1999 IRS Statistics of Income (SOI) Individual Edited Panel, which is a longitudinal dataset drawn from individual income tax returns and information returns.[10] The data comprise a panel of individuals from Tax Years 1999 to 2008. The advantage of these data is that they provide an accurate measure of what is reported to the IRS on income tax returns—and thus tax status. They also allow us to study the behavior of self-employed individuals; those who previous research suggests would be the most responsive. The disadvantage is that the dataset includes little demographic information, which precludes structural modeling of the response to taxation.

The panel has been augmented by matching all of the primary and secondary SSNs within the panel to the SOI-processed information returns databases for Forms W-2 (information on wages and withholdings), Forms 5498 (contributions to retirement accounts), Forms 1099-SSA (Social Security benefits), and Forms 1099-R (income from retirement accounts and pensions). Separate observations are created for primary and secondary taxpayers who were in the sample in 1999. The panel is a stratified random sample, which oversamples high-income returns. Sampling weights allow estimation of population aggregates.

We use information from several tax forms for the analysis. Our measure of gross Social Security benefit comes from Form 1099-SSA, an information return the Social Security Administration produces to report benefits for each recipient. Tax-exempt interest and the amount of Social Security benefits that are included

[10] For more information on the SOI Individual Income Tax Return Panel, see Weber and Bryant (2005). Pierce (2011) documents an extended version of the panel (through 2008).

in AGI come from Form 1040. Our dataset also includes reported self-employment income from Form 1040 Schedule SE.

FIGURE 3. Effect of Introducing a Kink in the Budget Constraint

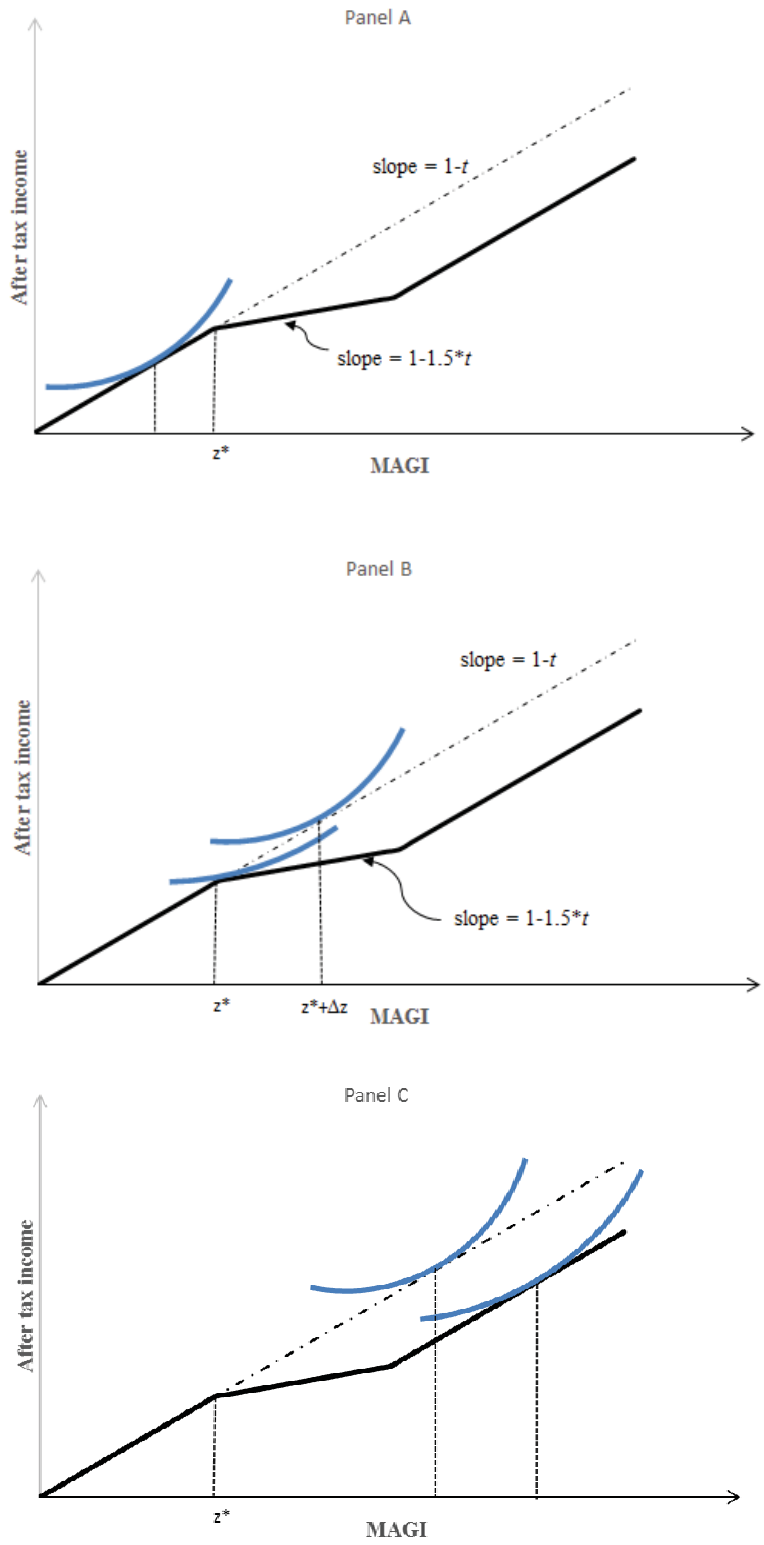
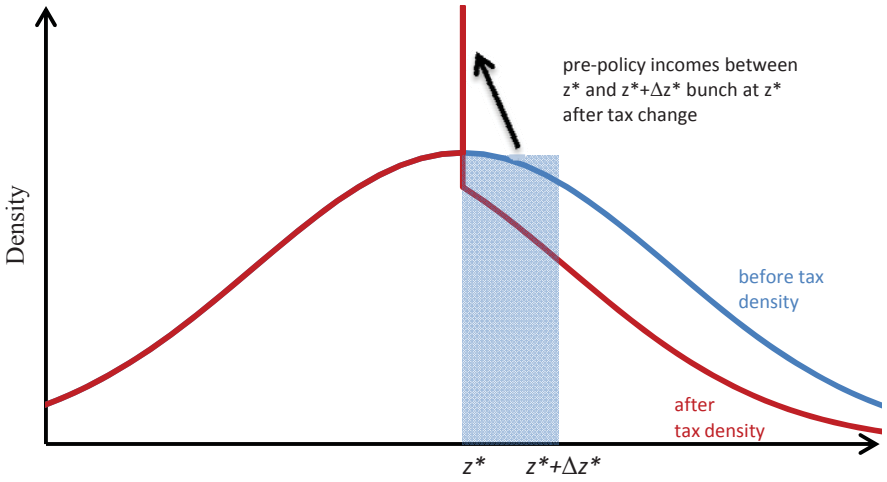


FIGURE 4. Illustration of Bunching at Threshold (z*) in Simple Utility Maximization Framework



Administrative data are not immune from measurement error. For example, self-employed taxpayers often misreport their income to the IRS.[11] The data, however, accurately reflect pre-audit information that taxpayers report to the tax authorities, and the resulting tax liability. Therefore, any behavioral response to the taxation of Social Security benefits should be evident on the tax return.

The sample starts with 112,823 records in 1999, but diminishes to 106,655 by 2008 (see Table 1). We are primarily interested in the subsample of taxpayers age 62 and over, which includes 23,535 individuals in 1999 and 36,530 in 2008. The weighted sample includes 153.6 million individuals in 1999, 28.6 million of whom are age 62 and over. Attrition within the panel is primarily due to death, but taxpayers may also drop out in years in which their income falls below the filing threshold. Because our sample has been supplemented with information returns, particularly earnings from the W-2 and Social Security benefit payments from Form 1099-SSA, we will continue to observe almost all individuals who are not required to file an income tax return. The sample of interest—taxpayers age 62 and over—actually increases over time, a reflection of an aging sample population.

TABLE 1. 1999 SOI Edited Panel Sample Sizes, 1999–2008

Tax Year	Total Sample		Subsample with Primary Taxpayer Age 62 or Over	
	Unweighted	Weighted	Unweighted	Weighted
1999	112,823	153,578,941	23,535	28,574,758
2000	112,804	153,904,818	24,797	29,666,609
2001	112,783	154,187,313	25,990	30,637,473
2002	112,528	154,118,136	27,282	31,720,317
2003	112,058	153,648,715	28,536	32,640,930
2004	111,144	152,282,996	30,269	34,106,473
2005	110,048	150,512,455	31,918	35,426,133
2006	108,946	148,771,365	33,380	36,666,559
2007	107,844	147,034,343	34,740	37,831,748
2008	106,655	145,134,423	36,530	39,309,668

Note: Total sample excludes returns receiving disability payments and those where the primary taxpayer is younger than 23.

[11] Based on audit data, only 43 percent of nonfarm proprietor income (i.e., small business income) was voluntarily reported on tax returns in 2001 (Internal Revenue Service 2006).

All told, the panel includes 755,087 observations for married individuals and 352,546 for singles, representing multiple annual observations for most individuals (see Table 2). Applying sample weights, that represents 906.9 million married filers and 606.3 million single filers. Most of the sample is too young to qualify for Social Security benefits (see Table 1); only 21.4 percent of married individuals and 18.6 percent of singles have Social Security benefits.[12]

TABLE 2. Summary Statistics, Pooled Sample: 1999–2008

	Married		Single	
	Unweighted	Weighted	Unweighted	Weighted
Number of Returns	755,087	906.9M	352,546	606.3M
Self-employed (%)	28	19	9.6	7.3
With SSB income (%)	23.5	21.4	20.1	18.6
Social Security Benefit ($)	24,371	20,489	15,171	13,711
SSB in AGI ($)	14,802	8,605	5,415	3,036
MAGI ($)	3,146,938	110,823	1,050,723	37,518
Wage Earners				
Wage income ($)	508,710	75,881	112,727	27,040
Self-Employed				
Self-employment income ($)	428,950	37,070	341,050	20,111
Wage income ($)	1,085,714	65,144	646,428	17,189

VI. Results

Figure 5 reports the distribution of MAGI relative to the first exempt amount calculated using the IRS Panel. A value of -1,000 on the x-axis means $1,000 below the threshold. Most of the panels are restricted to the sample of taxpayers who have been claiming Social Security benefits for at least 1 year under the logic that it may take time to understand the tax rules. Results are very similar if that restriction is lifted, and also are similar at the second threshold for taxation (see Appendix). Relative MAGI is measured in 2008 dollars. All of the histograms are weighted by population weights; unweighted histograms (not shown) look similar.

To examine bunching evidence statistically, we compare the empirical density, represented by the dots in the scatter plot with smoothed distributions, indicated by the solid line, of MAGI in the vicinity of the threshold in the right panel of each figure. The smoothed distribution is fitted by a quadratic form of MAGI, excluding the observations within $1,000 of the threshold. The grey band indicates the 95 percent confidence interval, reflecting the underlying variability of the data. The simple empirical test of bunching is whether observations near the threshold fall outside the confidence band (reflecting normal sample variability).

Unlike the histograms for the SSET, EITC, or Danish tax system reported in earlier studies, there is no visual evidence of bunching near the MAGI threshold, indicated by the red line, either for all taxpayers or for the self-employed subsample.

It is possible that married and single taxpayers respond differently to the taxation of benefits. Single taxpayers have an easier optimization problem to solve so this is a cleaner test of the bunching hypothesis. Presumably singles have more control of their own MAGI than individual spouses have in managing joint MAGI. Figure 6 shows the MAGI distribution separately for married and single households. Although there is no evidence of bunching for wage earners, there is a hint of bunching to the left of the threshold for single taxpayers with income from self-employment.

[12] Younger adults may qualify for Social Security disability benefits, but those individuals have been excluded from our sample.

All told, the evidence would seem to allay concerns that taxpayers might be over-reacting to the taxation of Social Security benefits. Responses appear to be modest, at most. There is only weak evidence of response for single taxpayers with self-employment income.

FIGURE 5. MAGI Distribution Around the First Threshold

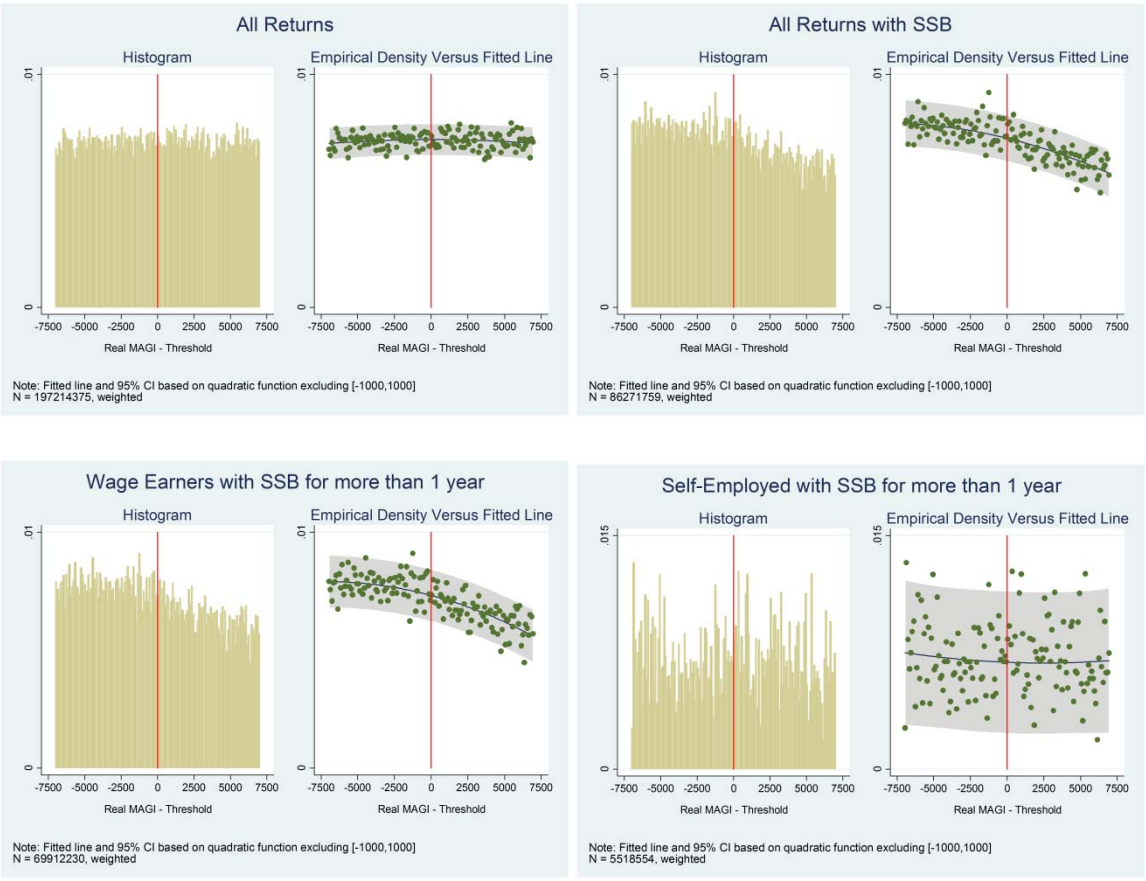
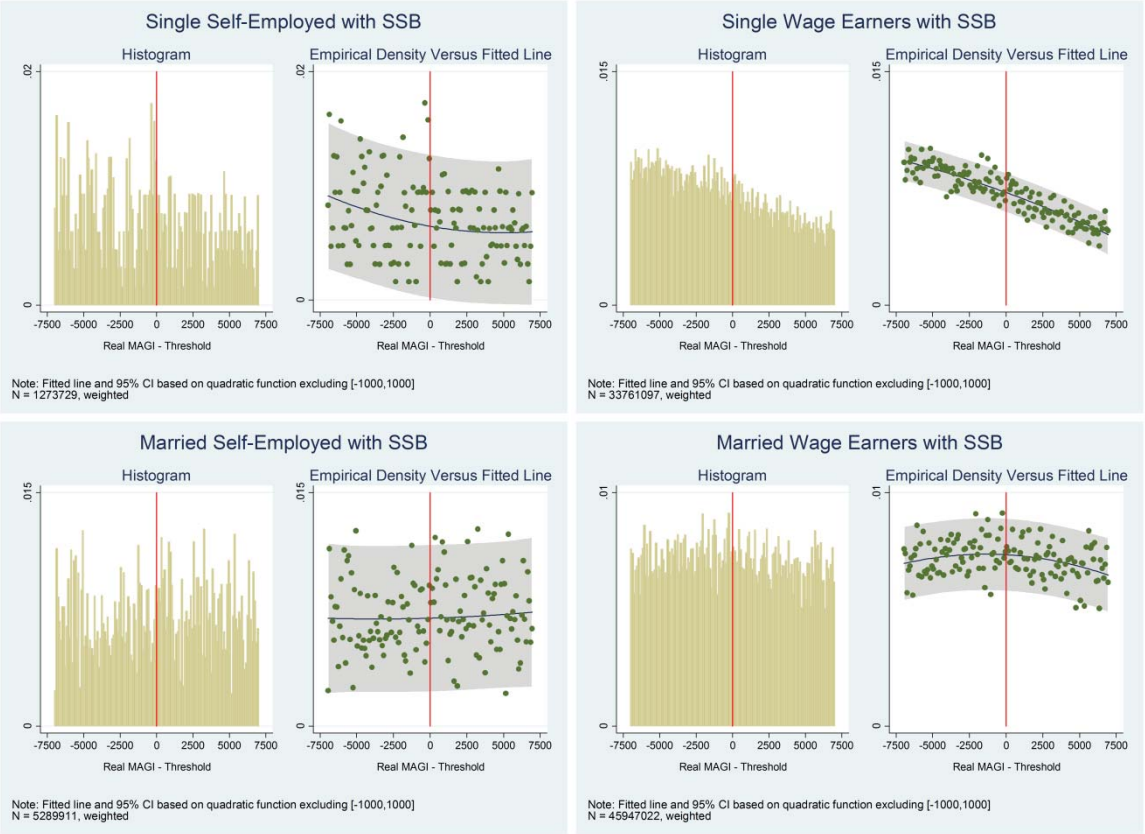


FIGURE 6. MAGI Distribution Around the First Threshold by Marital Status



VII. Conclusion

The taxation of Social Security benefits creates high effective marginal tax rates, which gives older workers an incentive to reduce their labor and nonlabor income below the taxable threshold. However, the tax rules are also quite complex. While in theory taxpayers have an unambiguous incentive to reduce income in the neighborhood of the threshold, the practical effect of these complex incentives is an empirical question. If taxpayers respond to those incentives, there could be significant efficiency costs as well as implications for Social Security's and the nation's finances as older workers would be paying less income and payroll taxes. Moreover, the issue is important as the nation considers tax reform options, which might include changing the way Social Security is taxed.

This study uses administrative data from tax and information returns to examine the distribution of Social Security recipients in the neighborhood of the taxation thresholds. There is little evidence of a response. We examined married and single individuals with and without self-employment income. Only single, self-employed people show any evidence of reducing income to avoid the tax and the response is much smaller and less precisely estimated than the response Saez (2010) found to the kink in the EITC benefit schedule. Overall, the findings suggest that older taxpayers have little understanding of the incentive effects of taxing Social Security.

In future work, we plan to look at how taxation affects labor force participation and the timing of capital gains realizations; capital gains face a much larger proportional rise in tax rates than other income and the timing of capital gains realization is comparatively easy to manipulate. We also plan to look at whether the taxation of benefits affects when individuals first claim Social Security benefits.

APPENDIX
Graphical Examination of Bunching Around the Second Taxation Threshold

This appendix shows graphs of the empirical density of tax returns around the second (higher) threshold at which Social Security benefits are phased into taxable income at an 85 percent rate (increasing marginal effective tax rates by 85 percent), There is no significant evidence of bunching around this threshold for Social Security recipients.

FIGURE A1. MAGI Distribution Around the Second Threshold

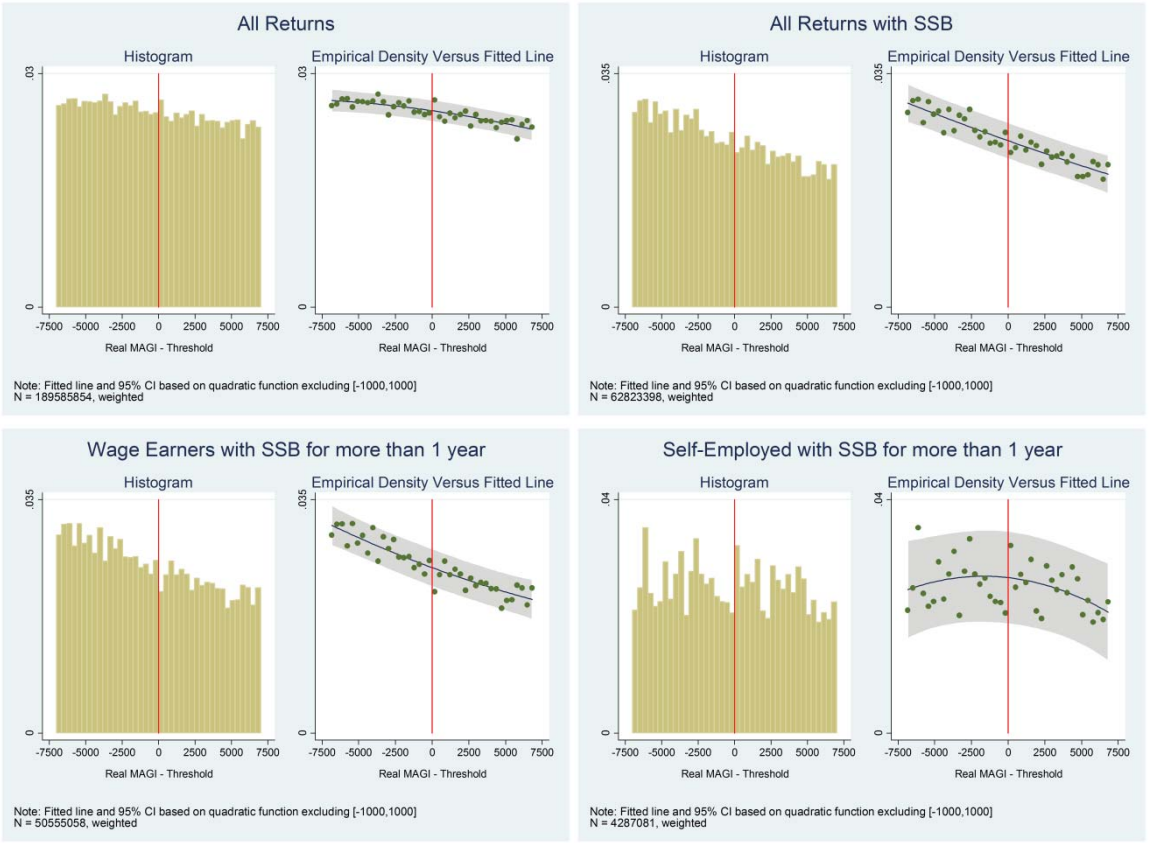
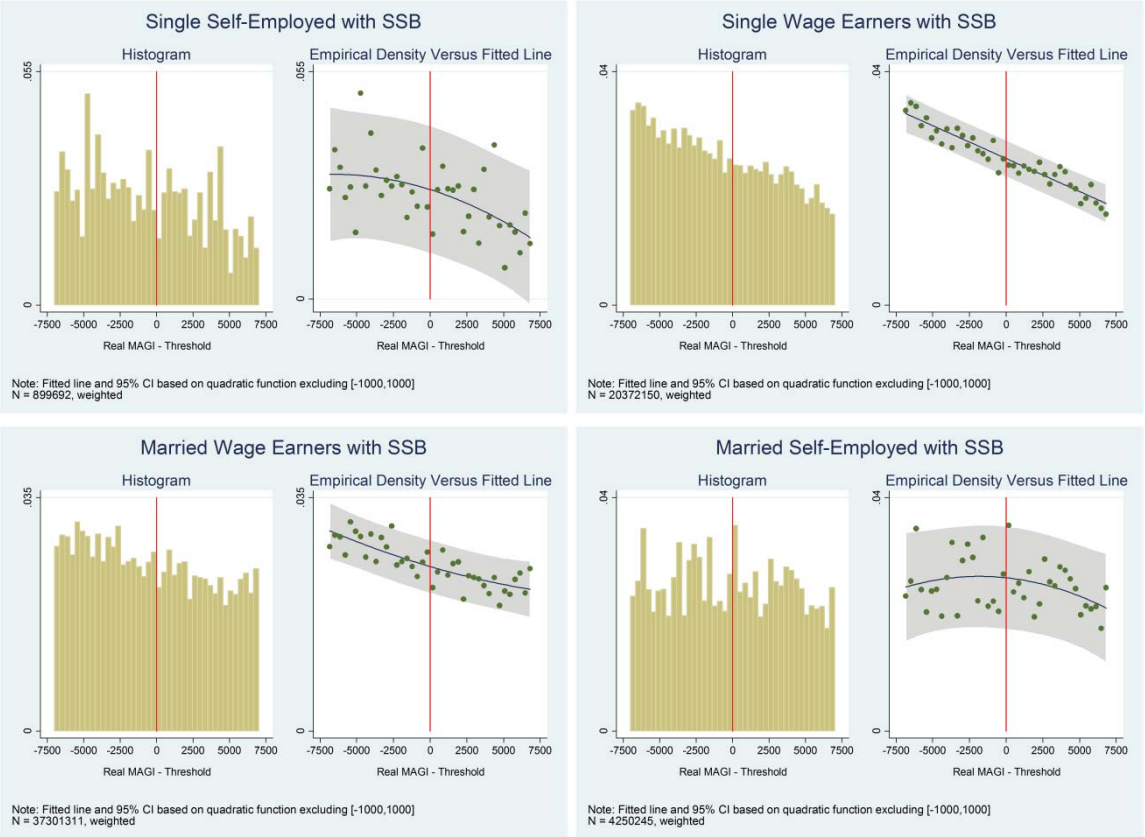


FIGURE A2. MAGI Distribution Around the Second Threshold by Marital Status



References

Benitez-Silva, Hugo, and Frank Heiland. 2007. "The Social Security Earnings Test and Work Incentives." *Journal of Policy Analysis and Management* 26(3): 527–555.

Burtless, Gary, and Robert A. Moffitt. 1985. "The Joint Choice of Retirement Age and Post Retirement Hours of Work," *Journal of Labor Economics.* 3:209–236.

Burman, Leonard E. 1999. *The Labyrinth of Capital Gains Taxation: A Guide for the Perplexed.* Washington, DC: Brookings Institution Press.

Burman, Leonard; Norma Coe; and Liu Tian. 2011. "Do Older Taxpayers Respond to the Taxation of Social Security Benefits?" Manuscript.

Chetty, Raj; John N. Friedman; Tore Olsen; and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics* 126(2): 749–804.

Coile, Courtney; Peter Diamond; Jonathan Gruber; and Alain Jousten. 2002. "Delays in Claiming Social Security Benefits," *Journal of Public Economics,* 84(3): 357–385.

Engelhardt, Gary V., and Anil Kumar. 2009. "The Repeal of the Retirement Earnings Test and the Labor Supply of Older Men." *Journal of Pension Economics and Finance* 8(4): 429–50.

Friedberg, Leora. 2000. "The Labor Supply Effects of the Social Security Earnings Test." *The Review of Economics and Statistics* 82(1): 48–63.

Friedberg, Leora, and Anthony Webb. 2009. "New Evidence on the Labor Supply Effects of the Social Security Earnings Test." *Tax Policy and the Economy* 23: 1–35.

Gruber, Jonathan, and Peter Orszag. 2003. "Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt?" *National Tax Journal* 56(4):755–773.

Gustman, Alan L., and Thomas L. Steinmeier. 1985. "The 1983 Social Security Reforms and Labor Supply Adjustments of Older Individuals in the Long Run." *Journal of Labor Economics* 3(2): 237–253.

Gustman, Alan L., and Thomas L. Steinmeier. 2005. "The Social Security Early Retirement Age in a Structural Model of Retirement and Wealth." *Journal of Public Economics* 89(2-3): 441–463.

Heider, Steven J., and David S. Loughran. 2008. "The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data." *Journal of Human Resources* 43(1): 57–87.

Internal Revenue Service. 2006. "IRS Updates Tax Gap Estimates." Available at http://www.irs.gov/uac/IRS-Updates-Tax-Gap-Estimates.

Liebman, Jeffrey. 1997. "The Impact of the EITC on Incentives and Income Distribution." In *Tax Policy and the Economy*, ed. James Poterba, 83–119. Cambridge: MIT Press.

Liebman, Jeffrey, and Serena Goodman. 2008. "The Taxation of Social Security Benefits as an Approach to Means Testing." Working Paper NB08-02. Cambridge, MA: National Bureau of Economic Research.

Liebman, Jeffrey B.; Erzo F.P. Luttmer; and David G. Seif. 2009. "Labor Supply Responses to Marginal Social Security Benefits: Evidence from Discontinuities." *Journal of Public Economics* 93(11-12): 1208–1223.

Liebman, Jeffrey B., and Erzo F. P. Luttmer. 2012. "The Perception of Social Security Incentives for Labor Supply and Retirement: The Median Voter Knows More Than You'd Think," in Jeffrey Brown, ed., *Tax Policy and the Economy, Volume 26*. Cambridge, MA: National Bureau of Economic Research.

Page, Timothy, and Conway, Karen. 2011. "The Labor Supply Effects of Taxing Social Security Benefits." University of New Hampshire (unpublished manuscript).

Pierce, Kevin. 2011. "General Description Booklet for the SOI Edited Panel Retirement File." Internal Revenue Service (unpublished manuscript).

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (August 2010): 180–212

Song, Jae, and Joyce Manchester. 2007. "New Evidence on Earnings and Benefit Claims Following Changes in the Retirement Earnings Test in 2000." *Journal of Public Economics* 91(3-4): 669–700.

Weber, Michael, and Victoria Bryant. 2005. "The 1999 Individual Income Tax Return Edited Panel," *2005 Proceedings of the American Statistical Association*, Social Statistics Section, Government Statistics Section, American Statistical Association.

Return Preparer Industry Analysis[1]

Patrick Langetieg, Mark Payne, and Melissa Vigil,
Office of Research: Taxpayer Analysis & Modeling, Internal Revenue Service

In 2010, the IRS adopted regulations aimed at establishing standards among tax return preparers. The objective was to improve voluntary compliance by supporting the paid preparer community and providing oversight of the industry with the goal of reducing errors on tax returns. By January 1, 2011, preparers were required to register with the IRS in order to receive a preparer tax identification number (PTIN) that is to be entered on the returns completed by the preparer. Preparers pay an annual fee for the PTIN. Additionally, those without professional credentials were required to fulfill new continuing education requirements and pass a new standardized test designed to ensure that minimum competencies or proficiencies are met.[2] Concurrently, larger preparers were statutorily required to electronically file the tax returns they prepared. IRS implementation of the e-file mandate required preparers in processing year 2011 to electronically submit returns if they expected to prepare and file at least 100 returns. This threshold was lowered to 11 or more returns in processing year 2012.

This report summarizes individual income tax preparer industry trends, trends in individual income tax return accuracy, and preliminary evidence regarding effects the preparer registration initiative may have had on these trends.

Trends in the number of preparers from processing years 2004[3] through 2013 are examined based on information provided on the individual income tax return.[4] The number of preparers in each processing year is determined by the count of unique identification numbers entered in the preparer identification field on the tax return. Throughout this report we refer to preparers, but it is important to note that we are referring to the self-identification of said preparers who have signed an individual income tax return using either a PTIN, social security number (SSN), employer identification number (EIN), or some other number.

Preparer tax identification numbers were first introduced in processing year 1999, but preparers may have entered a PTIN, an SSN, an EIN, or some other number in this field. Prior to the initiative, relatively few preparers obtained a PTIN, because there was no requirement to do so. To the extent that individual preparers may have used a PTIN for some returns and an SSN or EIN on other returns, the number of preparers will be overstated. It is also possible that non-professional individuals, assisting friends and family, sign the return and enter their SSN.

Some preparers sign some returns using an SSN and others using a PTIN. This results in a count of identifiers that exceeds the associated number of actual preparers. We were able to address this overstatement by finding all the identifiers a preparer used when preparing returns—both over time and across returns. We then developed a unique, consistent preparer identifier for tabulation and analysis. As shown in Figure 1, this adjustment reduces the count of preparers by about ninety thousand in processing year 2004. This difference narrows over time, particularly after the registration initiative, demonstrating that preparers are now more

[1] The authors wish to acknowledge helpful comments and assistance from Ed Emblom, John Guyton, Janice Hedemann, Drew Johns, David Ludlum, Dayanand Manoli, Karen Masken, Esmeralda Stuk and the Compliance Data Warehouse group at IRS.

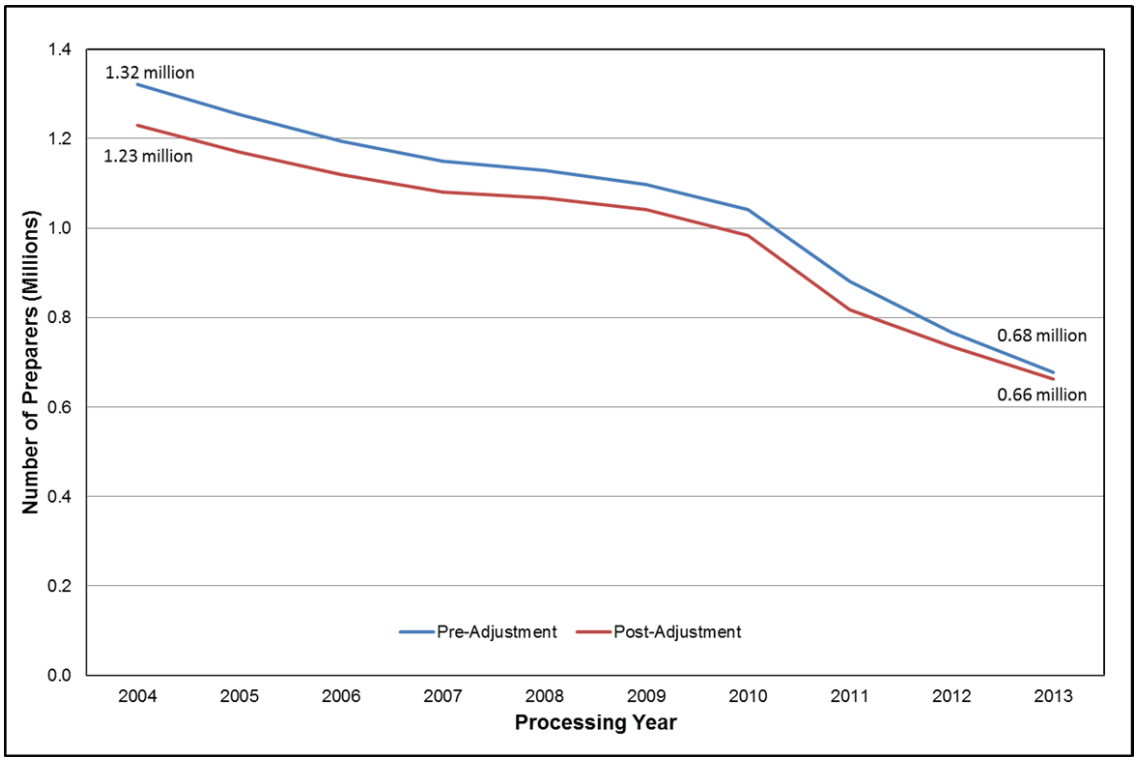
[2] On Friday, Jan. 18, 2013, the United States District Court for the District of Columbia enjoined the Internal Revenue Service from enforcing the regulatory requirements for registered tax return preparers. In accordance with this order, tax return preparers covered by this program are not required to complete competency testing or secure continuing education. The ruling does not affect the regulatory practice requirements for CPAs, attorneys, enrolled agents, enrolled retirement plan agents or enrolled actuaries.

[3] Analysis begins in processing year 2004 due to data processing issues prior to this period. PTINs consist of a leading 'P' followed by a numeric value. Prior to PY2004, PTINs were processed without modification and the leading 'P' resulted in some data loss. After PY2004, the leading character value was stored separately from the numeric portion enabling consistent identification of preparers.

[4] Counts in the Appendix tables and throughout this report were generated using the prep_tin and prep_ein variables entered on the tax return which appear in the irtf_entity table in the Compliance Data Warehouse (CDW). These variables correspond to the PTIN and EIN fields on the tax forms. Included in the counts are the returns for which there was a PTIN present on the form, as well as the cases in which there was no PTIN present on the form, but for which there was an EIN. Most of the data rely on the individual return transaction file data housed in CDW. In addition, the PTIN database, which retains information from registrants, was matched onto the preparers identified in the CDW irtf_entity table.

consistently signing returns using their PTINs. Some redundant identification persists with respect to preparers identified by EIN, as we are unable to unambiguously link an EIN to a single preparer. These EIN-identified preparers may already be represented on other returns by an SSN or PTIN. They may also represent more than one actual preparer.

FIGURE 1. Comparison of the Number of Identifiable Preparers Before and After Adjusting for a Consistent Preparer Identifier, Processing Years 2004–2013



Changes in return accuracy are measured based on errors for which the IRS systematically checks the entire individual taxpayer population. While this excludes many important types of errors, it allows an unbiased, albeit limited, measurement of certain aspects of return accuracy.

The intent of this analysis is to understand how the industry and preparers have changed over time and inform our understanding of how the industry and the taxpayer base continue to respond to changes in the regulation of preparers. It is also important to note that the counts in this report do not necessarily reflect preparers who registered as part of the preparer initiative, except where noted.

Preparer and Prepared Return Trends

The number of preparers who prepared returns from processing years 2004 to 2013 is presented below in Table 1. In the years leading up to the initiative year, the number of preparers declined steadily from 1.23 million preparers in processing year 2004 to 0.98 million preparers in 2010. In 2011, the first year of the initiative, the number of preparers declined by about 17 percent to 0.82 million preparers, and then by an additional 10 percent the second and third years of the initiative, to settle at 0.66 million preparers. As mentioned above, some of this decline can be explained by increased uniformity observed among preparers when signing tax returns, especially after adoption of the initiative. Another explanation may be that per the initiative, preparers who are supervised by a PTIN holder should not be signing tax returns, although they are required to obtain a PTIN. Instead, these preparers' supervisors are required to sign the return.

TABLE 1. Preparers and Prepared Returns, Processing Years 2004–2013*

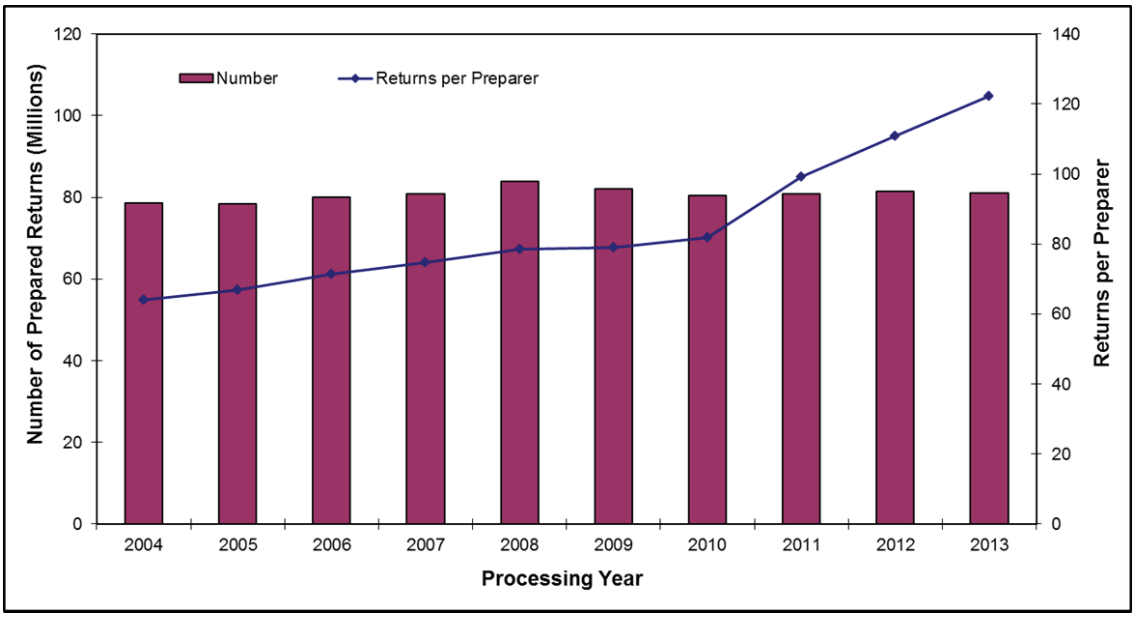
Processing Year	All Individual Returns	Growth	Number of Preparers (millions)	Growth	Prepared Returns (millions)	Growth	Average Returns Prepared per Preparer	Growth
2004	131.3		1.23		78.62		64	
2005	131.2	0.0%	1.17	-4.8%	78.35	-0.4%	67	4.6%
2006	132.4	0.9%	1.12	-4.3%	80.00	2.1%	71	6.7%
2007	133.8	1.1%	1.08	-3.5%	80.84	1.0%	75	4.7%
2008	140.5	5.0%	1.07	-1.2%	83.90	3.8%	79	5.1%
2009	139.5	-0.7%	1.04	-2.5%	82.16	-2.1%	79	0.4%
2010	137.6	-1.4%	0.98	-5.5%	80.51	-2.0%	82	3.7%
2011	140.0	1.8%	0.82	-16.9%	80.94	0.5%	99	21.0%
2012	142.4	1.7%	0.73	-10.1%	81.42	0.6%	111	11.9%
2013	141.7	-0.5%	0.66	-9.7%	80.99	-0.5%	122	10.2%

* Source: RAS:R:TAM tabulations using the IRTF table from December 2013 CDW
Excludes preparers that solely prepared SS/PR/NR/NR-EZ, stimulus, or TETR returns and no others.
Data also exclude all volunteer preparers (e.g., VITA, TCE).

The decrease in the number of preparers does not appear, however, to have noticeably reduced the overall number of prepared returns. Ignoring processing year 2008, which had an unusually high number of prepared returns due to the large volume of stimulus claimant returns processed that year, the total number of prepared returns remained relatively constant between processing years 2007 and 2013. While the absolute number of prepared returns has remained constant, it has declined as a percent of overall returns.

A necessary correlate of these two trends is an increasing number of returns prepared per preparer. With the exception of processing year 2009, the average number of returns prepared per preparer consistently increased, as illustrated in Figure 2. In processing year 2011 the number increased to an average of 99 returns prepared per preparer compared to an average of 82 returns the previous year. In processing years 2012 and 2013, the average number increased further to an average of 111 and 122 returns per preparer, respectively.

FIGURE 2. Number of Prepared Returns and Number of Returns per Preparer, Processing Years 2004–2013

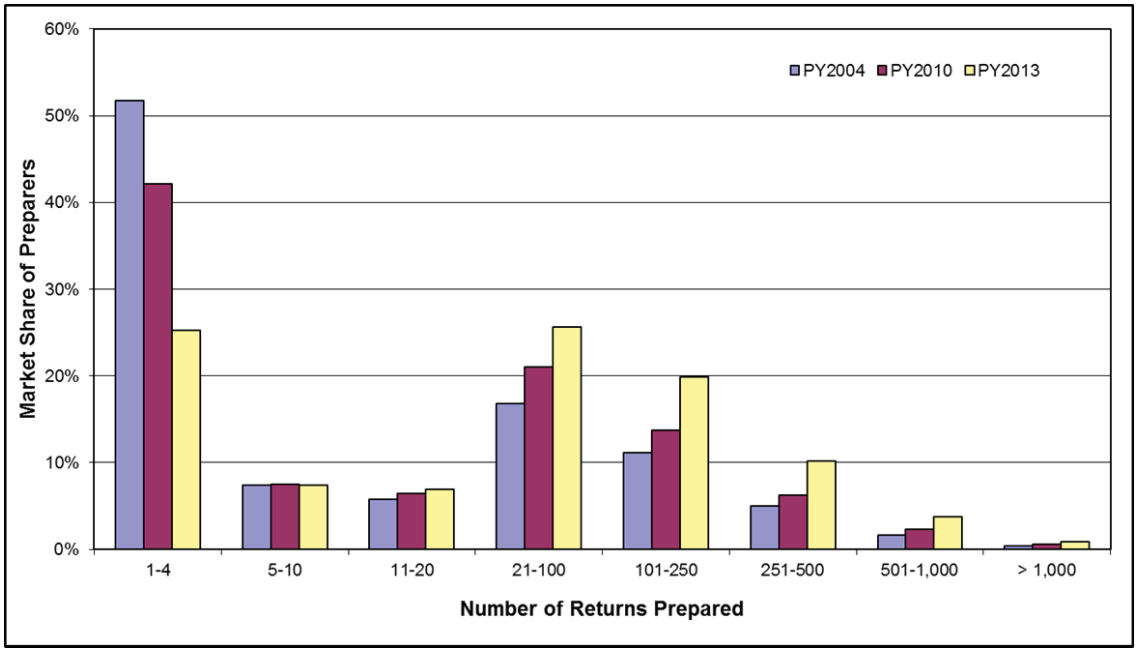


Preparer Industry Trends

The data generally point to increased consolidation in the return preparer market. Figure 3 presents market shares by segmented volumes prepared for select processing years.

Consolidation is evident in nearly all of the segments. The percentage of preparers preparing fewer than five returns per year declined from 52 percent in processing year 2004, to only 25 percent by processing year 2013.[5] In contrast, the percentage of preparers in the 21 to 100 return segment increased by 53 percent, the 101 to 250 return segment increased by 82 percent, and the market shares of all return segments above 250 returns doubled.

FIGURE 3. Market Share of Preparers by Volume Segment, Selected Processing Years



As shown in Figure 4 and Table A2 in the Appendix, the number of preparers was declining prior to the initiative. From processing years 2004 to 2010, the overall number of preparers declined by 20 percent, and all preparer segments preparing fewer than 500 returns either experienced a decline or remained constant. After the initiative, from processing years 2010 to 2013, the number of preparers dropped 33 percent, with the decline occurring primarily among those preparing fewer than 250 returns. In contrast, preparers preparing the largest number of returns increased substantially, by 20 percent prior to the initiative, and then stabilized at about 1.0 percent after the initiative.

At the same time, larger volume preparers are preparing a larger share of returns as shown in Figure 5. In processing year 2004, 83 percent of all prepared returns were prepared by those who prepared more than 100 returns, while the corresponding segments prepared about 87 percent of all prepared returns in processing year 2013.

Throughout the study period, the share of returns dropped for the smallest volume segments and shifted to segments preparing 251 to 1,000 returns.

[5] More than 60 percent of the smallest preparers, those preparing fewer than five returns, prepared only one return throughout the study period.

FIGURE 4. Percentage Change in the Number of Preparers by Volume Segment, Selected Ranges of Processing Years

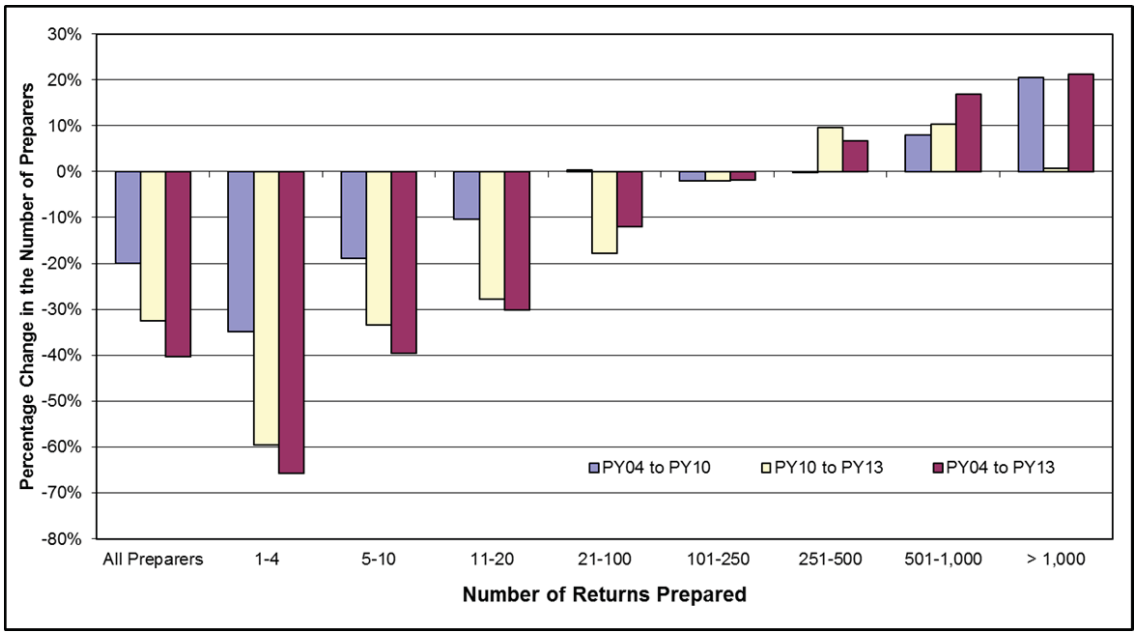


FIGURE 5. Share of Returns Prepared by Preparer Segment, Selected Processing Years

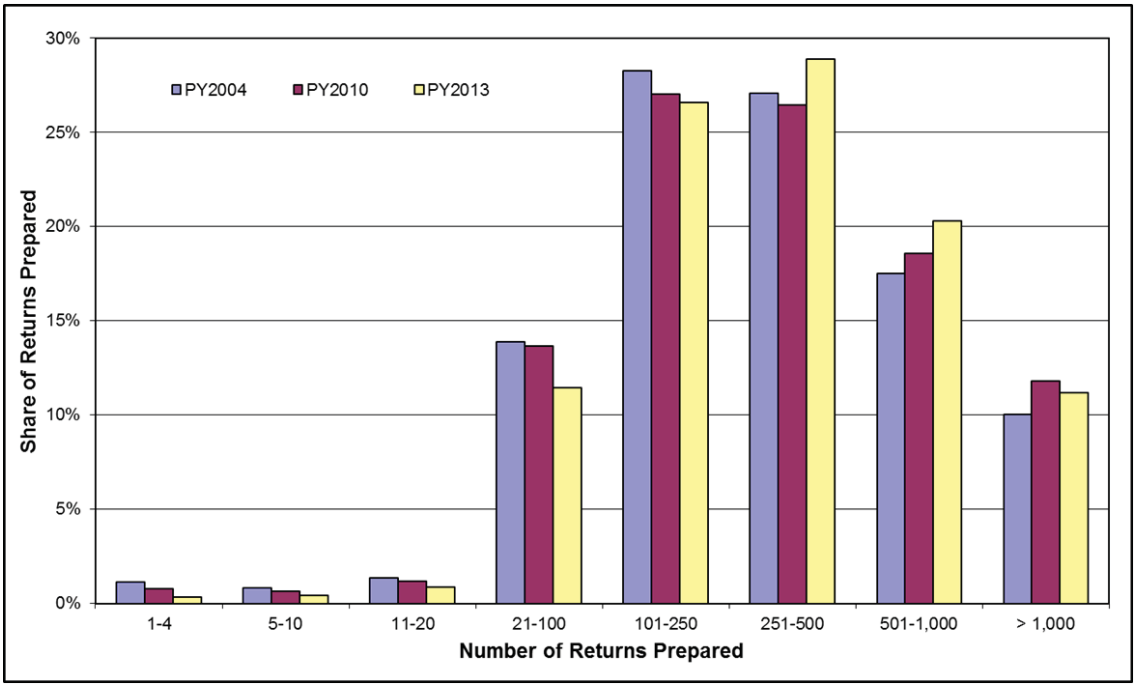
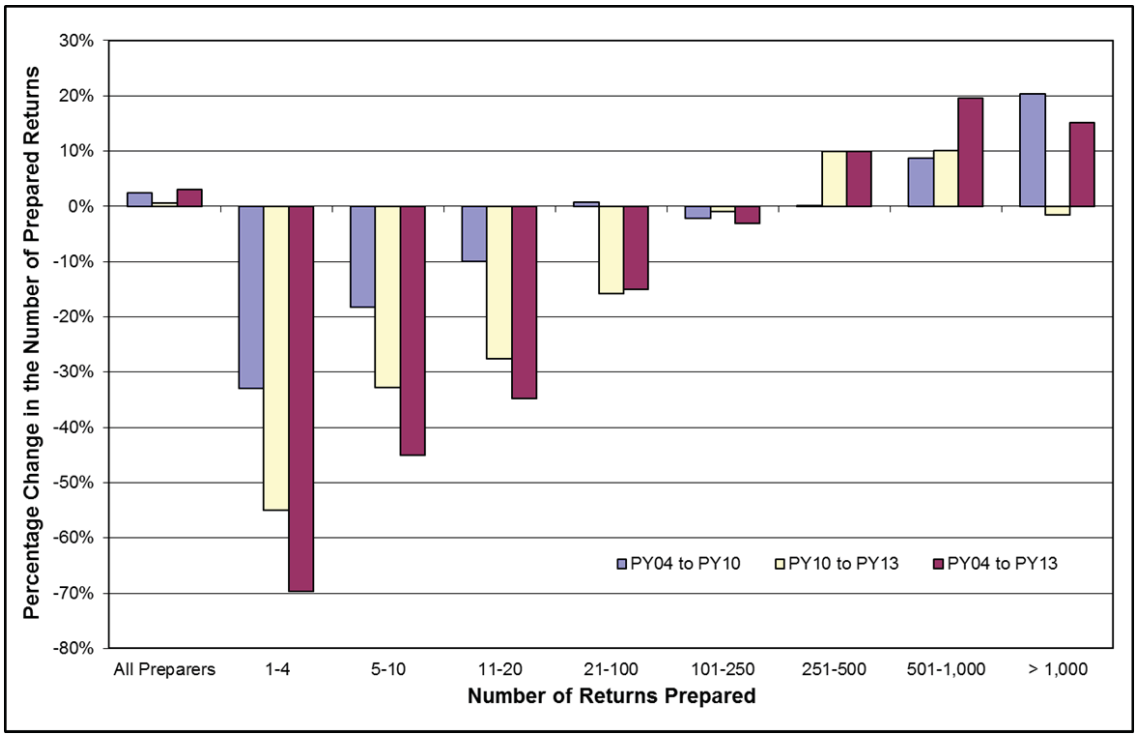


Figure 6 and Appendix Table A2 present growth in the number of prepared returns by volume segment. The overall number of prepared returns increased modestly, by 2 percent between processing years 2004 and 2010, and increased by only 1 percent in the years following the initiative. Similar to the results presented in Figure 4, the number of prepared returns by the smaller segments declined substantially since processing year 2004 and growth was observed only among the largest segments. The exception is among the largest prepared volume segments, which contracted by 2 percent following the initiative.

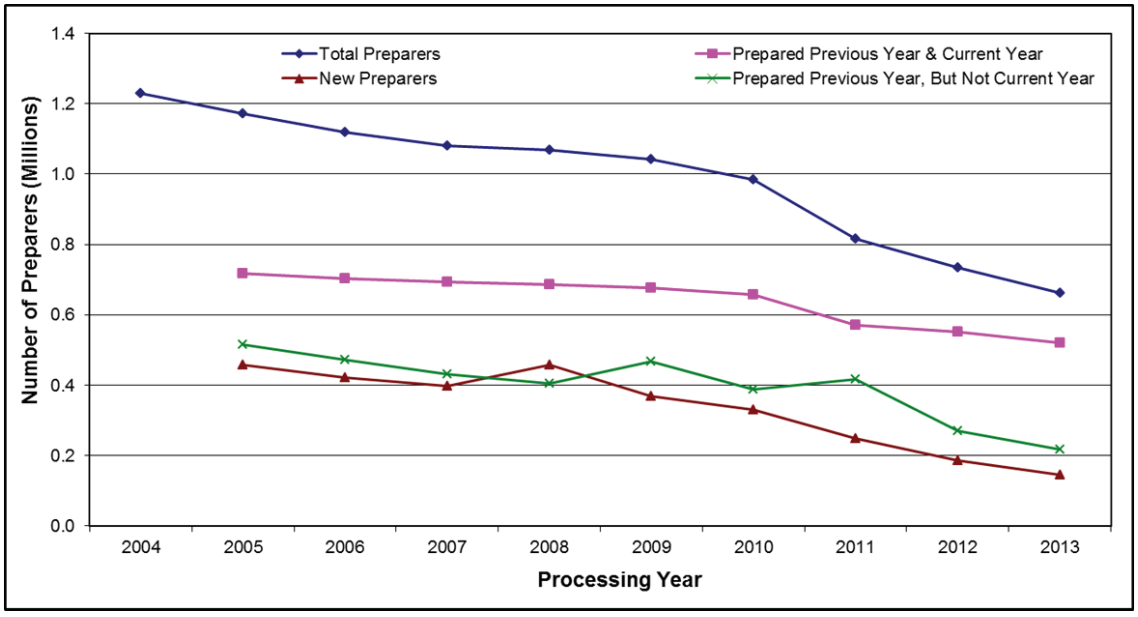
FIGURE 6. Percentage Change in the Number of Prepared Returns by Volume Segment, Selected Ranges of Processing Years



Preparer Industry Dynamics

Figure 7 presents data on the dynamics of the return preparer industry. The data compare pairs of adjacent years. Included are preparers who did not prepare returns in the previous year, but did prepare in the current year (new preparers), preparers who prepared returns in both the previous and current years, and those who prepared the previous year but not the current year (prepared previous year, but not current year).

FIGURE 7. Preparer Status, Processing Years 2004–2013



A larger than normal attrition rate in processing year 2011, combined with fewer preparers entering the market, and a substantial decrease in the preparers who prepared both years, result in a smaller preparer market. The number of new preparers entering the market continued to decline in processing years 2012 and 2013. Furthermore, preparers who prepared both years declined by an average of 7 percent after processing year 2010, compared with an average of 2 percent before the initiative. However, the attrition rate among all preparers in processing years 2012 and 2013 was lower than all previous years. See Appendix Table A3 for additional data.

New preparers typically are not fully replacing those who stop preparing. This phenomenon was particularly strong in processing year 2011, nearly tripling from the previous year, but moderating the following processing year. As Appendix Table A1 shows, most of this volatility is driven by preparers who prepare fewer than five returns.

Characteristics of PTIN Holders

The data up to this point included all preparers, regardless of whether they held a PTIN or not. In this next section we look only at preparers who hold a PTIN and use it to sign returns they prepare. Table 2 and Table 3 present data on the professional credentials of PTIN holders by volume segment for processing years 2011 through 2013.

About 546,000 preparers who held a PTIN in processing year 2011 prepared individual income tax returns. This represents 67 percent of all preparers who prepared returns that year. Of those, about 238,000 self-reported credentials as CPAs, enrolled agents, or attorneys.[6] This number represents 29 percent of all preparers who prepared returns in processing year 2011. Among credentialed PTIN holders, 31 percent self-reported they were CPAs, 7 percent enrolled agents, 2 percent attorneys and 3 percent other credentials. A relatively large share, 48 percent of PTIN holders who prepared fewer than five returns, self-reported as credentialed. Most of these preparers are CPAs and attorneys. In contrast, a relatively larger share of enrolled agents prepared larger volumes of returns.

The number of preparers with a PTIN in processing year 2012 increased to about 556,000, or about 76 percent of all preparers who prepared returns in processing year 2012. Forty-five percent of these preparers self-reported credentials, with similar composition among the types of credentials as in processing year 2011. Aggregate data for processing year 2013 remain relatively stable with continued increases in the percentage of preparers holding a PTIN and credentials, and similar composition among the types of credentials held.

Figure 8 compares the percent of preparers who held a PTIN by volume segment in processing years 2011 to 2013. Over 90 percent of those preparing more than 10 returns had registered by processing year 2013.

Trends in Return Accuracy

The IRS regulates tax return preparers with the goal of supporting more accurate return preparation. Developing a strong baseline understanding of relationships between return accuracy and how returns are prepared can inform regulatory efforts. The focus in this section is on errors for which the IRS systematically checks the entire individual taxpayer population. While this excludes many important types of errors, it allows a direct measurement of certain aspects of return accuracy. This section analyzes return accuracy by focusing on math errors[7] and potential AUR mismatches[8] for individual income tax returns by type of preparer, preparation method, and submission method.

---

[6] Many preparers hold more than one credential. For example, a preparer can both be a CPA and an attorney, CPA and enrolled agent, etc. These preparers were ranked to eliminate any overlap. They are ranked in the following order: CPA, enrolled agent, attorney, other credential.

[7] Math errors refer to all types of errors that fall under the math error authority of Title 26 of the United States Code, as described in Section 6213(b). They include a variety of conditions such as computational errors, incorrectly transcribed values, and omitted entries identified during the processing of tax returns.

[8] AUR is an abbreviation for the IRS Automated Underreporter program, the automated analysis and processing of potential underreported or over-deducted issues identified by matching tax returns against information returns provided by third parties. In the remainder of this report, the term "potential AUR mismatch" will be replaced by "AUR mismatch." What is being referred to here is the finding in the AUR computer matching program of an inconsistency between a line on the return and the information documents reported to the IRS for that taxpayer. Among the potential mismatches that result from this process, a significant number are false positives. Only about one quarter to one third of these potential mismatches are selected for review by the IRS and an even smaller proportion is sent a notice. Based on the IRS review, several hundred thousand are removed from the caseload ("screened out") and some of those taxpayers receiving a notice adequately explain the inconsistency.

TABLE 2. Professional Credentials of PTIN Holders Who Prepared Returns in 2011–2013, by Volume Segment: Levels*

Volume Segment (Number of Returns Prepared)	Number of Preparers	PTIN Holders	Type of Credential				
			Credentialed Preparers	CPA	Enrolled Agent**	Attorney	Other Credential***
2011							
All Preparers	817,004	546,272	238,252	169,340	38,166	12,906	17,840
1–4	274,705	63,141	30,242	22,185	2,358	4,177	1,522
5–10	65,198	41,846	20,028	14,897	1,820	2,053	1,258
11–20	56,655	44,754	19,945	14,804	2,139	1,621	1,381
21–100	190,919	174,839	68,117	50,129	9,266	3,209	5,513
101–250	135,779	130,888	55,798	39,760	10,530	1,280	4,228
251–500	64,566	62,644	30,937	20,292	7,952	415	2,278
501–1,000	23,251	22,504	10,937	6,293	3,337	132	1,175
> 1,000	5,931	5,656	2,248	980	764	19	485
2012							
All Preparers	734,386	556,202	248,976	174,762	41,276	13,639	19,299
1–4	217,297	70,686	32,825	23,579	2,861	4,496	1,889
5–10	55,058	42,732	21,050	15,487	2,027	2,193	1,343
11–20	49,602	43,659	19,923	14,710	2,209	1,617	1,387
21–100	181,253	172,389	70,263	51,323	9,856	3,413	5,671
101–250	134,917	132,046	57,919	40,698	11,273	1,341	4,607
251–500	66,060	65,015	32,603	21,086	8,540	421	2,556
501–1,000	24,148	23,751	11,888	6,806	3,640	129	1,313
> 1,000	6,051	5,924	2,505	1,073	870	29	533
2013							
All Preparers	663,054	541,509	249,960	175,035	42,472	13,457	18,996
1–4	167,582	71,746	32,867	23,465	3,018	4,497	1,887
5–10	49,176	39,819	20,684	15,336	2,111	2,096	1,141
11–20	45,945	41,229	19,897	14,766	2,237	1,592	1,302
21–100	170,175	162,261	69,710	50,919	10,061	3,362	5,368
101–250	131,886	129,380	58,229	40,834	11,424	1,335	4,636
251–500	67,654	66,818	33,678	21,500	8,978	413	2,787
501–1,000	24,585	24,299	12,325	7,092	3,746	137	1,350
> 1,000	6,051	5,957	2,570	1,123	897	25	525

* Source: RAS:R:TAM tabulations using IRTF and PTIN tables from December 2013 CDW

** Includes enrolled actuaries.

*** Includes enrolled retirement plan agents, certified acceptance agents, and state regulated tax return preparers.

Excludes preparers that solely prepared SS/PR/NR/NR-EZ, stimulus, or TETR returns and no others.

Data also exclude all volunteer preparers (e.g., VITA, TCE).

TABLE 3. Professional Credentials of PTIN Holders Who Prepared Returns in 2011–2013 by Volume Segment: Shares*

Volume Segment (Number of Returns Prepared)	Percent of All Preparers Who Held a PTIN	All Credentialed PTIN Holders	Percent of PTIN Holders Who are Credentialed				
			Credentialed Preparers	CPA	Enrolled Agent**	Attorney	Other Credential***
2011							
All Preparers	67%	44%	31%	7%	2%	3%	17,840
1–4	23%	48%	35%	4%	7%	2%	1,522
5–10	64%	48%	36%	4%	5%	3%	1,258
11–20	79%	45%	33%	5%	4%	3%	1,381
21–100	92%	39%	29%	5%	2%	3%	5,513
101–250	96%	43%	30%	8%	1%	3%	4,228
251–500	97%	49%	32%	13%	1%	4%	2,278
501–1,000	97%	49%	28%	15%	1%	5%	1,175
> 1,000	95%	40%	17%	14%	0%	9%	485
2012							
All Preparers	76%	45%	31%	7%	2%	3%	19,299
1–4	33%	46%	33%	4%	6%	3%	1,889
5–10	78%	49%	36%	5%	5%	3%	1,343
11–20	88%	46%	34%	5%	4%	3%	1,387
21–100	95%	41%	30%	6%	2%	3%	5,671
101–250	98%	44%	31%	9%	1%	3%	4,607
251–500	98%	50%	32%	13%	1%	4%	2,556
501–1,000	98%	50%	29%	15%	1%	6%	1,313
> 1,000	98%	42%	18%	15%	0%	9%	533
2013							
All Preparers	82%	46%	32%	8%	2%	4%	18,996
1–4	43%	46%	33%	4%	6%	3%	1,887
5–10	81%	52%	39%	5%	5%	3%	1,141
11–20	90%	48%	36%	5%	4%	3%	1,302
21–100	95%	43%	31%	6%	2%	3%	5,368
101–250	98%	45%	32%	9%	1%	4%	4,636
251–500	99%	50%	32%	13%	1%	4%	2,787
501–1,000	99%	51%	29%	15%	1%	6%	1,350
> 1,000	98%	43%	19%	15%	0%	9%	525

* Source: RAS:R:TAM tabulations using IRTF and PTIN tables from December 2013 CDW

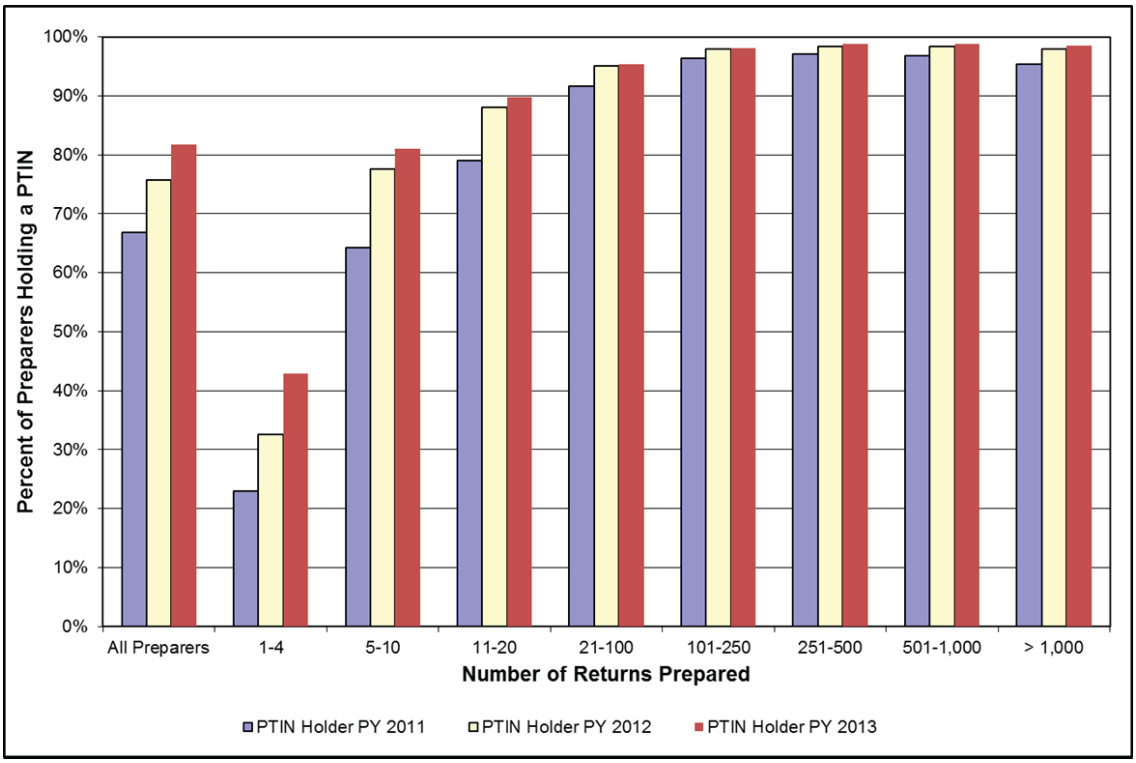
** Includes enrolled actuaries.

*** Includes enrolled retirement plan agents, certified acceptance agents, and state regulated tax return preparers.

Excludes preparers that solely prepared SS/PR/NR/NR-EZ, stimulus, or TETR returns and no others.

Data also exclude all volunteer preparers (e.g., VITA, TCE).

FIGURE 8. Percent of Preparers, by Volume Segment, and by Who Held a PTIN for Processing Years 2011 to 2013



The percentage of returns with a math error is substantially lower for returns prepared with tax preparation software and electronically filed. After controlling for use of software and e-filing, the math error rate for returns prepared by PTIN holders is lower than the rate associated with returns prepared by preparers who do not hold a PTIN. Among e-filed returns, preparers who self-report credentials have similar math error rates to those who do not; however, the math error rate for paper returns prepared by self-reported credentialed preparers is lower than the rate associated with non-PTIN holders and PTIN holders without credentials.

The percentage of returns that have at least one AUR mismatch is influenced more significantly by return complexity[9] and the number of income and deduction items that can be matched to information returns than by preparation and submission method. But, the type of preparer appears to be a factor in AUR mismatch rates, which are lower for returns prepared by PTIN holders than for returns prepared by non-PTIN holders. AUR mismatches are also less common for preparers who prepare a relatively large number of returns.

Math Errors

The analysis indicates that math errors are much more likely to occur on self-prepared paper returns than on self-prepared software returns or paid-prepared returns. The error rate for self-prepared paper returns is 28 times greater than for paid-prepared returns and 21 times greater than for returns prepared by taxpayers using software. The self-prepared paper returns' math error rate is 37.0 percent while for paid-prepared returns it is 1.3 percent and for software-prepared returns it is 1.8 percent.

[9] In the more detailed tables provided on AUR mismatches in the Appendix, returns are classified into three complexity categories—simple, intermediate, and complex—according to the definition in the Electronic Tax Administration IMF database. Simple returns are those without any schedules attached. Intermediate returns are Form 1040A returns with schedules or Form 1040 returns with Schedules A, B, D, Additional Child Tax Credit, Educational Credits, Child Care Credit, Credit for the Elderly, or Earned Income Tax Credit. Complex returns are Form 1040 returns with schedules C, E or F or other schedules.

As shown in Table 4, the comparatively low paid-prepared return math error rate appears to come in large part from the greater likelihood that such returns have been prepared with the use of software and are submitted electronically. For example, returns prepared by hand by paid preparers have a math error rate of 18.0 percent. This rate drops to 5.5 percent when the paid preparer uses software to prepare the return, but files the return by mail. It drops to 0.6 percent when the paid preparer uses software and submits the return electronically. Software helps taxpayers avoid math errors by ensuring that all of the computations are done correctly, line amounts from specific schedules are accurately transferred to the 1040 form, and by accurately determining eligibility for certain credits and deductions, given taxpayer inputs for income, filing status, family structure, etc. Electronic filing provides additional filters before returns are accepted, including ensuring that social security numbers and names are valid and entered accurately, that claimed dependents have not been previously claimed on another return, that all necessary schedules and forms are included with the return, and that the return is signed.[10]

TABLE 4. Percentage of Returns with a Math Error by Preparation and Submission Method, Tax Year 2010

Submission Method	Preparation Method		
	Paid Preparer	Self-Prepared	Total
E-file	0.6	1.1	0.8
Software-Prepared Paper Return	5.5	5.3	5.4
Hand-Prepared Paper Return	18.0	37.0	35.5
Total	1.3	6.8	3.6

Source: RAS:R:TAM. Analysis of data from CDW: IRTF through 3/2014.
NOTE: 1040PR, 1040NR, 1040NR-EZ and 1040SS and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.

Similarly, when self-preparing taxpayers use software but do not e-file, their error rate is 5.3 percent (compared to the 37.0 percent rate for paper returns self-prepared by hand) and falls to 1.1 percent when these returns are e-filed.

Math error rates also differ across types of preparers (Table 5). Returns prepared by preparers who hold a PTIN have fewer math errors than returns prepared by other preparers, even after considering their different rates of using software and e-filing their returns. While e-filed, software-prepared paper returns, and hand-prepared returns of PTIN holders have math error rates of 0.6 percent, 5.2 percent, and 15.1 percent, respectively, the error rates for other paid returns are 0.9 percent, 7.2 percent, and 26.1 percent.

TABLE 5. Percentage of Returns with a Math Error by Type of Preparer and Submission Method, Tax Year 2010

Submission Method	Preparer Type	
	PTIN Holders	Non-PTIN Holders
E-file	0.6	0.9
Software-Prepared Paper Return	5.2	7.2
Hand-Prepared Paper Return	15.1	26.1
Total	1.2	4.1

Source: RAS:R:TAM. Analysis of data from CDW: IRTF, IMF and Return Preparer Registration Database through 3/2014.
NOTE: 1040PR, 1040NR, 1040NR-EZ and 1040SS and returns prepared at VITA sites are excluded. Limited to returns filed in 2011. Preparers are considered as having a PTIN if it was obtained by the end of 2011.

[10] To the extent that some returns rejected by the e-file filters may transfer to paper submission this would tend to lower the math error rate for e-filing and raise it for paper returns. For instance, in the case of divorced parents claiming the same dependents on separate returns, the first return that is e-filed would be accepted while the second return would be rejected. If the second filer then files on paper, this return would be subject to a math error even if this taxpayer is the primary provider for the children. A further risk with rejected returns is that the taxpayer may not file the return at all.

Software-prepared paper returns and hand-prepared returns prepared by self-reported, credentialed, PTIN holders have lower error rates than returns prepared by PTIN holders who do not self-report such credentials and by non-PTIN holders. While 4.4 percent of the software-prepared paper returns by credentialed preparers have math errors and 12.7 percent of their hand-prepared returns have math errors, these numbers are 6.1 percent and 19.4 percent respectively for non-credentialed preparers. On the other hand, e-filed returns prepared by such credentialed preparers have a slightly higher math error rate (0.7 percent) than the rest of the e-filed paid-prepared returns (0.6 percent) (see Appendix: Table A7).

Math error rates also differ across preparers according to the number of returns they prepare. While preparers signing 10 or fewer returns have math errors on 5.8 percent of their returns, those who prepare more than 100 returns have math errors on only about 1.1 percent of their returns. An important reason for this difference is the fact that preparers who prepare a larger number of returns are significantly more likely to e-file their returns and more likely to use software. But, even after controlling for the different rates of electronic submission and software use across the market segments, those preparing 20 or fewer returns have a larger math error rate than those preparing more than 20 returns (Table 6 and Table 7). This difference can be partly accounted for by the fact that those preparing 20 or fewer returns are less likely to be PTIN holders and, as we saw above, non-PTIN holders have a higher error rate, even after controlling for different rates of using software and e-filing. While about 82 percent of returns were signed by PTIN holders in the 20 or fewer return segments, this figure was 97 percent for the more than 20 return segments.

TABLE 6. Percentage of Returns with a Math Error by Number of Returns Prepared and Submission Method, Tax Year 2010

Number of Returns	Submission Method	Error Rate	Share of Returns
1–4	Paid with Software E-File	2.3%	31.3%
	Paid with Software on Paper	6.5%	56.3%
	Paid by Hand	27.9%	12.4%
5–10	Paid with Software E-File	1.0%	47.2%
	Paid with Software on Paper	5.2%	46.1%
	Paid by Hand	22.5%	6.7%
11–20	Paid with Software E-File	0.8%	58.9%
	Paid with Software on Paper	4.9%	36.9%
	Paid by Hand	20.8%	4.2%
21–100	Paid with Software E-File	0.6%	79.5%
	Paid with Software on Paper	5.1%	18.9%
	Paid by Hand	17.7%	1.6%
101–250	Paid with Software E-File	0.6%	90.0%
	Paid with Software on Paper	5.2%	9.4%
	Paid by Hand	15.4%	0.6%
251–500	Paid with Software E-File	0.6%	91.1%
	Paid with Software on Paper	5.2%	8.5%
	Paid by Hand	13.4%	0.4%
501–1,000	Paid with Software E-File	0.7%	91.0%
	Paid with Software on Paper	5.8%	8.7%
	Paid by Hand	13.9%	0.3%
>1,000	Paid with Software E-File	0.7%	88.7%
	Paid with Software on Paper	6.2%	11.0%
	Paid by Hand	13.2%	0.2%
Preparers with ID number	Paid with Software E-File	0.6%	88.3%
	Paid with Software on Paper	5.4%	11.0%
	Paid by Hand	17.0%	0.7%
Preparers without ID number	Paid with Software E-File	1.0%	34.9%
	Paid with Software on Paper	7.2%	56.4%
	Paid by Hand	26.2%	8.7%
All Preparers	Paid with Software E-File	0.6%	87.8%
	Paid with Software on Paper	5.5%	11.4%
	Paid by Hand	18.0%	0.7%
	All Methods	1.3%	100.0%

Source: RAS:R:TAM. Analysis of data from CDW: IRTF and IMF through 3/2014.
NOTE: 1040PR, 1040NR, 1040NR-EZ and 1040SS and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.
The line "no preparer id number" shows the error rate for returns that are signed by a third party preparer but no identifying number for the individual preparer was entered on the return.

TABLE 7. Percentage of Returns with a Math Error by Number of Returns Prepared and Preparer Type, Tax Year 2010

Number of Returns	Preparer Type	Error Rate	Share of Returns
1–4	PTIN Holder	4.8%	50.2%
	No PTIN	10.9%	49.8%
5–10	PTIN Holder	3.6%	85.1%
	No PTIN	8.8%	14.9%
11–20	PTIN Holder	2.7%	90.3%
	No PTIN	7.2%	9.7%
21-100	PTIN Holder	1.6%	95.4%
	No PTIN	4.5%	4.6%
101–250	PTIN Holder	1.1%	97.3%
	No PTIN	2.8%	2.7%
251–500	PTIN Holder	1.0%	97.5%
	No PTIN	2.2%	2.5%
501–1,000	PTIN Holder	1.1%	97.2%
	No PTIN	2.1%	2.8%
>1,000	PTIN Holder	1.4%	96.2%
	No PTIN	1.8%	3.8%
Total	PTIN Holder	1.2%	96.7%
	No PTIN	3.4%	3.3%
	No preparer ID number	6.7%	0.0%
	All Preparers	1.3%	0.0%

Source: RAS:R:TAM. Analysis of data from CDW: IRTF and IMF through 3/2014.
NOTE: 1040PR, 1040NR, 1040NR-EZ and 1040SS and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.
Preparers are considered as having a PTIN if it was obtained by the end of 2011. The line "no preparer id number" shows the error rate for returns that are signed by a third party preparer but no identifying number for the individual preparer was entered on the return.

The long-term trends in tax return preparation and submission methods, including the increased use of software, increases in e-filing rates, and decreases in the number of returns prepared by preparers who prepare relatively few returns, have all contributed to reducing the rate of math errors (see Figure 9). In the period prior to tax year 2010, the main forces driving these preparation and submission trends were technological and financial in nature, including better and less expensive software, increased computer and internet access and literacy among taxpayers and preparers, and the fact that e-filed returns generate faster refunds and facilitated a variety of associated financial products. These trends by themselves reduced the rate of math errors from 4.6 percent in tax year 2000 to 1.9 percent in tax year 2009.[11] But, in tax years 2010 and 2011, two IRS administrative actions—the e-file mandate and the return preparer initiative—accelerated these trends in tax preparation and submission methods. The combined effect of the prevailing trends and these initiatives further reduced the overall rate of math errors to 1.3 percent in tax year 2011. The math error rate for paid prepared returns fell from 0.7 percent in tax year 2009 to 0.5 percent in tax year 2011.

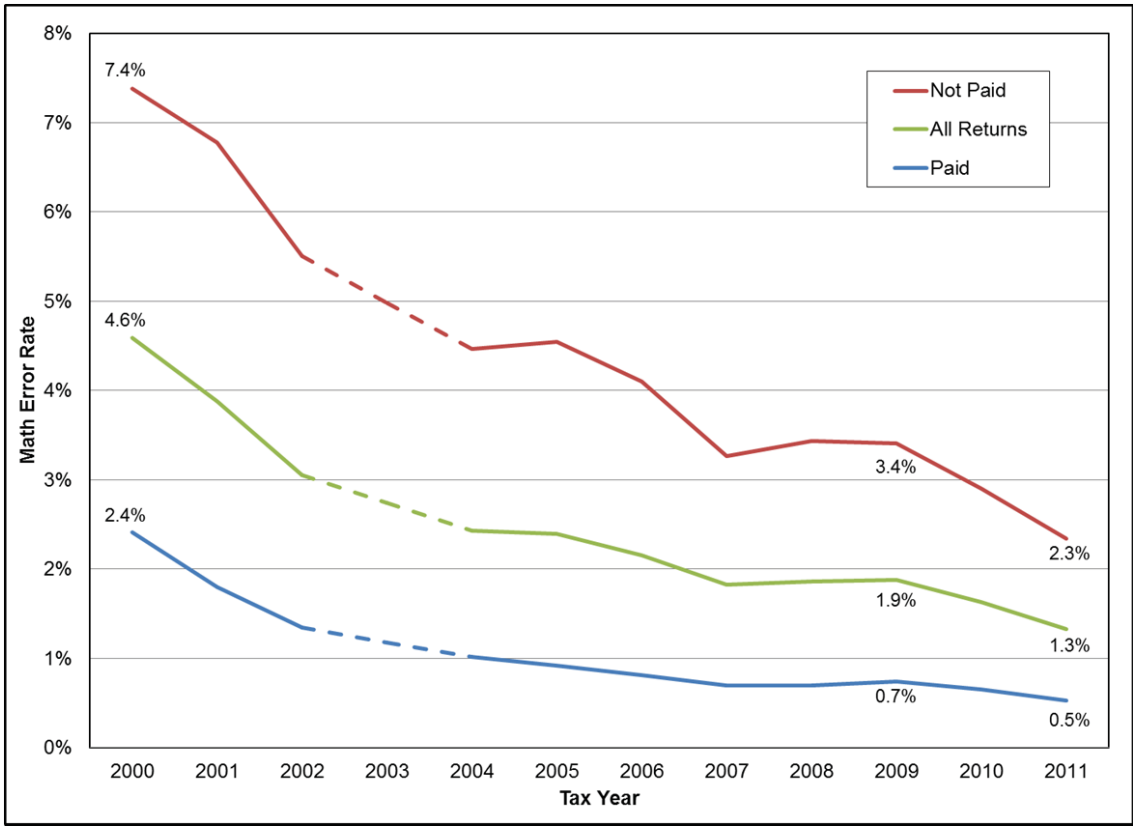
IRS implementation of the e-file mandate required preparers in processing year 2011 to electronically submit returns if they expected to prepare and file at least 100 returns. This threshold was lowered to 11 or more returns in processing year 2012. At the same time, starting at the beginning of 2011, preparers of individual tax returns were required under the return preparer initiative to obtain and use a PTIN on the returns they prepare. These new regulations increased the rate of e-filing for paid-prepared returns and accelerated the pre-existing trend increasing the share of returns prepared by larger volume preparers.

[11] In developing the math error rate estimates for each year, math errors associated with temporary or one-year tax credits, including the Rate Reduction Credit (2001), Recovery Rebate Credit (2008), Making Work Pay Credit (2008), and the First-Time Homebuyer Credit (2009 and 2010) are excluded. Tax year 2003 is omitted because of the difficulty of distinguishing errors related to recurring and non-recurring tax law changes involving the Child Tax Credit. The spike in tax year 2003 is due to an increase in errors related to the Child Tax Credit because of provisions in the Jobs and Growth Tax Relief Reconciliation Act of 2003, which raised the maximum credit per child to $1,000 and also provided that eligible taxpayers who claimed children on their 2002 return would receive advance payments in 2003 of up to $400 per child.

Compared to the existing trend, we estimate that about 6 million additional returns were e-filed in tax year 2010 and about 7 million additional returns were e-filed in tax year 2011 (Figure 10).[12] As a consequence of this higher rate of e-filing for paid prepared returns, we estimate that the number of returns with math errors was reduced by about 300,000 in tax year 2010 and about 220,000 in tax year 2011.

We also estimate that preparers preparing fewer than 100 returns prepared about 1.4 percent (1.1 million) fewer returns in tax year 2010, and 1.9 percent (1.5 million) fewer returns in tax year 2011, than would have been the case without these initiatives (Figure 11). We estimate that this acceleration in the rate of transfer of returns from smaller to larger volume preparers reduced the number of math errors by about 60,000 in tax year 2010 and 30,000 in tax year 2011.

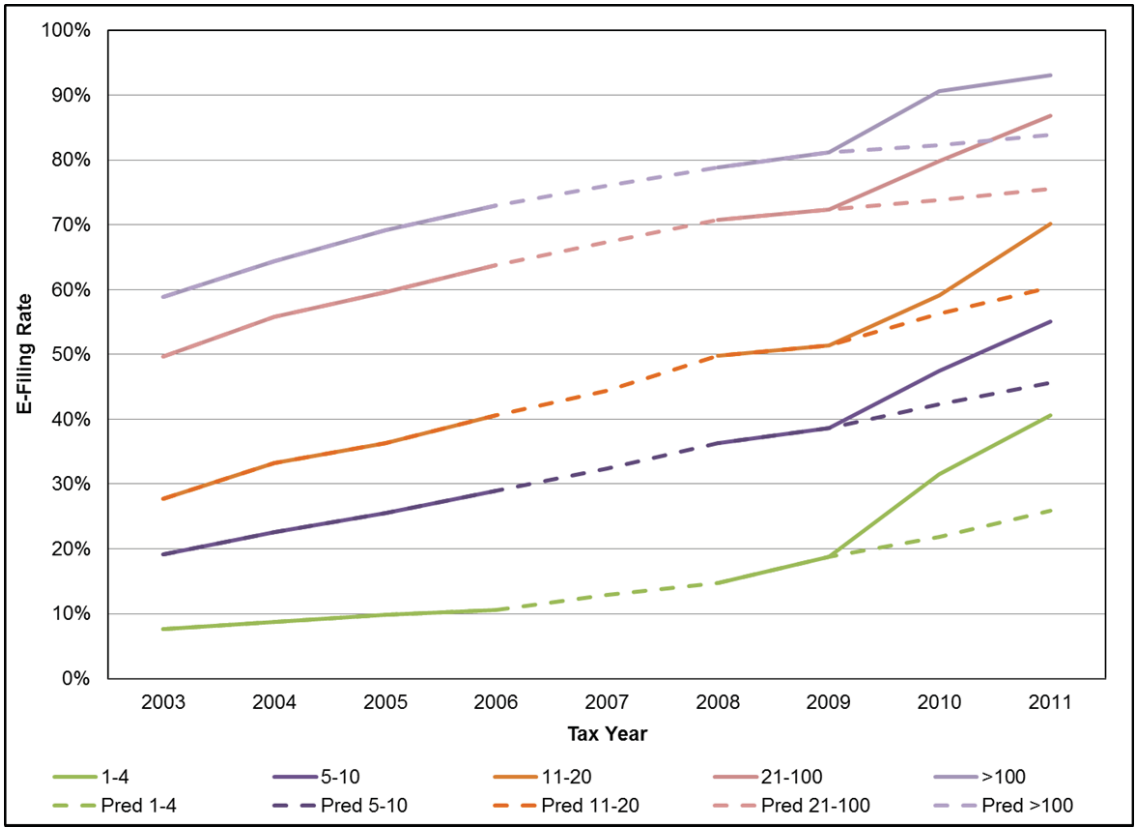
FIGURE 9. Math Error Rate for Returns With and Without Paid Preparers (Excluding Year-Specific Errors), Tax Years 2000–2011



NOTE: Math errors related to non-recurring credits, including the Rate Reduction Credit (2001), Recovery Rebate Credit (2008), Making Work Pay Credit (2009 and 2010), and the First Time Homebuyer Credit (2009 to 2011) are excluded. 2003 is omitted because of the difficulty of distinguishing errors related to recurring and non-recurring tax law changes involving the child tax credit.

[12] Tax year 2007 was excluded in generating the predicted trend line for e-file rates by volume segment and for the share of returns prepared by each volume segment because of the distorting influence of economic stimulus filings for that year.

FIGURE 10. Actual and Predicted Rates of E-Filing by Number of Returns Prepared, Tax Years 2003–2011



NOTE: Tax year 2007 was excluded in generating the predicted trend lines because of the distorting influence of economic stimulus filings in that year.

AUR Mismatches

AUR mismatch rates are more closely related to return complexity and the number of income and deduction items on the tax return that can be matched against information returns, than they are to preparation or submission method. For instance, just 8.5 percent of all returns with two or fewer matchable items have at least one AUR mismatch, while 23.8 percent of those with six or more matchable items do.[13] The weaker influence of preparation method on AUR mismatches is understandable, given that avoiding a mismatch depends on taxpayers receiving, securely storing, retrieving, and accurately reporting all of the relevant information provided on information returns—regardless of preparation method. Software and third-party preparation can help mitigate such errors by, for example, providing reminders from the previous year's return about items that may need to be reported, and by permitting the electronic uploading of information from financial institutions and employers. Despite the benefits of software and preparer assistance, however, avoiding AUR mismatches depends importantly on the taxpayer.

For all types of returns, paid-prepared returns have lower AUR mismatch rates than those self-prepared using tax preparation software. In addition, paid-prepared returns have fewer mismatches than returns self-prepared by hand for returns with three or more income or deduction items that can be matched. But returns self-prepared by hand have slightly fewer mismatches when just two or fewer items can be matched (see Appendix: Table A10).

[13] The rates reported here are the percent of returns with potential mismatches resulting from the computerized matching of information reported on tax returns with that reported by third parties on information returns (Forms 1099, 1098, W-2, etc.) Mismatches related to education credits are excluded since a large share of them are false positives and do not enter the potential work stream of the AUR program.

For paid-prepared returns, AUR mismatch rates are lower for PTIN holders than non-PTIN holders when all returns are considered and when they are broken out by the number of matched items on the return. The comparison between PTIN holders and non-PTIN holders is similar when one controls for income or return complexity rather than AUR items. The overall percentage of returns with mismatches is greater for self-reported credentialed PTIN holders than non-credentialed PTIN holders and non-PTIN holding preparers. Controlling for the number of potential AUR mismatch items shows a lower mismatch rate for credentialed preparers only in the case of returns with two or fewer AUR items but not in the case of returns with three or more AUR items. Credentialed preparers have lower mismatch rates for lower and middle levels of income and lower levels of complexity but higher mismatch rates for the higher level of income and the middle and higher levels of complexity (Appendix: Table A10).

AUR mismatches are less likely for preparers who prepare a larger number of returns, even when one controls for the different shares of preparers who are PTIN holders across the volume segments. While 17.3 percent of returns of those preparing 4 or fewer returns have mismatches, this rate diminishes to less than 13 percent for those preparing more than 500 returns (Appendix: Table A12). But, at each level of preparer activity, and regardless of the number of matchable items, the returns of PTIN holders are less prone to AUR mismatches than preparers who do not hold a PTIN (Table 8). We estimate that the acceleration in the trend in return preparation towards larger volume preparers that followed the adoption of the e-file mandate and the return preparer initiative reduced the number of returns with AUR mismatches by approximately 15,000 in each of tax years 2010 and 2011. Nonetheless, these numbers represent a very small fraction (less than 0.08 percent) of all mismatches in those years.

TABLE 8. Percentage of Returns with AUR Mismatches by Number of Returns Prepared, Preparer Type and Number of AUR Items, Tax Year 2010

Number of Returns	Preparer Type*	Number of AUR Items**			
		2 or fewer items	3 to 5 items	6 or more items	Overall
1–4	PTIN Holder	9.16	16.07	26.96	16.64
	No PTIN	9.00	18.38	30.31	18.04
5–10	PTIN Holder	8.90	15.19	26.79	15.92
	No PTIN	9.15	16.75	28.20	15.99
11–20	PTIN Holder	8.77	14.99	26.23	15.35
	No PTIN	9.27	16.81	28.16	15.71
21–100	PTIN Holder	8.74	14.16	25.28	14.46
	No PTIN	9.21	16.37	27.18	15.17
101–250	PTIN Holder	8.32	13.36	23.32	13.85
	No PTIN	8.71	15.38	25.29	14.41
251–500	PTIN Holder	7.60	12.91	21.77	13.35
	No PTIN	8.32	14.65	23.58	13.80
501–1,000	PTIN Holder	7.13	12.84	20.96	12.81
	No PTIN	9.68	14.95	23.19	14.31
>1,000	PTIN Holder	7.11	13.23	21.26	12.17
	No PTIN	12.22	15.04	23.21	14.82
Total	PTIN Holder	7.84	13.24	22.52	13.44
	No PTIN	9.39	15.52	25.16	14.70
	No preparer ID number***	9.51	16.77	27.78	15.65
	All Preparers	7.91	13.35	22.62	13.50

Source: RAS:R:TAM. Analysis of data from CDW: IRTF, CDW: AUR and CDW: Return Preparer Registration Database through 3/2014.
NOTE: 1040PR, 1040NR, 1040NR-EZ and 1040SS and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.
*Preparers are considered as having a PTIN if it was obtained by the end of 2011. Mismatches related to education credits are excluded.
**AUR Items classifies returns according to the number of income or deduction items on the return subject to the AUR matching process.
***The line "no preparer id number" shows the error rate for returns that are signed. by a third party preparer but no identifying number for the individual preparer was entered on the return.

TABLE A1. Preparers and Prepared Returns, Processing Years 2004–2013*

Processing Year	Total Preparers		Prepared Previous Year & Current Year			New Preparers			Prepared Previous Year, But Not Current Year				Total Returns Prepared		
	Number	Growth Rate	Number	Growth Rate	Share: Current Year	Number	Growth Rate	Share: Current Year	Number	Growth Rate	Share: Previous Year Attrition	Net Increase/ Decrease	Number	Share of All Returns	Returns per Preparer
All Preparers															
2004	1,229,259												78,624,459	100%	64
2005	1,170,754	-5%	717,457		61%	457,546		39%	516,157		42%	-58,611	78,347,514	100%	67
2006	1,119,959	-4%	702,821	-2%	63%	421,207	-8%	38%	472,182	-9%	40%	-50,975	80,002,267	100%	71
2007	1,080,748	-4%	693,250	-1%	64%	398,467	-5%	37%	430,778	-9%	38%	-32,311	80,841,652	100%	75
2008	1,067,268	-1%	686,983	-1%	64%	457,323	15%	43%	404,734	-6%	37%	52,589	83,901,684	100%	79
2009	1,040,919	-2%	676,234	-2%	65%	369,355	-19%	35%	468,072	16%	44%	-98,717	82,155,100	100%	79
2010	983,219	-6%	657,735	-3%	67%	329,504	-11%	34%	387,854	-17%	37%	-58,350	80,508,154	100%	82
2011	817,004	-17%	571,414	-13%	70%	249,117	-24%	30%	415,825	7%	42%	-166,708	80,943,432	100%	99
2012	734,386	-10%	551,476	-3%	75%	185,980	-25%	25%	269,055	-35%	33%	-83,075	81,424,450	100%	111
2013	663,054	-10%	521,119	-6%	79%	144,771	-22%	22%	216,337	-20%	29%	-71,566	80,988,703	100%	122
Prepared 1–4 returns															
2004	635,968												902,604	1%	1
2005	581,420	-9%	215,996		37%	365,424		63%	444,137		70%	-78,713	834,599	1%	1
2006	534,080	-8%	200,880	-7%	38%	333,200	-9%	62%	401,974	-9%	69%	-68,774	771,558	1%	1
2007	494,953	-7%	190,934	-5%	39%	304,019	-9%	61%	363,376	-10%	68%	-59,357	718,823	1%	1
2008	478,838	-3%	181,281	-5%	38%	297,557	-2%	62%	335,819	-8%	68%	-38,262	694,066	1%	1
2009	452,023	-6%	174,267	-4%	39%	277,756	-7%	61%	395,717	18%	83%	-117,961	656,514	1%	1
2010	414,238	-8%	161,599	-7%	39%	252,639	-9%	61%	312,683	-21%	69%	-60,044	605,890	1%	1
2011	274,705	-34%	117,115	-28%	43%	157,590	-38%	57%	322,575	3%	78%	-164,985	433,836	1%	2
2012	217,297	-21%	97,048	-17%	45%	120,249	-24%	55%	199,028	-38%	72%	-78,779	345,055	0%	2
2013	167,582	-23%	80,026	-18%	48%	87,556	-27%	52%	152,379	-23%	70%	-64,823	273,146	0%	2
Prepared 5–10 returns															
2004	91,030												642,086	1%	7
2005	88,041	-3%	72,263		82%	15,778		18%	17,454		19%	-1,676	621,991	1%	7
2006	84,327	-4%	69,481	-4%	82%	14,846	-6%	18%	16,853	-3%	19%	-2,007	595,631	1%	7
2007	81,519	-3%	66,922	-4%	82%	14,597	-2%	18%	15,955	-5%	19%	-1,358	576,866	1%	7
2008	78,838	-3%	64,077	-4%	81%	14,761	1%	19%	15,821	-1%	19%	-1,060	558,737	1%	7
2009	77,764	-1%	62,040	-3%	80%	15,724	7%	20%	16,965	7%	22%	-1,241	552,803	1%	7
2010	73,783	-5%	58,807	-5%	80%	14,976	-5%	20%	17,503	3%	23%	-2,527	524,693	1%	7
2011	65,198	-12%	49,267	-16%	76%	15,931	6%	24%	25,181	44%	34%	-9,250	463,548	1%	7
2012	55,058	-16%	43,758	-11%	79%	11,300	-29%	21%	19,465	-23%	30%	-8,165	394,028	0%	7
2013	49,176	-11%	38,839	-11%	79%	10,337	-9%	21%	14,333	-26%	26%	-3,996	352,579	0%	7

* Source: RAS:R:TAM tabulations using the IRTF table from December 2013 CDW
Excludes preparers that solely prepared SS/PR/NR/NR-EZ, stimulus, or TETR returns and no others. Data also exclude all volunteer preparers (e.g., VITA, TCE).

TABLE A1. Preparers and Prepared Returns, Processing Years 2004–2013*—Continued

Processing Year	Total Preparers		Prepared Previous Year & Current Year			New Preparers			Prepared Previous Year, But Not Current Year				Total Returns Prepared		
	Number	Growth Rate	Number	Growth Rate	Share: Current Year	Number	Growth Rate	Share: Current Year	Number	Growth Rate	Share: Previous Year Attrition	Net Increase/ Decrease	Number	Share of All Returns	Returns per Preparer
Prepared 11-20 returns															
2004	71,051												1,065,011	1%	15
2005	69,745	-2%	57,265		82%	12,480		18%	11,537		16%	943	1,046,944	1%	15
2006	67,528	-3%	55,692	-3%	82%	11,836	-5%	18%	11,514	0%	17%	322	1,014,301	1%	15
2007	66,458	-2%	54,255	-3%	82%	12,203	3%	18%	10,867	-6%	16%	1,336	998,581	1%	15
2008	66,262	0%	53,452	-1%	81%	12,810	5%	19%	11,480	6%	17%	1,330	997,053	1%	15
2009	66,651	1%	52,583	-2%	79%	14,068	10%	21%	12,269	7%	19%	1,799	1,004,138	1%	15
2010	63,645	-5%	51,201	-3%	80%	12,444	-12%	20%	13,623	11%	20%	-1,179	958,732	1%	15
2011	56,655	-11%	43,866	-14%	77%	12,789	3%	23%	17,222	26%	27%	-4,433	854,249	1%	15
2012	49,602	-12%	39,948	-9%	81%	9,654	-25%	19%	12,374	-28%	22%	-2,720	749,475	1%	15
2013	45,945	-7%	37,511	-6%	82%	8,434	-13%	18%	10,473	-15%	21%	-2,039	694,245	1%	15
Prepared 21–100 returns															
2004	206,095												10,895,848	14%	53
2005	206,993	0%	164,283		79%	42,710		21%	32,271		16%	10,439	10,956,905	14%	53
2006	206,052	0%	165,303	1%	80%	40,749	-5%	20%	32,166	0%	16%	8,583	10,962,208	14%	53
2007	209,246	2%	166,831	1%	80%	42,415	4%	20%	30,972	-4%	15%	11,443	11,120,789	14%	53
2008	209,512	0%	166,546	0%	79%	42,966	1%	21%	32,126	4%	15%	10,840	11,156,238	13%	53
2009	215,749	3%	171,456	3%	79%	44,293	3%	21%	33,453	4%	16%	10,840	11,440,877	14%	53
2010	206,862	-4%	171,372	0%	83%	35,490	-20%	17%	35,562	6%	16%	-72	10,981,868	14%	53
2011	190,919	-8%	152,328	-11%	80%	38,591	9%	20%	39,681	12%	19%	-1,090	10,255,770	13%	54
2012	181,253	-5%	150,970	-1%	83%	30,283	-22%	17%	29,214	-26%	15%	1,069	9,847,157	12%	54
2013	170,175	-6%	144,723	-4%	85%	25,452	-16%	15%	29,116	0%	16%	-3,664	9,254,160	11%	54
Prepared 101–250 returns															
2004	137,571												22,223,722	28%	162
2005	137,564	0%	122,020		89%	15,544		11%	9,515		7%	6,029	22,204,215	28%	161
2006	138,776	1%	123,653	1%	89%	15,123	-3%	11%	8,760	-8%	6%	6,363	22,435,723	28%	162
2007	138,555	0%	123,883	0%	89%	14,672	-3%	11%	8,612	-2%	6%	6,060	22,374,018	28%	161
2008	139,249	1%	124,829	1%	90%	14,420	-2%	10%	8,459	-2%	6%	5,961	22,514,079	27%	162
2009	137,786	-1%	125,735	1%	91%	12,051	-16%	9%	8,471	0%	6%	3,580	22,177,188	27%	161
2010	134,678	-2%	125,537	0%	93%	9,141	-24%	7%	7,584	-10%	6%	1,557	21,743,498	27%	161
2011	135,779	1%	120,528	-4%	89%	15,251	67%	11%	9,623	27%	7%	5,628	22,024,495	27%	162
2012	134,917	-1%	124,784	4%	92%	10,133	-34%	8%	7,833	-19%	6%	2,300	21,927,382	27%	163
2013	131,886	-2%	122,948	-1%	93%	8,938	-12%	7%	8,661	11%	6%	277	21,527,704	27%	163

* Source: RAS:R:TAM tabulations using the IRTF table from December 2013 CDW
Excludes preparers that solely prepared SS/PR/NR/NR-EZ, stimulus, or TETR returns and no others. Data also exclude all volunteer preparers (e.g., VITA, TCE).

TABLE A1. Preparers and Prepared Returns, Processing Years 2004–2013*—Continued

Processing Year	Total Preparers		Prepared Previous Year & Current Year			New Preparers			Prepared Previous Year, But Not Current Year				Total Returns Prepared		
	Number	Growth Rate	Number	Growth Rate	Share: Current Year	Number	Growth Rate	Share: Current Year	Number	Growth Rate	Share: Previous Year Attrition	Net Increase/ Decrease	Number	Share of All Returns	Returns per Preparer
Prepared 251–500 returns															
2004	61,896												21,276,546	27%	344
2005	61,358	-1%	59,086		96%	2,272		4%	1,112		2%	1,160	21,088,794	27%	344
2006	62,499	2%	60,238	2%	96%	2,261	0%	4%	818	-26%	1%	1,443	21,468,976	27%	344
2007	62,626	0%	60,529	0%	97%	2,097	-7%	3%	914	12%	1%	1,183	21,572,974	27%	344
2008	64,856	4%	62,655	4%	97%	2,201	5%	3%	920	1%	1%	1,281	22,400,835	27%	345
2009	62,201	-4%	60,287	-4%	97%	1,914	-13%	3%	1,034	12%	2%	880	21,455,136	26%	345
2010	61,714	-1%	60,025	0%	97%	1,689	-12%	3%	804	-22%	1%	885	21,285,244	26%	345
2011	64,566	5%	59,844	0%	93%	4,722	180%	7%	1,363	70%	2%	3,359	22,282,874	28%	345
2012	66,060	2%	64,117	7%	97%	1,943	-59%	3%	992	-27%	2%	951	22,794,711	28%	345
2013	67,654	2%	65,863	3%	97%	1,791	-8%	3%	1,187	20%	2%	604	23,390,668	29%	346
Prepared 501–1,000 returns															
2004	20,658												13,743,554	17%	665
2005	20,587	0%	20,122		98%	465		2%	115		1%	350	13,721,039	18%	666
2006	21,298	3%	20,832	4%	98%	466	0%	2%	87	-24%	0%	379	14,246,589	18%	669
2007	21,693	2%	21,221	2%	98%	472	1%	2%	69	-21%	0%	403	14,510,790	18%	669
2008	23,461	8%	22,942	8%	98%	519	10%	2%	93	35%	0%	426	15,736,072	19%	671
2009	22,577	-4%	22,086	-4%	98%	491	-5%	2%	141	52%	1%	350	15,126,807	18%	670
2010	22,288	-1%	21,754	-2%	98%	534	9%	2%	80	-43%	0%	454	14,931,200	19%	670
2011	23,251	4%	21,622	-1%	93%	1,629	205%	7%	159	99%	1%	1,470	15,556,346	19%	669
2012	24,148	4%	23,575	9%	98%	573	-65%	2%	131	-18%	1%	442	16,157,723	20%	669
2013	24,585	2%	24,085	2%	98%	500	-13%	2%	164	25%	1%	336	16,432,011	20%	668
Prepared >1,000 returns															
2004	4,990												7,875,088	10%	1,578
2005	5,046	1%	4,909		97%	137		3%	16		0%	121	7,873,027	10%	1,560
2006	5,399	7%	5,269	7%	98%	130	-5%	2%	10	-38%	0%	120	8,507,281	11%	1,576
2007	5,698	6%	5,572	6%	98%	126	-3%	2%	13	30%	0%	113	8,968,811	11%	1,574
2008	6,252	10%	6,100	9%	98%	152	21%	2%	16	23%	0%	136	9,844,604	12%	1,575
2009	6,168	-1%	6,013	-1%	97%	155	2%	3%	22	38%	0%	133	9,741,637	12%	1,579
2010	6,011	-3%	5,868	-2%	98%	143	-8%	2%	15	-32%	0%	128	9,477,029	12%	1,577
2011	5,931	-1%	5,422	-8%	91%	509	256%	9%	21	40%	0%	488	9,072,314	11%	1,530
2012	6,051	2%	5,900	9%	98%	151	-70%	2%	18	-14%	0%	133	9,208,919	11%	1,522
2013	6,051	0%	5,916	0%	98%	135	-11%	2%	24	33%	0%	111	9,064,190	11%	1,498

* Source: RAS:R:TAM tabulations using the IRTF table from December 2013 CDW

Excludes preparers that solely prepared SS/PR/NR/NR-EZ, stimulus, or TETR returns and no others. Data also exclude all volunteer preparers (e.g., VITA, TCE).

TABLE A2. Number, Market Shares, Percent Change, and Average—by Preparers, Returns Prepared, Volume Segment, and Selected Ranges of Processing Years*

Volume Segment (Number of Returns Prepared)	Preparers					Returns Prepared				
	PY2004	PY2010	PY2011	PY2012	PY2013	PY2004	PY2010	PY2011	PY2012	PY2013
	Number					Number				
All Preparers	1,229,259	983,219	817,004	734,386	663,054	78,624,459	80,508,154	80,943,432	81,424,450	80,988,703
1–4	635,968	414,238	274,705	217,297	167,582	902,604	605,890	433,836	345,055	273,146
5–10	91,030	73,783	65,198	55,058	49,176	642,086	524,693	463,548	394,028	352,579
11–20	71,051	63,645	56,655	49,602	45,945	1,065,011	958,732	854,249	749,475	694,245
21–100	206,095	206,862	190,919	181,253	170,175	10,895,848	10,981,868	10,255,770	9,847,157	9,254,160
101–250	137,571	134,678	135,779	134,917	131,886	22,223,722	21,743,498	22,024,495	21,927,382	21,527,704
251–500	61,896	61,714	64,566	66,060	67,654	21,276,546	21,285,244	22,282,874	22,794,711	23,390,668
501–1,000	20,658	22,288	23,251	24,148	24,585	13,743,554	14,931,200	15,556,346	16,157,723	16,432,011
> 1,000	4,990	6,011	5,931	6,051	6,051	7,875,088	9,477,029	9,072,314	9,208,919	9,064,190
	Market Shares					Market Shares				
1–4	52%	42%	34%	30%	25%	1%	1%	1%	0%	0%
5–10	7%	8%	8%	7%	7%	1%	1%	1%	0%	0%
11–20	6%	6%	7%	7%	7%	1%	1%	1%	1%	1%
21–100	17%	21%	23%	25%	26%	14%	14%	13%	12%	11%
101–250	11%	14%	17%	18%	20%	28%	27%	27%	27%	27%
251–500	5%	6%	8%	9%	10%	27%	26%	28%	28%	29%
501–1,000	2%	2%	3%	3%	4%	17%	19%	19%	20%	20%
> 1,000	0%	1%	1%	1%	1%	10%	12%	11%	11%	11%
	Percent Change					Percent Change				
	PY 04 to PY10	PY04 to PY13	PY10 to PY11	PY10 to PY12	PY10 to PY13	PY 04 to PY10	PY04 to PY13	PY10 to PY11	PY10 to PY12	PY10 to PY13
All Preparers	-20%	-40%	-17%	-25%	-33%	2%	3%	1%	1%	1%
1-4	-35%	-66%	-34%	-48%	-60%	-33%	-70%	-28%	-43%	-55%
5-10	-19%	-40%	-12%	-25%	-33%	-18%	-45%	-12%	-25%	-33%
11-20	-10%	-30%	-11%	-22%	-28%	-10%	-35%	-11%	-22%	-28%
21-100	0%	-12%	-8%	-12%	-18%	1%	-15%	-7%	-10%	-16%
101-250	-2%	-2%	1%	0%	-2%	-2%	-3%	1%	1%	-1%
251-500	0%	7%	5%	7%	10%	0%	10%	5%	7%	10%
501-1,000	8%	17%	4%	8%	10%	9%	20%	4%	8%	10%
> 1,000	20%	21%	-1%	1%	1%	20%	15%	-4%	2%	-2%

	Average Number of Returns Prepared by Preparer				
	PY2004	PY2010	PY2011	PY2012	PY2013
All Preparers	64	82	99	111	122
1-4	1	1	2	2	2
5-10	7	7	7	7	7
11-20	15	15	15	15	15
21-100	53	53	54	54	54
101-250	162	161	162	163	163
251-500	344	345	345	345	346
501-1,000	665	670	669	669	668
> 1,000	1,578	1,577	1,530	1,522	1,498

* Source: RAS:R:TAM tabulations using the IRTF table from December 2013 CDW
Excludes preparers that solely prepared SS/PR/NR/NR-EZ, stimulus, or TETR returns and no others. Data also exclude all volunteer preparers (e.g., VITA, TCE).

TABLE A3. Preparer Dynamics, Processing Years 2005–2013*

Processing Year	New Preparers			Prepared Previous Year, But Not Current Year				Prepared Both Years		
	Number	Growth Rate	Share: Current Year	Number	Growth Rate	Share: Previous Year Attrition	Net Increase/ Decrease	Number	Growth Rate	Share: Current Year
2005	457,546		39%	516,157		42%	-58,611	717,457		61%
2006	421,207	-8%	38%	472,182	-9%	40%	-50,975	702,821	-2%	63%
2007	398,467	-5%	37%	430,778	-9%	38%	-32,311	693,250	-1%	64%
2008	457,323	15%	43%	404,734	-6%	37%	52,589	686,983	-1%	64%
2009	369,355	-19%	35%	468,072	16%	44%	-98,717	676,234	-2%	65%
2010	329,504	-11%	34%	387,854	-17%	37%	-58,350	657,735	-3%	67%
2011	249,117	-24%	30%	415,825	7%	42%	-166,708	571,414	-13%	70%
2012	185,980	-25%	25%	269,055	-35%	33%	-83,075	551,476	-3%	75%
2013	144,771	-22%	22%	216,337	-20%	29%	-71,566	521,119	-6%	79%

* Source: RAS:R:TAM tabulations using the IRTF table from December 2013 CDW
Excludes preparers that solely prepared SS/PR/NR/NR-EZ, stimulus, or TETR returns and no others. Data also exclude all volunteer preparers (e.g., VITA, TCE).

TABLE A4. Percent of Returns with Math Errors by Preparation Method, Tax Year 2010

Preparation Method	With Errors	No Errors	Total
Self-prepared by hand	37.0	63.0	100.0
Self-prepared with software	1.8	98.2	100.0
Paid-prepared	1.3	98.7	100.0
Total	3.6	96.4	100.0

Source: RAS:R:TAM. Analysis of data from CDW: IRTF and IMF through 3/2014.
NOTE: Forms 1040PR, 1040NR, 1040NR-EZ, and 1040SS, and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.

TABLE A5. Percent of Returns with Math Errors by Preparation and Submission Method, Tax Year 2010

Preparation Method	With Errors	No Errors	Total
Self-prepared by hand	37.0	63.0	100.0
Self-prepared with software e-filed	1.1	98.9	100.0
Self-prepared with software on paper	5.3	94.7	100.0
Paid-prepared by hand	18.0	82.0	100.0
Paid-prepared with software e-filed	0.6	99.4	100.0
Paid-prepared with software on paper	5.5	94.5	100.0
Total	3.6	96.4	100.0

Source: RAS:R:TAM. Analysis of data from CDW: IRTF and IMF through 3/2014.
NOTE: 1040PR, 1040NR, 1040NR-EZ and 1040SS and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.

TABLE A6. Percent of Returns with Math Errors by Preparation and Submission Method and Preparer Type, Tax Year 2010

Preparation and Submission Method and Preparer Type	With Errors	No Errors	Total
Self-prepared by hand	37.0	63.0	100.0
Self-prepared with software e-filed	1.1	98.9	100.0
Self-prepared with software on paper	5.3	94.7	100.0
Paid with software e-filed PTIN holder	0.6	99.4	100.0
Paid with software e-filed no PTIN	0.9	99.1	100.0
Paid with software on paper PTIN holder	5.2	94.8	100.0
Paid with software on paper no PTIN	7.2	92.8	100.0
Paid by hand PTIN holder	15.1	84.9	100.0
Paid by hand no PTIN	26.1	73.9	100.0
Paid PTIN holder	1.2	98.8	100.0
Paid no PTIN	4.1	95.9	100.0
Total	3.6	96.4	100.0

Source: RAS:R:TAM. Analysis of data from CDW: IRTF and IMF through 3/2014.
NOTE: Forms 1040PR, 1040NR, 1040NR-EZ, and 1040SS, and returns prepared at VITA sites are excluded. Limited to returns filed in 2011. Preparers are considered as having a PTIN if it was obtained by the end of 2011.

TABLE A7. Percent of Returns with Math Errors by Preparation and Submission Method and Preparer Type, Tax Year 2010

Preparation and Submission Method and Preparer Type	With Errors	No Errors	Total
Self-prepared by hand	37.0	63.0	100.0
Self-prepared with software e-filed	1.1	98.9	100.0
Self-prepared with software on paper	5.3	94.7	100.0
Paid with software e-filed credentialed	0.7	99.3	100.0
Paid with software e-filed not credentialed	0.6	99.4	100.0
Paid with software on paper credentialed	4.4	95.6	100.0
Paid with software on paper not credentialed	6.1	93.9	100.0
Paid by hand credentialed	12.7	87.3	100.0
Paid by hand not credentialed	19.4	80.6	100.0
Paid credentialed	1.1	98.9	100.0
Paid not credentialed	1.5	98.5	100.0
Total	3.6	96.4	100.0

Source: RAS:R:TAM. Analysis of data from CDW: IRTF and IMF through 3/2014.
NOTE: Forms 1040PR, 1040NR, 1040NR-EZ, and 1040SS, and returns prepared at VITA sites are excluded. Limited to returns filed in 2011. The designation 'credentialed' means that the preparer self-reported in the registration for a PTIN that he/she is a Certified Public Accountant, Enrolled Agent, or an Attorney.

TABLE A8. Percent of Returns with Math Errors by Number of Returns Prepared by Tax Preparer, Tax Year 2010

Number of Returns	With Errors	No Errors	Total
1–4	7.9	92.4	100.0
5–10	4.4	95.8	100.0
11–20	3.2	96.9	100.0
21–100	1.7	98.2	100.0
101–250	1.1	98.9	100.0
251–500	1.1	98.9	100.0
501–1,000	1.2	98.8	100.0
>1,000	1.4	98.6	100.0
Total with preparer ID number	1.3	98.7	100.0
Without preparer ID number*	6.7	93.3	100.0
All returns	1.3	98.7	100.0

* Returns that are signed by a third-party preparer but no identifying number for the individual preparer was entered on the return.
Source: RAS:R:TAM. Analysis of data from CDW: IRTF and IMF through 3/2014.
NOTE: Forms 1040PR, 1040NR, 1040NR-EZ, and 1040SS, and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.

TABLE A9. Percent of Returns with Math Errors by Number of Returns Prepared by Tax Preparer and Preparer Type, Tax Year 2010

Number of Returns	Preparer Type*	With Errors	Share of Returns
1–4	Credentialed	4.5	26.8
	Not Credentialed	8.9	73.2
5–10	Credentialed	3.5	43.4
	Not Credentialed	5.0	56.6
11–20	Credentialed	2.8	42.4
	Not Credentialed	3.4	57.6
21–100	Credentialed	1.9	39.2
	Not Credentialed	1.7	60.8
101–250	Credentialed	1.2	44.6
	Not Credentialed	1.1	55.4
251–500	Credentialed	1.0	51.0
	Not Credentialed	1.1	49.0
501–1,000	Credentialed	0.9	49.1
	Not Credentialed	1.4	50.9
>1,000	Credentialed	1.0	39.7
	Not Credentialed	1.5	60.3
Total	Credentialed	1.1	41.3
	Not Credentialed	1.4	58.7

* Credentialed means that the preparer self-reported in the registration for a PTIN that he/she is a Certified Public Accountant, Enrolled Agent, or Attorney.
Source: RAS:R:TAM. Analysis of data from CDW: IRTF and IMF through 3/2014.
NOTE: Forms 1040PR, 1040NR, 1040NR-EZ, and 1040SS, and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.

TABLE A10. AUR Apparent Mismatch Rates by Preparation Method, Income, Complexity, and the Number of AUR Items on the Return, Tax Year 2010

Preparation Method		AUR Apparent Mismatch Rate (%)									
		Income[1]			Complexity[2]			Number of AUR Items[3]			Overall
		Less than $37,000	$37,000 to $75,000	Greater than $75,000	Simple	Intermediate	Complex	2 or fewer items	3 to 5 items	6 or more items	
Self-Prepared	By Hand	8.52	16.84	25.98	8.83	19.37	20.48	7.37	15.21	29.57	13.49
	With Software	10.48	17.63	23.86	10.13	17.37	19.83	9.83	15.56	26.42	15.06
	With Software E-Filed	11.06	21.22	23.42	10.56	17.57	20.04	10.55	15.58	26.02	15.92
	With Software on Paper	7.72	16.41	26.32	7.91	16.24	19.26	6.57	15.49	28.21	14.01
Paid Preparer	By Hand	8.57	21.23	26.29	8.52	15.33	18.38	8.03	15.07	26.21	15.24
	With Software E-filed	8.24	14.12	22.93	7.13	13.11	17.70	8.06	13.00	22.06	13.28
	With Software on Paper	7.09	13.81	27.28	7.42	13.68	19.45	6.85	15.86	27.45	14.29
	Credentialed	7.02	14.26	24.53	6.35	13.34	20.23	7.11	13.73	23.30	15.57
	Not Credentialed	8.59	14.70	21.58	7.49	13.10	14.96	8.18	13.08	21.35	12.07
	PTIN Holder using PTIN	8.05	14.38	23.36	7.04	13.12	17.94	7.84	13.24	22.52	13.44
	No PTIN	9.59	17.91	26.25	10.18	14.50	18.42	9.42	15.80	25.70	14.91
	All Paid Preparers	8.12	14.52	23.46	7.17	13.18	17.96	7.91	13.35	22.62	13.50
	All Returns	9.03	15.79	23.70	8.73	15.01	18.46	8.60	14.30	23.97	14.07

Source: RAS:R:TAM. Analysis of data from CDW: IRTF, AUR and Return Preparer Registration Database through 3/2014.

Note: Forms 1040PR, 1040NR, 1040NR-EZ, and 1040SS, and returns prepared at VITA sites are excluded. Limited to returns filed in 2011. Mismatches related to education credits are excluded. Preparers are considered as having a PTIN if it was obtained by the end of 2011. The designation 'credentialed' means that the preparer self-reported in the registration for a PTIN that he/she is a Certified Public Accountant, Enrolled Agent, or an Attorney.

[1]Income is defined as total positive income.

[2]Complexity is defined in agreement with the definition in the ETA IMF Marketing Database.

[3]The Number of AUR Items classifies returns according to the number of income or deduction items on the return subject to the AUR matching process.

TABLE A11. AUR Apparent Mismatch Rates by Paid Preparer Type, Submission Method, Income, Complexity, and Number of AUR Items, Tax Year 2010

Paid Preparation Method		AUR Apparent Mismatch Rate (%)									
		Income[1]			Complexity[2]			Number of AUR Items[3]			Overall
		Less than $37,000	$37,000 to $75,000	Greater than $75,000	Simple	Intermediate	Complex	2 or fewer items	3 to 5 items	6 or more items	
Credentialed	E-Filed	6.97	14.00	23.99	6.22	13.16	19.87	6.96	13.25	22.76	15.24
	Software on Paper	7.43	16.51	28.60	7.51	15.19	22.75	8.18	17.30	29.01	18.31
	By Hand	7.67	17.24	27.95	7.09	15.09	21.56	7.79	15.90	27.89	15.98
Not Credentialed	E-Filed	8.79	14.46	20.84	7.48	13.09	14.63	8.42	12.82	20.67	11.94
	Software on Paper	6.96	16.21	25.36	7.39	12.97	16.33	6.41	14.87	25.53	12.80
	By Hand	8.78	17.21	25.49	8.84	15.39	17.26	8.09	14.85	25.56	14.09
PTIN Holder	E-Filed	8.17	14.17	22.84	7.01	13.08	17.69	7.97	12.93	22.01	13.27
	Software on Paper	6.98	16.10	27.12	7.33	13.54	19.51	6.78	15.76	27.41	14.86
	By Hand	7.55	16.38	25.46	7.24	14.50	18.01	7.06	14.32	25.42	13.75
No PTIN	E-Filed	10.24	17.79	24.96	11.01	14.23	17.89	10.26	15.27	24.33	14.65
	Software on Paper	7.86	17.86	27.87	8.15	14.73	19.05	7.36	16.56	27.78	15.16
	By Hand	11.27	19.60	28.94	11.95	17.67	19.49	10.67	17.20	28.75	16.61
	All Paid Preparers	8.12	14.52	23.46	7.17	13.18	17.96	7.91	13.35	22.62	13.50

Source: RAS:R:TAM. Analysis of data from CDW: IRTF, AUR and Return Preparer Registration Database through 3/2014.

NOTE: Forms 1040PR, 1040NR, 1040NR-EZ, and 1040SS, and returns prepared at VITA sites are excluded. Limited to returns filed in 2011. Mismatches related to education credits are excluded. Preparers are considered as having a PTIN if it was obtained by the end of 2011. The designation 'credentialed' means that the preparer self-reported in the registration for a PTIN that he/she is a Certified Public Accountant, Enrolled Agent, or an Attorney.

[1]Income is defined as total positive income.

[2]Complexity is defined in agreement with the definition in the ETA IMF Marketing Database.

[3]The Number of AUR Items classifies returns according to the number of income or deduction items on the return subject to the AUR matching process.

TABLE A12. AUR Apparent Mismatch Rates by Number of Returns Prepared by Tax Preparer, Income, Complexity, and Number of AUR Items, Tax Year 2010

Number of Returns	AUR Apparent Mismatch Rate (%)									
	Income[1]			Complexity[2]			Number of AUR Items[3]			Overall
	Less than $37,000	$37,000 to $75,000	Greater than $75,000	Simple	Intermediate	Complex	2 or fewer items	3 to 5 items	6 or more items	
1–4	9.24	17.34	28.90	8.57	15.67	22.30	9.08	17.22	28.54	17.34
5–10	9.10	16.20	27.05	8.20	14.85	20.89	8.94	15.43	26.96	15.93
11–20	9.03	16.03	26.72	7.98	14.54	20.58	8.83	15.17	26.37	15.39
21–100	8.95	15.16	25.68	7.56	13.98	20.13	8.77	14.26	25.35	14.49
101–250	8.42	14.42	24.08	7.21	13.43	19.03	8.33	13.42	23.36	13.87
251–500	7.80	14.08	22.81	6.78	12.99	17.70	7.62	12.95	21.80	13.36
501–1,000	7.57	14.25	22.00	6.73	12.83	16.39	7.21	12.90	21.01	12.85
>1,000	7.75	15.11	21.98	7.88	12.44	15.04	7.33	13.30	21.32	12.27
Had Preparer ID Number	8.10	14.49	23.42	7.15	13.16	17.95	7.90	13.32	22.58	13.48
No Preparer ID Number	9.90	18.57	28.14	10.45	15.19	18.85	9.51	16.77	27.78	15.65
All Returns	8.12	14.52	23.46	7.17	13.18	17.96	7.91	13.35	22.62	13.50

Source: RAS:R:TAM. Analysis of data from CDW: IRTF, AUR and Return Preparer Registration Database through 3/2014.

NOTE: 1040PR, 1040NR, 1040NR-EZ and 1040SS and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.

Mismatches related to education credits are excluded. The line "no preparer id number" shows the error rate for returns that are signed by a third party preparer but no identifying number for the individual preparer was entered on the return.

[1]Income is defined as total positive income.

[2]Complexity is defined in agreement with the definition in the ETA IMF Marketing Database.

[3]The column labeled AUR Items classifies returns according to the number of income or deduction items on the return subject to the AUR matching process.

TABLE A13. AUR Apparent Mismatch Rates by Number of Returns Prepared, Registration Status, Income, Complexity, and Number of AUR Items, Tax Year 2010

Number of Returns	Registration Status	AUR Apparent Mismatch Rate (%)									
		Income[1]			Complexity[2]			Number of AUR Items[3]			Overall
		Less than $37,000	$37,000 to $75,000	Greater than $75,000	Simple	Intermediate	Complex	2 or fewer items	3 to 5 items	6 or more items	
1–4	PTIN Holder	9.01	16.10	27.67	8.22	15.02	21.60	9.16	16.07	26.96	16.64
	No PTIN	9.46	18.59	30.17	8.91	16.37	22.95	9.00	18.38	30.31	18.04
5–10	PTIN Holder	8.96	15.77	26.89	8.03	14.60	21.05	8.90	15.19	26.79	15.92
	No PTIN	9.76	18.53	28.24	9.10	16.21	19.89	9.15	16.75	28.20	15.99
11–20	PTIN Holder	8.94	15.72	26.61	7.86	14.36	20.67	8.77	14.99	26.23	15.35
	No PTIN	9.81	18.87	28.17	9.02	16.19	19.66	9.27	16.81	28.16	15.71
21–100	PTIN Holder	8.90	15.02	25.61	7.51	13.91	20.16	8.74	14.16	25.28	14.46
	No PTIN	9.83	17.90	27.49	8.50	15.43	19.48	9.21	16.37	27.18	15.17
101–250	PTIN Holder	8.40	14.35	24.05	7.18	13.40	19.05	8.32	13.36	23.32	13.85
	No PTIN	9.21	17.22	25.60	8.38	14.51	18.34	8.71	15.38	25.29	14.41
251-500	PTIN Holder	7.77	14.02	22.78	6.74	12.97	17.70	7.60	12.91	21.77	13.35
	No PTIN	8.69	16.48	24.41	8.44	13.59	17.59	8.32	14.65	23.58	13.80
501–1,000	PTIN Holder	7.50	14.17	21.96	6.59	12.79	16.39	7.13	12.84	20.96	12.81
	No PTIN	9.71	17.37	23.89	11.61	14.00	16.60	9.68	14.95	23.19	14.31
>1,000	PTIN Holder	7.64	14.91	21.93	7.46	12.42	15.02	7.11	13.23	21.26	12.17
	No PTIN	10.48	20.23	23.63	18.19	13.08	15.62	12.22	15.04	23.21	14.82
Total with ID Number	Credentialed	7.02	14.26	24.53	6.35	13.34	20.23	7.11	13.73	23.30	15.57
	Not Credentialed	8.57	14.65	21.45	7.45	13.07	14.87	8.16	13.02	21.23	12.01
No preparer ID number		9.90	18.57	28.14	10.45	15.19	18.85	9.51	16.77	27.78	15.65
All Returns		8.12	14.52	23.46	7.17	13.18	17.96	7.91	13.35	22.62	13.50

Source: RAS:R:TAM. Analysis of data from CDW: IRTF, AUR and Return Preparer Registration Database through 3/2014.

NOTE: Forms 1040PR, 1040NR, 1040NR-EZ, and 1040SS, and returns prepared at VITA sites are excluded. Limited to returns filed in 2011. Mismatches related to education credits are excluded. Preparers are considered as having a PTIN if it was obtained by the end of 2011. The designation 'credentialed' means that the preparer self-reported in the registration for a PTIN that he/she is a Certified Public Accountant, Enrolled Agent, or an Attorney.

[1]Income is defined as total positive income.

[2]Complexity is defined in agreement with the definition in the ETA IMF Marketing Database.

[3]The Number of AUR Items classifies returns according to the number of income or deduction items on the return subject to the AUR matching process.

TABLE A14. AUR Apparent Mismatch Rates by Number of Returns, Credentials, Income, Complexity, and Number of AUR Items, Tax Year 2010

Number of Returns	Credentialed Status	AUR Apparent Mismatch Rate (%)									
		Income[1]			Complexity[2]			Number of AUR Items[3]			Overall
		Less than $37,000	$37,000 to $75,000	Greater than $75,000	Simple	Intermediate	Complex	2 or fewer items	3 to 5 items	6 or more items	
1–4	Credentialed	8.26	15.98	28.81	7.70	15.63	23.91	8.84	16.92	28.16	18.96
	Not Credentialed	9.46	17.77	28.94	8.75	15.68	21.65	9.13	17.33	28.74	16.81
5–10	Credentialed	7.89	15.68	27.92	7.43	14.95	23.44	7.87	15.97	27.90	18.32
	Not Credentialed	9.58	16.52	25.73	8.49	14.79	18.25	9.27	15.03	25.56	14.33
11–20	Credentialed	7.57	15.52	27.65	7.18	14.60	23.19	7.94	15.77	27.08	18.02
	Not Credentialed	9.55	16.36	25.21	8.24	14.51	17.66	9.07	14.76	25.22	13.71
21–100	Credentialed	7.39	15.38	27.34	6.97	14.28	23.16	7.73	15.43	26.58	18.04
	Not Credentialed	9.40	15.04	22.73	7.69	13.86	16.20	8.99	13.59	23.05	12.52
101–250	Credentialed	7.14	14.59	25.52	6.61	13.56	21.49	7.43	14.27	24.31	16.63
	Not Credentialed	8.92	14.31	21.14	7.40	13.36	15.09	8.60	12.82	21.25	11.95
251–500	Credentialed	6.95	14.03	23.81	6.30	13.13	19.65	7.12	13.36	22.44	15.18
	Not Credentialed	8.26	14.14	20.73	7.02	12.89	14.52	7.84	12.59	20.38	11.75
501–1,000	Credentialed	6.82	13.72	22.56	6.00	12.97	18.11	6.75	12.80	21.26	14.10
	Not Credentialed	7.97	14.68	21.09	7.13	12.74	14.30	7.41	12.98	20.57	11.88
>1,000	Credentialed	6.87	13.97	22.10	5.92	12.99	16.79	6.35	12.97	21.25	13.11
	Not Credentialed	8.05	15.66	21.87	8.66	12.24	13.87	7.62	13.45	21.38	11.90
Total with ID Number	Credentialed	7.02	14.26	24.53	6.35	13.34	20.23	7.11	13.73	23.30	15.57
	Not Credentialed	8.57	14.65	21.45	7.45	13.07	14.87	8.16	13.02	21.23	12.01
No preparer ID number		9.90	18.57	28.14	10.45	15.19	18.85	9.51	16.77	27.78	15.65
All Returns		8.12	14.52	23.46	7.17	13.18	17.96	7.91	13.35	22.62	13.50

Source: RAS:R:TAM. Analysis of data from CDW: IRTF, CDW: AUR and CDW: Return Preparer Registration Database through 3/2014.
NOTE: 1040PR, 1040NR, 1040NR-EZ and 1040SS and returns prepared at VITA sites are excluded. Limited to returns filed in 2011.
Mismatches related to education credits are excluded. The designation 'credentialed' means that the preparer self-reported in the registration for a PTIN that he/she is a Certified Public Accountant, Enrolled Agent, or an Attorney.
The line "no preparer id number" shows the error rate for returns that are signed. by a third party preparer but no identifying number for the individual preparer was entered on the return.
[1]Income is defined as total positive income.
[2]Complexity is defined in agreement with the definition in the ETA IMF Marketing Database.
[3]AUR Items classifies returns according to the number of income or deduction items on the return subject to the AUR matching process.

Changes in EITC Eligibility and Participation, 2005–2009

Maggie R. Jones[1]

Center for Administrative Records Research and Applications, U.S. Census Bureau

1 Introduction

Participation in the Earned Income Tax Credit (EITC) has long been of interest to economists and policymakers. Over the program's lifespan, several estimates of "take-up" have been produced, all of which relied on survey data for an estimate of the eligible population. Under a data-sharing agreement with the Internal Revenue Service (IRS), the Census Bureau has undertaken the calculation of EITC take-up using linked survey and tax data, with estimates produced and reported to the IRS for years 2005 to 2009. An IRS report (Plueger, 2009) was published describing the data linkage and modeling used to produce the estimates for 2005. The current work takes these estimates a step further, examining changes in eligibility and take-up over the 5 years in question. The purpose of the work is threefold. First, because estimates of EITC eligibility and take-up are scarce, one purpose is simply to publish general information about eligibility and take-up estimates for years not currently available. A second purpose is to describe changes in eligibility for the program and its take-up, broken out by demographic group, over what has become known as the Great Recession. The final purpose is to assess, from these changes, the extent to which the EITC as a policy targeted those groups most affected by the recession.

This work contributes to the literature by providing more precise estimates of EITC take-up than have been available previously. It also is unique in that it describes the program's availability to its intended target populations during a deep economic downturn. As in all recessions, the latest downturn affected types of skill and family groups differently, and the work presented here makes an attempt to assess the effectiveness of the EITC in reaching these groups. These assessments are essentially descriptive in nature, and further work is needed to assess the dynamic connection between eligibility and unemployment.

The paper proceeds as follows: Section 1 provides background on the EITC and summarizes some previous literature regarding its take-up, as well as key literature on the take-up of social programs in general during hard economic times. Section 2 describes the data used, including details on the linking of records, the sample selection, and outside sources of data; it also provides details on the generation of summary statistics and their presentation. Section 3 describes the difference-in-differences model used to analyze differential eligibility and take-up among key demographic and skill groups, going into detail on the model used. Section 4 presents the results, and section 5 concludes with a summary of the findings and areas for possibly future research.

2 Background

2.1 The EITC

The EITC has become the largest cash-transfer program in the United States. It is a refundable tax credit, meaning that it provides a credit to taxpayers even if they have no Federal income tax liability. The vast majority of participants receive the credit upon filing their taxes; less than 2 percent take advantage of a program that allows employers to distribute an expected credit over the course of the tax year in an employee's paycheck.

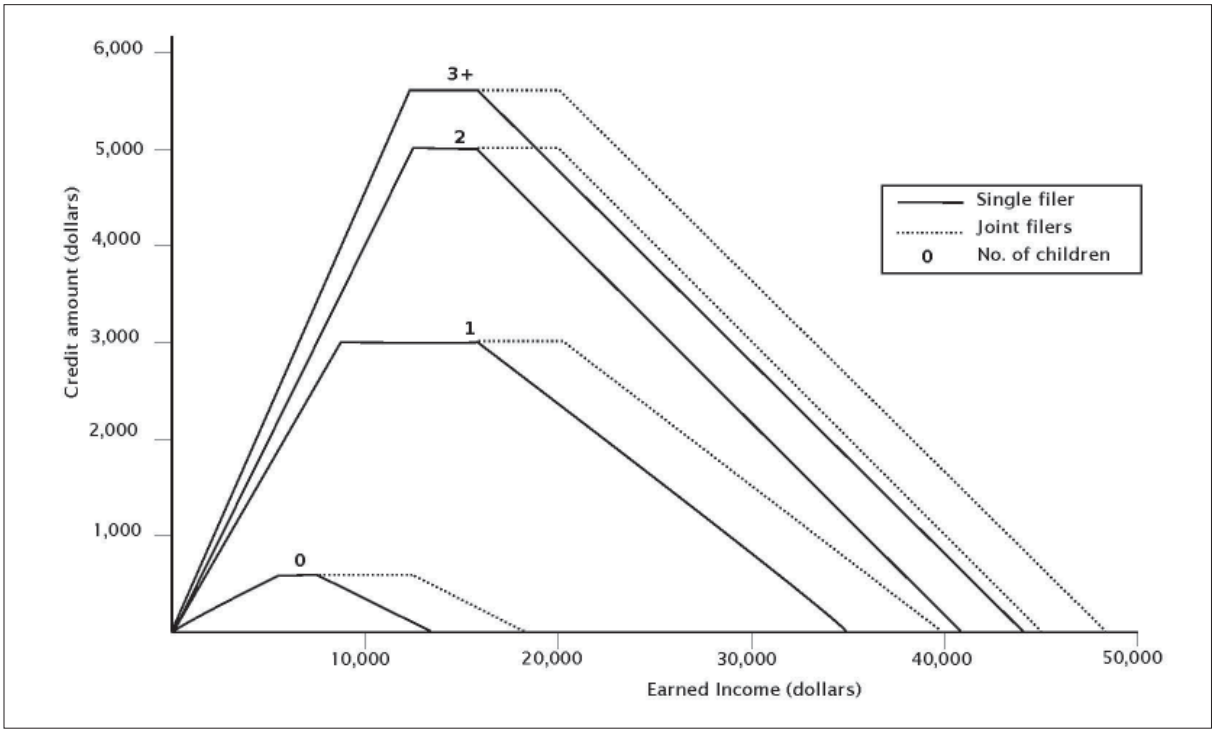
Lawmakers' original intent for instituting the credit was as an offset to payroll taxes, which represent a disproportionately high percentage of the earnings of low-income workers. At its inception in 1975, credit rates

[1] This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on technical, statistical, or methodological issues are those of the author and not necessarily those of the Census Bureau.

were quite low, but the program has seen substantial expansions over the subsequent decades. In particular, the Clinton administration's revamping of the credit in the mid-1990s, in tandem with welfare reform, greatly expanded the phase-in rate and the maximum credit a family could receive. The credit formula currently takes the shape shown in Figure 1. The credit phases in at a percentage of earnings until leveling off; it then phases out as a percentage of income until the credit reaches zero. Phase-in percent, phase-out percent, and maximum credit level are determined by family structure.

As shown in Figure 1, a person must have some earned income to be eligible for the credit, and having total income above a certain amount results in ineligibility. Thus labor-force participants may become ineligible through one of two major pathways: either they have zero earnings over the entire tax year (and are not married to a spouse who had earnings), or they earn too much. There are other eligibility requirements that earners must meet, including a limit on investment income. Expansions to the credit, discussed further below, changed the program cutoffs for married families (in 2005) and those with three or more children (in 2009). Table 1 lists the program parameters for Tax Year 2009.

FIGURE 1. EITC Schedule for Tax Year 2009



2.2 Previous Research

The rate of EITC participation among eligible earners has received some attention, but its calculation has presented a challenge. Most previous studies of EITC participation have focused on cross-sectional data. The first paper to estimate and report on EITC participation was Scholz (1990). This work immediately identified the challenges involved in calculating a precise participation rate, in that the calculation of the denominator—eligible earners—involved using self-reported information on earnings and family structure from the CPS ASEC and the Survey of Income and Program Participation (SIPP). Scholz's estimate is in the neighborhood of 70 percent for 1979 and 1984, early years for the program and during a time when credits were comparatively low. In later work, Scholz (1994) examined rates for 1990, finding participation in that year to be between 80 percent and 86 percent.

TABLE 1. Earned Income Tax Credit Parameters, Tax Year 2009

Family type	Phase-in rate (%)	Minimum income for maximum credit	Maximum credit	Phase-out rate (%)	Phaseout range	
					Beginning income	Ending income
No children	7.65	5,970	457	7.65	7,470	13,440
Married	7.65	5,970	457	7.65	12,470	18,440
One child	34.00	8,950	3,043	15.98	16,420	35,463
Married	34.00	8,950	3,043	15.98	21,420	40,463
Two children	40.00	12,570	5,028	21.06	16,420	40,295
Married	40.00	12,570	5,028	21.06	21,420	45,295
Three children	45.00	12,570	5,657	21.06	16,420	43,279
Married	45.00	12,570	5,657	21.06	21,420	48,279

Source: Urban-Brookings Tax Policy Center summary

Two recent calculations of EITC take-up rates are Plueger (2009), who calculated an overall take-up of 75 percent (with a confidence interval between 73 percent and 77 percent) for Tax Year 2005, and Caputo (2011), who used the National Longitudinal Survey of Youth (NLSY) to examine eligibility and take-up for Tax Years 1999 to 2005. Plueger's work is of particular interest because it used the same data source and methodology for eligibility calculation as used in this paper. Plueger estimated that 16 percent of eligible EITC claimants do not receive the credit because they do not file taxes, while 9 percent filed taxes but did not claim the credit. Caputo's estimates of take-up overall are much smaller, ranging between 53 percent and 64 percent for the NLSY79 sample. This is possibly due to the fact that the NLSY relies on asking the participant if he or she filed, while the tax data used in Plueger reflects true filings, including that done by tax preparers. Caputo (2011) also looks at predictors of take-up, finding that women, food stamp recipients, those with more children, and those separated, divorced, or widowed were all more likely to participate in the program. Caputo considers this evidence that the EITC reaches its target population relatively well.

The question addressed in this paper is the effect of the economic downturn on EITC eligibility and take-up. There is a wide-ranging literature on what happens to social-program eligibility and caseloads over the business cycle, although not much exists on the relationship of the EITC to economic downturns. Hotz et al. (2003) analyzed the income dynamics of families to look at how much "churning" existed in EITC receipt, finding that 74 percent of new EITC recipients lose eligibility within 2 years. Moreover, the main reason why families become ineligible is that their earnings increase beyond the eligibility range. There is a significant probability of families returning to eligibility, however, with approximately 35 percent of families becoming eligible 5 years after the end of an EITC spell. The results are suggestive of the sensitivity of previously low-income families to economic conditions.

This suggestion is further supported by studies of participation in other social programs, such as Aid to Families with Dependent Children (AFDC) or Temporary Assistance to Needy Families (TANF) and the Supplemental Nutrition Assistance Program (SNAP). Blank (1997) and Blank (2001) examine take-up rates for AFDC, finding a long-term increase in eligibility between 1977 and 1995. Increases in caseloads overall were related to economic conditions that increased eligibility, although changes were also induced by State-level policy changes and changes in demographics. Similarly, Grogger (2003) finds economic conditions explain much of the initial entry onto welfare rolls during the era of welfare reform; using SIPP data from 1993–1999, Grogger also finds that policy changes to welfare and the EITC also explain much of initial entry and, for the EITC, reentry into welfare. When changes in caseloads were decomposed, welfare reform explained 12 percent of the decline in welfare participation, the EITC 10 percent, and the unemployment rate 5 percent over years 1993 to 1999.

Finally, because EITC receipt is linked to the labor market, recent work on the effect of recessions on workers is relevant. Specifically, the use of the EITC during recessions is of interest, although the only work on this matter (Williams and Maag, 2008) indicates that EITC use may increase or decrease. Clearly, if a worker

loses his or her job entirely for a tax year, the lack of earnings would make him or her ineligible for the EITC. However, cases exist in which previously ineligible earners may enter eligibility: 2-earner families may become single-earner, earners may be employed part of the year, and earners may become underemployed either through fewer hours worked or lower-paid employment. A feature of the recent recession discussed by Elsby et al. (2010) is the way in which loss of total labor input (defined as the product of employment and hours per worker) is split between "bodies" (that is, number of employed individuals) and "hours" of work time. For the most recent recession, the ratio of the first to the second is approximately 70:30. Elsby et al. (2010) and Hoynes et al. (2012) each examine which demographic groups are hardest hit during recessionary periods. Each considers the latest recession the "deepest" downturn since the Great Depression, but considers the recession similar to past recessions in terms of its differential impact on certain groups. These include young, male, and minority workers as well as those with less educational attainment.

3 Data and Methods

The data and matching process used to generate the file studied in this analysis are described at length for Tax Year 2005 in Plueger (2009). The matching process changed little between 2005 and 2009, and any differences are discussed below.

The study uses data from the CPS ASEC-IRS matched file for Tax Years 2005 to 2009. IRS data sets include the universe of Form 1040 filers ("1040 data"); the subset of Form 1040 filers who received the EITC,[2] combined with a subset of filers who received a notice that they were potentially eligible for the EITC ("EITC data"); and the universe of Form W-2 earnings records ("W2 data"). Census data include the CPS ASEC, as well as data for earners who were in modeling files used previously by the IRS and Census to calculate take-up. These original modeling files reflect modeling algorithms that used only CPS data.

Records were linked using a process whereby individuals in each data set are given a unique key, called a Protected Identification Key (PIK), based on comparing name, address, and date of birth to the same variables in a master reference file. All data were then merged using this unique identifier, with other identifying information (such as name and Social Security Number) stripped. Only those observations that received the unique key are used in the analysis. Furthermore, a match is used only if CPS earnings were not imputed or allocated. Table 2 gives an account of the quality of the records match for each year. The final count is the total number of records used in the analysis for all years, It reflects the universe of CPS earners who could be matched to a unique identifier in the master file and who had modeled data from the original CPS modeling.

TABLE 2. Sample Construction

	2005		2006		2007		2008		2009	
Total CPS sample	208,562	100%	206,639	100%	206,404	100%	207,921	100%	209,802	100%
Imputed earnings	19,450	9%	20,204	10%	18,243	9%	18,926	9%	20,458	10%
Edited earnings	19,587	9%	20,490	10%	20,831	10%	19,698	9%	20,154	10%
Not PIKed	16,131	8%	15,150	1%	18,473	9%	18,547	9%	16,801	8%
In analysis sample	153,394	74%	150,795	73%	148,857	72%	150,750	73%	152,389	73%
Earners	72,447	35%	71,044	34%	71,629	35%	72,318	35%	72,603	35%
Earners with modeled data	67,289	32%	65,919	32%	66,116	32%	72,318	35%	72,603	35%
Modeled number, all years	344,245									

The number of EITC recipients in a given tax year is easy to estimate from the 1040 and EITC data; this number, however, tells us nothing about how many people were eligible for the credit. Eligibility modeling has relied on other data sources, and in this case relies on the CPS ASEC. The CPS ASEC provides important pieces of information for a tax unit that, if the same information is used when filing taxes, helps determine EITC

[2] Note that some of these EITC recipients had their EITC claim disallowed in full or in part by the IRS during the return filing process or later, and others presumably made errors related to eligibility that were not detected by the IRS. This paper does not account for the impact of noncompliance on the take-up rate.

eligibility for the unit. This includes the number of children in a household, their ages, and their relationship to the unit filer; and the unit's adjusted gross income and earnings for individuals and spouses (if married).

However, the CPS data lack certain key elements necessary for determining whether a tax unit meets the eligibility rules for the EITC. The first concerns who in a household claims a child for tax purposes. While family relationships can be established, there is no variable in the CPS that reliably assigns a child to dependency on an adult or married couple. In households where there are two related earners who file separate taxes and otherwise meet eligibility requirements (a mother and grandmother, for example), the CPS provides no information on who claims any children in the household.

The second issue concerns income and earnings. The accuracy of CPS ASEC earnings has been widely studied, with a general finding that earnings are reported with error (Bollinger (1998); Bound and Krueger (1989)). This error can be attributable to rounding or ball-parking—an error that does not vary with other variables—or to systematic under-or overreporting, an error that is negatively associated with earnings level for men.

The analysis file, therefore, was refined to update and improve EITC eligibility modeling. Part of this process involved substituting in values from 1040, W2, or EITC data when available and appropriate. These included values for earnings, adjusted gross income, and investment returns and dividends; and variables related to household structure, filing status, and claimed children. Married persons filing separately were removed from eligibility to be consistent with EITC rules. Qualifying children who were modeled as being dependent on one adult, but were claimed by another in the tax data, were reassigned to the claimant. Finally, using the matched data allowed for checking when a possible eligible was actually claimed on someone else's tax return, which would disqualify him or her from EITC participation. Table 3 lists the variables used in this analysis and their source.

TABLE 3. Sources of Variables

Variable	Source
Eligibility determination	
Wages/Earnings	W2; 1040* if W2 missing; CPS for non-filers
Adjusted Gross Income	1040; CPS for nonfilers
Dependents	1040; CPS for nonfilers
Filing type/Marital status	1040; CPS for nonfilers
Sanction status & 1040	1040
Cohabitation & CPS \\	CPS
Grandparents & CPS \\	CPS
State economic variables	
Unemployment rate	Bureau of Labor Statistics
Federal/State minimum wages	U.S. Department of Labor
Control variables	
State	CPS
Year	IRS/CPS (survey given in March of next year)
Supplemental Security Income	CPS
TANF	CPS
SNAP	CPS
Sex	CPS
Race	CPS
Hispanic origin	CPS
Education	CPS

*1040 data include 1040 files, EITC recipient files, and CP09/27 files for each year. The last is a record of 1040 filers who did not claim the EITC, but were sent a notice about their potentially being eligible for it.

In the CPS ASEC, person weights sum to the population level for the civilian, noninstitutionalized U.S. population. With the sample restrictions outlined above, the weights no longer sum to the population count. A major restriction is the removal of observations with allocated or imputed earnings data. To handle this issue, the missing data were assumed to be missing at random,[3] and weights on the remaining sample were inflated accordingly. All summary tables report the population-level estimates.

This report addresses two main questions regarding the EITC: did more people become eligible for the credit, either due to program expansion or economic forces, between 2005 and 2009; and did take-up rates change? A higher raw number of EITC eligibles might be due to an increased number of taxpayers overall or due to a higher proportion of taxpayers moving into credit eligibility. Therefore, any EITC eligibility increases over time were examined within the population that are possible 1040 filers (who are referred to here occasionally as "earners"). These include both those who were required by law to file for the tax year and those who were not. Those with positive earnings for a tax year, but who were not required to file a 1040, likely had earnings that placed them within EITC eligibility. Those who chose not to file represent a substantial portion of the population who forego the credit.[4]

Possible 1040 filers were determined using CPS ASEC data, supplemented by earnings reported on W2 forms. Based on this overall sample population of possible 1040 filers, the population of interest are those who were modeled as eligible for EITC receipt in a given tax year based on income, earnings, investment income, and number of dependents ("EITC eligibles"). The "overall participation rate" is the proportion of all eligibles who actually received a credit. "Eligible nonfilers" are those modeled as eligible for the EITC who either did not file a 1040 or filed a 1040 but did not file for the credit. Since this analysis deals only with EITC eligibles, other populations of possible interest were not examined—for example, the population who were not modeled as eligible but who did receive an EITC.

4 Results

4.1 Increases in Eligibility and Participation

In the summary tables that follow, all estimates are reported using the population weights from the CPS. Standard errors for these estimates were calculated using the CPS replicate weights. Table 4 shows increases between 2005 and 2009 in EITC eligibility and EITC participation.[5] By way of comparison, the IRS estimates that the number of required returns grew by about 1.2 percent between 2005 and 2009, and the number of required returns actually filed increased by about 2.8 percent.[6]

Meanwhile, changes in EITC eligibility and take-up outpaced the 1040 rate, with an increase of 14.1 percent in participants and a 12.7-percent increase in EITC eligibility. While the rate of 1040 participation did not increase significantly, the take-up rate for the EITC increased to about 79 percent in 2009 from about 77 percent in 2005. The highest rate of take-up occurred in 2007, with about 81 percent of those eligible filing for the EITC.

[3] That is, the methodology implicitly assumes that EITC eligibles and EITC claimants are represented among the omitted people the same way they are represented among the included people (and, hence, in the entire population).

[4] Note that both of these groups may include people who had nonwage income not reported to the IRS or to Census (and so not in our data), which may have made them ineligible for the EITC.

[5] All comparative Statements in this report have undergone statistical testing, and, unless otherwise noted, all comparisons are statistically significant at the 5-percent significance level or less.

[6] These numbers were calculated using the official estimates from the IRS.

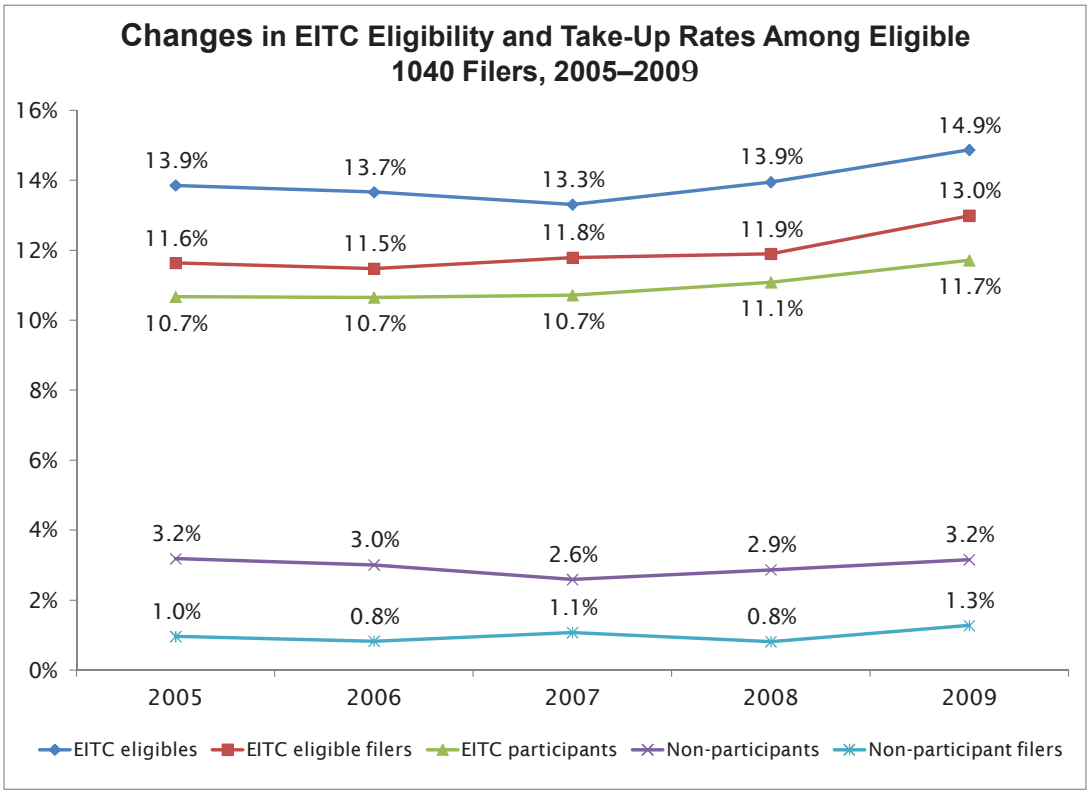
TABLE 4. Change in EITC Eligibility and Take-Up, 2005–2009

	Participants (Thousands)	Eligibles (Thousands)	Take-Up Rate (Standard Error)
2005	15,547	20,185	77.03 (0.51)
2006	15,642	20,062	77.97 (0.51)
2007	15,967	19,827	80.53 (0.52)
2008	16,678	20,992	79.45 (0.49)
2009	17,913	22,742	78.77 (0.47)
% change 05-09*	14.1	11.9	

*Computed as $(N_t - N_{t-1}) / (N_t + N_{t-1})$

Figure 2 shows eligibility, take-up, and nonparticipation rates in the EITC for those who were possible 1040 filers. The top line displays the proportion of all earners who were eligible for the EITC, which increased from 13.9 percent in 2005 to 14.9 percent in 2009. The second line shows the proportion of actual 1040 filers who were eligible for the EITC, who had a more marked increase, going from 11.6 percent in 2005 to 13 percent in 2009. Rates for EITC take-up followed a similar trend, increasing from 10.7 percent in 2005 to 11.7 percent in 2009.

FIGURE 2. Trends in EITC Eligibility and Participation, 2005—2009



Meanwhile, the proportion of all earners (possible 1040 filers) who were nonparticipants in the EITC remained steady over time, ranging between 2.6 percent and 3.2 percent. Nonparticipants who filed a 1040 also

showed little change over time, remaining at about 1 percent over the time period. These percentages, when translated into population-level estimates, correspond to between 3.9 million and 4.8 million nonparticipants each year who appeared to be eligible. Of these, 1.2 to 1.9 million filed 1040s.

4.2 Changes in Eligibility: Program Expansion or Economics?

Any analysis of changes in EITC eligibility and take-up in this time period must take into account the fact that a nontrivial proportion of earners were made eligible for the credit due to expansions to the program, which occurred in 2005, 2008, and 2009. Beginning in 2002, the start-point for the phase-out range was extended by $1,000 for those filing "Married Filing Jointly" (relative to other filers); further extensions occurred in 2005, in 2008, and again in 2009. Table 5 lists the year and the extension amount. Table 5 also lists a credit expansion that occurred for all filers with three or more children in 2009, for whom the phase-in rate increased from 40 percent (the rate for families with two or more children in previous years) to 45 percent. This phase-in change affected the maximum credit a filer could receive, from $5,028 to $5,657. Any filers who fell into the expanded program parameters were identified in the data.

TABLE 5. Changes in Program Parameters, 2005–2009

Year	Expansion Category	Expansion Amount*
2005	Married filing jointly	$2,000
2006	Married filing jointly	$2,000
2007	Married filing jointly	$2,000
2008	Married filing jointly	$3,000
2009	Married filing jointly	$5,000
2009	Three-child expansion	45% phase-in rate; $5,657 max credit

* Difference relative to other filers in the range of income eligible for the maximum credit.

Any other changes in participation rates over this period are assumed to be due to economic forces, although this term is being used in a broad sense. The recession that began in December 2007 and ended in June 2009 was associated with high rates of unemployment, which in turn might be reflected in lower end-of-year earnings, either through job loss or underemployment in terms of hours or weeks worked.

For the results that follow, the full sample of possible 1040 filers was used, in recognition that earners can move into the expanded eligibility region through program expansion or through economic forces (although these groups are certainly not mutually exclusive). Later, when I analyze eligibility using fixed-effects regressions, the subsample of earners covered by expansions is recoded as noneligible to examine changes in eligibility when program parameters are held constant except for inflation changes.

Table 6 shows the change in the rate of EITC eligibility between 2005 and 2009 among eligible 1040 filers, decomposing the percent change into two components: that which occurred due to expanded program parameters and that which occurred due to other (presumed to be economic) forces. Each rate was calculated using the person weights to arrive at a population-level estimate. Standard errors for the estimates were calculated using the CPS replicate weights. The first three columns show the probability for all earners in 2005 and 2009 and the change between the 2 years; columns 4 through 6 calculates the same statistics for only those earners who fell into the expansion category (in other words, in each case the denominator is the same: the universe of possible 1040 filers).

In general, most subgroups saw increases in eligibility over the period, with the exception of female earners and earners eligible to file "single." Black alone earners experienced a decrease, but the change was not statistically different from 0. Some increases were more marked than others. Male earners experienced a change in rate of eligibility of 17 percent, with a nearly 12-percent change attributable to economic forces. Both lower and more highly educated earners experienced greater increases. Those with a BA or BS degree or higher saw an overall percentage change of nearly 20 percent, with 14 percent attributable to other forces. Joint filers experienced the highest increase in eligibility, with a nearly 24-percent change, split almost evenly between

program expansion and other forces. Finally, earners with children experienced higher percent changes over the period than did those with no children: about 10 percent for those with one child and 16 percent for those with more than one. Other groups' changes were not statistically significant.

TABLE 6. Changes in Rates of EITC Eligibility Due to Program Expansion and Other Forces, by Demographic Characteristics, 2005–2009

	All			Expansion			% Change expansion	% Change other
	2005 Total	2009 Total	Change 05–09	2005 Total	2009 Total	Change 05–09		
Female	20.16 (0.30)	19.72 (0.29)	-0.44	0.13 (0.02)	0.39 (0.04)	0.26	1.32	-3.51
Male	9.87 (0.17)	11.78 (0.18)	1.92	0.37 (0.03)	1.03 (0.06)	0.66	6.09	11.62
White alone	12.14 (0.16)	13.47 (0.16)	1.33	0.27 (0.03)	0.78 (0.04)	0.51	3.99	6.42
Black alone	23.65 (0.58)	23.05 (0.52)	-0.60	0.31 (0.07)	0.74 (0.10)	0.43	1.86	-4.42
Other race	13.29 (0.63)	13.85 (0.62)	0.56	0.30 (0.10)	0.98 (0.20)	0.68	5.01	-0.87
Non-Hispanic	12.15 (0.16)	12.81 (0.16)	0.66	0.22 (0.02)	0.66 (0.04)	0.44	3.50	1.79
Hispanic	25.15 (0.60)	27.19 (0.49)	2.04	0.63 (0.10)	1.51 (0.12)	0.88	3.35	4.46
Less than HS	20.03 (0.50)	22.22 (0.49)	2.19	0.31 (0.06)	1.12 (0.12)	0.81	3.84	6.54
HS graduate	16.98 (0.32)	17.97 (0.30)	0.99	0.34 (0.04)	0.89 (0.06)	0.55	3.16	2.50
Some college	15.11 (0.30)	16.17 (0.30)	1.07	0.34 (0.05)	0.82 (0.07)	0.48	3.05	3.78
BA/BS or more	5.38 (0.21)	6.55 (0.20)	1.18	0.13 (0.03)	0.47 (0.05)	0.34	5.68	14.04
No children	5.70 (0.14)	5.95 (0.13)	0.25	0.09 (0.02)	0.23 (0.03)	0.14	2.42	1.89
One child	32.67 (0.54)	36.23 (0.60)	3.56	0.55 (0.07)	1.38 (0.12)	0.84	2.43	7.91
More than one child	30.33 (0.48)	35.59 (0.46)	5.26	0.79 (0.09)	2.69 (0.17)	1.90	5.75	10.20
Single filer	16.31 (0.23)	15.98 (0.22)	-0.33	{NA}	0.03 (0.01)	0.03	{NA}	-2.27
Joint filer	10.38 (0.22)	13.19 (0.21)	2.80	0.67 (0.22)	1.93 (0.09)	1.25	10.69	13.12

Each column shows the rate specified, with the percent change in the rate reported in the last two columns, first for those who became eligible due to the expansions in the program and then for all others. Standard errors for each rate (shown in parentheses) were calculated using the CPS replicate weights.

Table 7 shows a similar analysis for changes in EITC take-up contingent upon eligibility. In this case, the denominator for each column is the universe of those modeled as EITC eligibles. For take-up rates, differences between those in the expansion range and those not is not particularly relevant. Take-up is determined based on the choice of an individual to file for the credit, rather than on parameters that may have opened up eligibility. Moreover, cell sizes for certain demographic groups became too small for reporting purposes when looked at within program expansion. Therefore, Table 7 reports only the overall take-up for the different demographic groups.

Take-up rates increased for most groups, with statistically significant changes for the following: male, Black alone and other race, non-Hispanic, some college and college educated, no children, single filers, and

those in the phase-in region of the benefit. Those with less than a high school education saw a drop of 10 percent over the period, and Hispanics experienced deep decreases in participation (both statistically significant). Other groups' changes were not statistically significant.

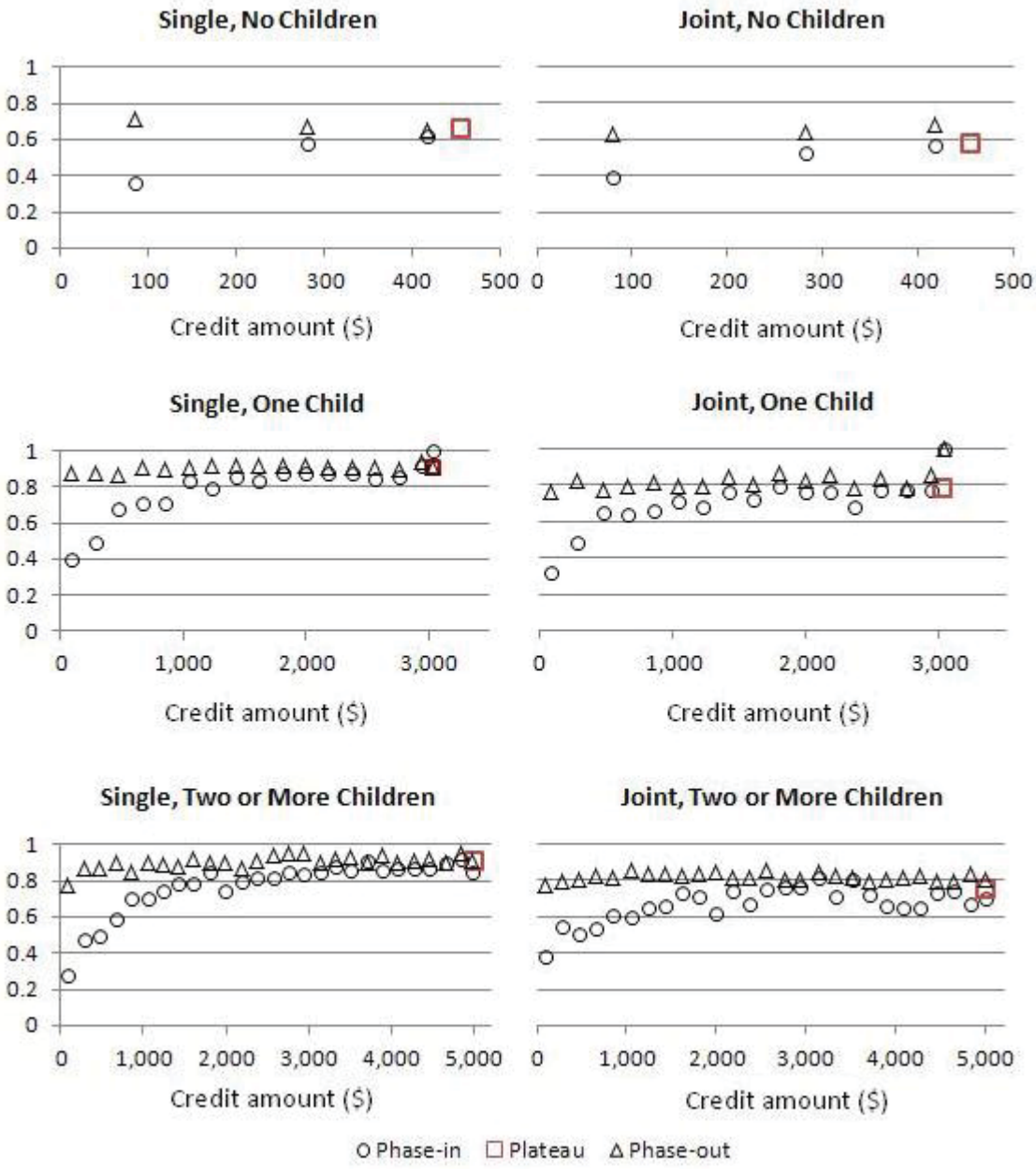
Of particular interest is the rate of take-up by credit amount. However, because of the "U" shape of the EITC benefit function, looking only at amount of credit gives an incomplete picture. Figure 3 shows the rate of take-up for the EITC based on credit amount and according to whether the earner is in the phase-in, plateau, or phase-out region. The lowest participation rate occurs where the credit is extremely low and the earner is in the phase-in region. This likely reflects earners who do not file taxes for the tax year in question. Slightly higher rates are seen for those in the plateau region and those in the phase-out for the same credit amount. Interestingly, a drop-off in participation occurs for those in the phase-out and plateau at the maximum credit amount, while those in the phase-in close to the maximum amount exhibit increasing participation. These patterns are beyond the scope of this paper, but will be examined more closely in later work.

TABLE 7. Changes in EITC Take-Up by Demographic Characteristics, 2005–2009

	2005 Total	2009 Total	Change 05–09	% Change
Female	80.81 (0.64)	81.75 (0.60)	0.94	1.16
Male	72.13 (0.83)	75.59 (0.69)	3.46	4.69
White alone	76.85 (0.63)	77.76 (0.57)	0.91	1.17
Black alone	78.29 (1.15)	81.91 (0.98)	3.63	4.53
Other race	74.17 (2.39)	81.58 (1.69)	7.42	9.52
Non-Hispanic	75.63 (0.60)	81.09 (0.48)	5.46	6.97
Hispanic	81.51 (0.93)	72.21 (1.05)	-9.30	-12.10
Less than High School	79.51 (1.11)	72.16 (1.14)	-7.35	-9.69
High School graduate	78.57 (0.80)	81.44 (0.74)	2.87	3.59
Some college	77.52 (0.92)	82.11 (0.76)	4.59	5.75
BA/BS or more	64.60 (1.99)	73.47 (1.52)	8.87	12.85
No children	56.10 (1.21)	65.23 (1.15)	9.13	15.05
One child	86.15 (0.70)	85.33 (0.67)	-0.82	-0.96
More than one child	84.33 (0.71)	82.94 (0.66)	-1.39	-1.66
Single filer	75.45 (0.69)	78.39 (0.56)	2.93	3.81
Joint filer	80.52 (0.79)	79.47 (0.75)	-1.04	-1.30
Phase-in	64.15 (1.06)	68.23 (0.95)	4.08	6.16
Plateau	83.11 (1.40)	81.31 (1.11)	-1.80	-2.19
Phase-out	83.13 (0.64)	84.39 (0.53)	1.26	1.50

Rates are based on the population-level estimate of EITC eligibles (denominator) and those paid the EITC (numerator).

FIGURE 3. Rate of Take-Up for the EITC, 2005–2009, Based on Credit Amount and According to Whether the Earner is in the Phase-In, Plateau, or Phase-Out Region†



†The charts for Two or More Children exclude rates for credit amounts larger than the 2-Child plateau ($5,028) since these credit amounts are available only to those claiming three or more children. Take-up rates looked similar when eligible earners with 3 or more children were graphed separately. All credit amounts are in 2009 dollars.

To sum up: higher rates of eligibility were seen for nearly all groups, with particularly marked increases for male earners (who were hit hard by the recession), those with more education, and those with a family structure (married and more children), which fall under more generous EITC parameters. The next section looks further into these changes.

5 Fixed-Effects Models

Assessing relative changes over time between a "treatment" and "control group" (or multiple groups) can be tackled using fixed-effects regressions. In a two-period situation, such a model reduces down to difference-in-differences. With more than one time period and repeated cross-sectional data, dummy variables are generated for each period and group, with further variation often handled using place-fixed effects.

The main question to be addressed is the impact of the recession on EITC eligibility rates. Among the major impacts of the recession of 2007–2009 was local unemployment. The overall U.S. unemployment rate increased substantially over this period, but some States and counties were affected more than others. To examine the association between the recession and increased eligibility for and participation in the EITC, I use State unemployment rates as a source of variation in economic conditions. Economic conditions by State may also be reflected in the wage distribution of workers; therefore I also include the State value for the median wage and the 20th percentile of wage by year. These controls take into account differences between States that are relatively stable over time (the overall wage distribution), as well as a time-varying economic factor (the unemployment rate) that may have increased at a greater pace for some States compared with others.

Also worth examining is how eligibility rates changed over time by demographic group. To address this question, I add to models of eligibility dummy variables for each characteristic of interest multiplied by a linear time term. This captures the year-by-year change for the group in question over the time period, using 2005 as the base year. As an alternative, I could have pooled pre- and post-recession years and used a difference-in-difference model. Doing so leads to similar results once the time-trend coefficients are summed over the 4 years. However, in terms of unemployment, deciding on a pre- and post-period is problematic. While the recession officially began in December 2007, unemployment lagged slow economic growth. For example, the national monthly unemployment rate at the end of 2007 was 5 percent, and it increased gradually over 2008 to end at 7.3 percent. The highest rates of unemployment—greater than 9 percent—occurred in 2009.

The baseline model, regressing State economic indicators on eligibility in a fixed-effect framework, is

$$y_{ist} = \alpha + \beta z_{st} + \gamma x_{ist} + \sigma_s + \tau_t + \varepsilon_{ist}$$

where y takes on a 1 if an individual in State s and in year t is eligible for the EITC and 0 otherwise.[7] This model is run first parsimoniously—with just the economic indicators—and then with the full set of individual characteristics, group dummies, and group-specific trends. Individual characteristics include age and age squared, and binary terms indicating race (White alone, Black alone, Other race[8]); Hispanic origin; education (four categories); filing type (Joint or Single); and number of dependent children (none, one, or more than one).

Adding individual characteristics tells us only the contribution of each characteristic to the "after" rate, not how rates change for a given characteristic over time. Therefore, I run a fixed-effects model using characteristic-specific linear time terms, which captures the year-by-year change in eligibility for each characteristic. The equation for this model can be expressed as

$$y_{ist} = \alpha + \beta z_{st} + \gamma x_{ist} + T_\tau + (X_{ist} \times T_\mu) + \sigma_s + \varepsilon_{ist}$$

where T is a linear time term and $X_{ist} \times T_\mu$ refers to binary characteristics multiplied by the linear term for time. The interaction between time and the characteristic in question can be interpreted as the additional change in eligibility experienced by this group year-by-year, holding other characteristics constant. Using a linear probability model rather than a probit or logit allows for a straightforward interpretation of the interaction (Ai and Norton, 2003).[9]

Take-up was examined using a similar model, with inclusion of individual-level measures of other program use, including the log values of Supplemental Security Income (SSI), Temporary Assistance to Needy Families (TANF), and Supplemental Nutrition Assistance Program (SNAP) benefits. Moreover, since the ben-

[7] This work focuses on individual characteristics, so each CPS sample would have to be viewed as a year panel. This is why the standard errors are clustered on the State level.

[8] Included in Other race are American Indian/Alaska Native alone, Asian alone, and Other.

[9] Other models were examined, with similar results: A logit yielded marginal effects very similar to the coefficients reported in the linear model. Models were also run using each characteristic in turn as the "difference-in-differences" estimator. The coefficients yielded by this method did not differ significantly from those in the joint specification.

efit level of the EITC may affect take-up, I included dummy variables for being in the phase-in or phase-out region (with eligibility for the maximum benefit forming the comparison group).

Each model was weighted using the CPS ASEC population weights, corrected as described earlier. For each model, standard errors were corrected for both the "dimension" problem (the use of State-level variables with individual level units of observation) and autocorrelation by clustering the standard errors at the State level (Bertrand et al., 2000). Models were also run with the standard errors corrected using the CPS ASEC replicate weights. Standard errors were smaller using the weights, so the more conservative results using clustered standard errors are reported. No coefficient moved from significant to nonsignificant (or vice versa) at the 5-percent level between models.

5.1 Predicting Eligibility

Table 8 displays the results for eligibility. The first three models are results when all those estimated to be eligible for the EITC are coded "1" for the dependent variable. The second three models are results when those estimated to have become eligible due to program expansion are recoded to "0," thus holding constant the EITC program parameters. The overall eligibility rate for the first three columns is 13.3 percent, and for the second three columns, 13.0 percent. For the full population of eligibles, unemployment rate is not a predictor of eligibility unless the linear trend and the interaction effects are included in the model. For the more restricted population, unemployment is a predictor of eligibility when characteristics are included. Depending on which population one considers and the model in question, a 1-percent increase in the unemployment rate predicts a change in probability of approximately 0.2 percent. Median wage influences EITC eligibility in a way one would expect, with a higher median associated with a lower probability of eligibility.

The interaction effects appear in columns 3 and 6. In each case, the constant term reflects the rate in 2005 for single, White alone, female, non-Hispanic earners with a high school degree and no children. The coefficient on the time trend indicates the change per year, averaged over the time period, that this base group experienced. For male earners, the average year-by-year growth in EITC eligibility is 0.4 percent whether the full or more restricted population is used. Those with less than a high school education experienced a year-by-year change in eligibility of -0.5 percent regardless of which sample was used. Those with one child saw a growth rate of about 0.7 percent using either population, while those with more than one child experienced a nearly 1 percent year-by-year increase for the full population. The latter effect drops by nearly a third once the population is restricted to nonexpanders. Similarly, joint filers experienced a 0.6-percent year-by-year change in the full population, but only a 0.4-percent change in the restricted population. These results are not surprising given that the target populations for the expansions were married filers and those with three or more children. It should be noted that in all cases for which a coefficient term was statistically significant for the full population, the same coefficient derived when holding the 2005 parameters constant was statistically different from the first. However, the inclusion of expanders changes only the magnitude of the coefficients, not the direction.

It is interesting to examine the extent to which families with children differed in terms of their status. A triple interaction term indicating any children times filing jointly times time was added to the full regression. The coefficient on the interaction term indicates that joint filers with children experienced increasing rates of eligibility compared with single filers with children (marginally significant at $p < 0.07$). However, the coefficient is not different from zero when those who were in the expansion category were recoded to ineligible. Thus, the increasing eligibility for those with children seen in the main results appears to be largely driven by married earners, although much of this occurred due to program expansion.

Men and low-education earners of both sexes were at higher risk for unemployment during the Great Recession, but while male earners experienced increasing eligibility over time compared to the base group, those with less than a high school education experienced decreasing rates. When looking within the population of only those with less than a high school degree, men were increasingly more likely to report not working at all for the entire year compared with women with otherwise the same characteristics. The results lend support to the "underemployment" hypothesis—that families entered eligibility due to the retention of one spouse's earnings, thus leading to an increase overall in the eligible population. The increasing eligibility of men also provides some evidence for the underemployment hypothesis; the interaction term expresses the year-by-year change in eligibility experienced by single, white, male workers with a high school degree and no children. This population reported decreasing weeks of work each year over the same time period (-0.57 weeks each year,

$p < 0.001$). The analysis provides suggestive evidence regarding those who received benefit from the EITC, and those whose labor-market experience was negative enough to leave them out of eligibility. Earners were protected by marriage, and working any amount during the course of a tax year ensured that male earners benefited from eligibility. Those with the least amount of education experienced a drop-off in eligibility due to a complete lack of earnings and of weeks worked over entire tax years. These dynamics between employment and eligibility will be further examined in future work.

TABLE 8. Linear Probability Models: Dependent Variable Is Eligibility for the EITC.

	Model 1 β, (SE)	Model 2 β, (SE)	Model 3 β, (SE)	Model 4 β, (SE)	Model 5 β, (SE)	Model 6 β, (SE)
Unemployment rate	0.04 (0.14)	-0.18 (0.12)	0.22*** (0.07)	0.10 (0.11)	0.24* (0.09)	0.15* (0.07)
Minimum wage (log)	1.59 (1.23)	0.66 (1.13)	0.44* (1.18)	1.11 (1.17)	0.20 (1.07)	0.19 (1.14)
Median wage (÷100)	-0.02* (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.01* (0.01)
20th percentile wage (÷100)	0.00 (0.01)	0.00 (0.01)	0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.01 (0.01)
Time			-0.57*** -0.14			-0.46** -0.13
Interactions of Variable X Time						
Male			0.36** (0.12)			0.37** (0.11)
Black alone			0.00 (0.11)			0.00 (0.11)
Other race			-0.32 (0.23)			-0.34 (0.23)
Hispanic			0.09 (0.19)			0.02 (0.19)
Less than HS			-0.50** (0.14)			-0.54*** (0.14)
Some college			-0.20 (0.11)			-0.18 (0.10)
BS/BA or more			-0.01 (0.10)			0.06 (0.09)
Joint filer			0.58*** (0.14)			0.35** (0.12)
One child			0.77*** (0.16)			0.68*** (0.16)
More than one child			0.99*** (0.15)			0.70*** (0.13)
Constant	14.65*** (2.72)	18.62*** (2.46)	18.88*** (3.06)	15.15*** (2.52)	18.90*** (2.31)	18.81*** (3.03)
X variables	No	Yes	Yes	No	Yes	Yes
Year	Yes	Yes	No	Yes	Yes	No
State	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.01	0.21	0.21	0.01	0.21	0.21
F test			22.17			17.44
Observations	344,245					

***p ≤ .001, **p ≤ .01, *p ≤ .05. Population comprises those who were eligible for the EITC from 2005 to 2009. Take-up is defined as ``1'' for those who filed for and received the EITC and ``0'' for those who did not. Coefficients and standards errors are multiplied by 100. Standard errors clustered at the State level in parentheses. F-tests report the joint significance of the interaction terms.

5.2 Predicting Participation

Table 9 shows results when participation is examined. The sample is restricted to those who were modeled as eligible, thus the results can be interpreted as rates of change in take-up contingent upon eligibility. The sample

is therefore smaller in models 4 through 6, since I have removed from eligibility those who became eligible due to program expansion.

TABLE 9. Linear Probability Models: Dependent Variable Is Participation in the EITC.

	Model 1 β, (SE)	Model 2 β, (SE)	Model 3 β, (SE)	Model 4 β, (SE)	Model 5 β, (SE)	Model 6 β, (SE)
Unemployment rate	-0.49 (0.59)	-0.20 (0.61)	-0.50* (0.20)	-0.50 (0.56)	-0.21 (0.58)	-0.54** (0.19)
Minimum wage (log)	3.33 (4.19)	3.39 (4.18)	6.90* (3.31)	1.75 (4.38)	1.93 (4.38)	5.47 (3.41)
Median wage (100)	-0.02 (0.02)	-0.01 (0.02)	0.01 (0.02)	-0.02 (0.02)	-0.01 (0.02)	0.01 (0.02)
20th percentile wage (100)	0.05 (0.04)	0.02 (0.04)	0.03 (0.02)	0.04 (0.04)	0.02 (0.03)	0.03 (0.02)
Unemployment comp (log)	0.77*** (0.08)	0.69*** (0.10)	0.71*** (0.11)	0.78*** (0.08)	0.70*** (0.11)	0.71*** (0.12)
Supplemental Security Income (log)	-1.77*** (0.13)	-1.10*** (0.13)	-0.70*** (0.13)	-1.79*** (0.13)	-1.12*** (0.13)	-0.72*** (0.13)
TANF (log)	-0.35 (0.20)	-0.82*** (0.19)	-0.27 (0.18)	-0.38 (0.20)	-0.85*** (0.19)	-0.29 (0.19)
SNAP (log)	0.55*** (0.10)	0.05 (0.09)	0.62*** (0.10)	0.57*** (0.10)	0.05 (0.10)	0.62*** (0.10)
Time			2.28*** (0.64)			2.37** (0.68)
Interactions of Variable X Time						
Male			0.49 (0.43)			0.48 (0.43)
Black alone			-0.30 (0.39)			-0.35 (0.40)
Other race			0.58 (0.57)			0.49 (0.57)
Hispanic			-2.62*** (0.38)			-2.55 (0.39)
Less than HS			-1.56*** (0.36)			-1.64 (0.36)
Some college			0.56 (0.36)			0.53 (0.37)
BS/BA or more			0.56 (0.64)			0.66 (0.60)
Joint filer			-0.37 (0.42)			-0.35 (0.41)
One child			-1.67*** (0.44)			-1.62*** (0.43)
More than one child			-1.69** (0.49)			-1.67** (0.49)
Phase-in			0.72* (0.36)			0.76* (0.35)
Phase-out			0.4 (0.27)			0.42 (0.25)
Constant	73.54*** (10.43)	63.49*** (12.38)	63.77*** (8.91)	77.11*** (10.91)	65.82*** (12.92)	66.44*** (9.11)
X variables	No	Yes	Yes	No	Yes	Yes
Year	Yes	Yes	No	Yes	Yes	No
State	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.01	0.09	0.13	0.01	0.09	0.13
F test			18.9			18.45
Obs.	48,148			46,661		

***p ≤ .001, **p ≤ .01, *p ≤ .05. Population comprises those who were eligible for the EITC from 2005 to 2009. Take-up is defined as ``1'' for those who filed for and received the EITC and ``0'' for those who did not. Coefficients and standards errors are multiplied by 100. Standard errors clustered at the State level in parentheses. F-tests report the joint significance of the interaction terms.

State unemployment rates negatively affected participation rates, with a 1-percent increase in the rate associated with a 0.5-percent drop in the probability of take-up in both the full and restricted population of eligibles. While this is not a surprising finding, it is unclear what mechanism is at work to lead to this result. The receipt of unemployment insurance has a positive effect on take-up, with a 10-percent increase in benefit equating to a 0.7-percent increase in take-up. Food stamp receipt has a similar effect on participation. Meanwhile, other program participation had a negative effect on take-up, although these results are significant across the board only for SSI. Greater TANF receipt is associated with negative take-up only in models 2 and 5, which do not include the full set of time interactions.

These findings are somewhat consistent with other analyses of EITC take-up. For example, Caputo (2006) found that any food stamp receipt tripled the odds of filing for the EITC. He did not, however, find significant effects for SSI or TANF. Caputo hypothesized that, because food stamps and SSI have higher income eligibility thresholds than TANF, the latter two programs were more likely to have an influence on take-up (since higher income people are, in general, more likely to participate). However, the findings here indicate that SSI receipt is negatively correlated with EITC take-up. Because I include the value of the variable rather than a simple indicator, the case may be that a higher level of SSI—rather than any participation—discourages EITC take-up. The log minimum wage is also a predictor of EITC take-up, but only in the full sample and with the full set of interaction terms. This is likely due to the fact that States with a State-level EITC, which may induce higher take-up, tend to also have higher minimum wages.

For demographic predictors, increasing eligibility did not necessarily translate into increasing take-up. Those with children experienced decreasing take-up rates compared with those without children, both in the full and restricted models. Men and joint filers, however, did not see a year-by-year change in take-up that was statistically significant. Those with less education experienced not only a year-by-year decline in eligibility, but a year-by-year decline in take-up, as well. The results may reflect new EITC eligibles not yet correctly negotiating the program, an idea that will be taken up in future work.

Finally, those in the phase-in range of the EITC experienced year-by-year participation increases compared with those at the plateau. Historically, take-up in this region of the credit has been low, which is of concern for policy-makers as this is the lowest-earning group of eligibles. Many in this group do not file a 1040, and may not know that the EITC is available to them.

6 Conclusion

The work presented in this paper was intended to provide descriptive information on the changes experienced in eligibility for the EITC and its take-up over the Great Recession. The objective of the work is twofold: to report on general estimates of eligibility and take-up over the years in question and to break down changes in eligibility and take-up by demographic groups that experience labor-market downturns differentially. The paper contributes to the literature by: providing information on eligibility and take-up using administrative records linked to survey data, which improves the accuracy of estimates; and analyzing changes in eligibility over a large-magnitude change in the health of the economy.

Findings indicate that eligibility for the credit increased overall during the recession, and most demographic groups experienced increases in eligibility for the EITC over the time period when looked at individually. Take-up contingent upon eligibility, however, remained constant. Two groups—men and the low-skilled—are of particular interest in the analysis, since they experienced disproportionately negative labor market outcomes. While men overall experienced increases in eligibility, those with low education experienced decreases when other characteristics are held constant. This finding gives some suggestive evidence that low skill simultaneously predicts particularly poor labor market outcomes and EITC eligibility. Because those with less education are also less likely to be married, it is probable that the combination of total loss of earnings and zero spouse earnings conspire to lower eligibility for this group.

Work remains to be done on the dynamics of employment versus eligibility. The descriptive information presented here indicates that groups that were affected more strongly by the economic downturn (male and low-skilled workers) experienced either increasing or decreasing rates of eligibility over the time frame compared with a base group. Since an individual may become ineligible either by having zero earnings over a tax

year or by having too much earnings, aggregate eligibility may change for certain family and skill groups based on full unemployment or underemployment. The evidence presented here gives an indication that there may be groups within the target population for the EITC that do not benefit in an economic downturn because of the program's tie to work. Further study on the EITC during recessions is necessary to understand the full impact of the credit.

References

Ai, C., Norton, E., 2003. Interaction terms in logit and probit models. Economics letters 80 (1), 123–129.

Bertrand, M., Duflo, E., Mullainathan, S., 2000. How much should we trust difference in-differences estimates? The Quarterly Journal of Economics 119 (1), 249–275.

Blank, R., 1997. What causes public assistance caseloads to grow? Tech. rep., National Bureau of Economic Research.

Blank, R., 2001. What causes public assistance caseloads to grow? Journal of Human Resources 36 (1), 85–118.

Bollinger, C. R., 1998. Measurement error in the current population survey: A non-parametric look. Journal of Labor Economics 16 (3), 576–594.

Bound, J., Krueger, A. B., 1989. The extent of measurement error in longitudinal earnings data: do two wrongs make a right?

Caputo, R., 2011. TANF and EITC use: A study. U.S. Social Welfare Reform, 105–141.

Elsby, M., Hobijn, B., Sahin, A., 2010. The labor market in the great recession. Tech. rep., National Bureau of Economic Research.

Grogger, J., 2003. Welfare transitions in the 1990s: The economy, welfare policy, and the EITC.

Hotz, V., Mullin, C., Scholz, J., 2003. Trends in EITC take-up and receipt for California's welfare population, 1992-1999. Los Angeles: University of California, Los Angeles.

Hoynes, H., Miller, D., Schaller, J., 2012. Who suffers during recessions? Tech. rep., National Bureau of Economic Research.

Plueger, D., 2009. Earned Income Tax Credit participation rate for Tax Year 2005. IRS Research Bulletin.

Scholz, J.K., 1990. The participation rate of the Earned Income Tax Credit. University of Wisconsin-Madison, Institute for Research on Poverty.

Scholz, J.K., 1994. The Earned Income Tax Credit: Participation, compliance, and antipoverty effectiveness. National Tax Journal 47, 63–63.

Williams, R., Maag, E., 2008. The recession and the Earned Income Tax Credit. Tech. rep., The Urban Institute. 23.

2

Business Compliance Behavior

Beers ◆ LoPresti ◆ San Juan

Klassen ◆ Lisowsky ◆ Mescall

Bickers ◆ Lloyd ◆ Nair ◆ Slyuzberg

Factors Influencing Voluntary Compliance by Sole Proprietors: Preliminary Survey Results

Tom Beers, Eric LoPresti, and Eric San Juan,
IRS Taxpayer Advocate Service, Internal Revenue Service

A principal goal of the IRS is to maximize the extent to which taxpayers pay their taxes voluntarily and timely. To do so, the IRS needs to understand why they comply (or don't comply). The National Taxpayer Advocate 2007 Annual Report to Congress, which included a review of existing research (the "2007 Review") and the National Taxpayer Advocate 2010 Annual Report to Congress, which included a proposal for this research (the "2010 Proposal") identified numerous types of noncompliance, as shown below.[1]

TABLE 1. Typology of Noncompliance[2]

Type	Description
Procedural	Failed to follow complicated procedural rules, such as quarterly filing requirements
Lazy	Failed to follow burdensome procedural rules, such as recordkeeping requirements
Unknowing	Misunderstood the legal rules
Asocial	Motivated by economic gain
Brokered	Acted on the advice of a professional
Symbolic	Perceived the law or the IRS as unfair
Social	Acted in accordance with social norms and peer behavior
Habitual	Knowingly repeated previous noncompliance

The 2007 Review and 2010 Proposal also identified various factors driving taxpayer compliance decisions. The Taxpayer Advocate Service (TAS) conducted a study to investigate whether and how these factors affect voluntary compliance by sole proprietors (*i.e.*, those who file Form 1040, Schedule C, *Profit or Loss from Business*), as described in the 2010 Proposal. The factors are reflected in the following table:

[1] See National Taxpayer Advocate 2007 Annual Report to Congress vol. 2, 138-50 (Marjorie E. Kornhauser, *Normative and Cognitive Aspects of Tax Compliance*) [hereinafter "2007 Review"]; National Taxpayer Advocate 2010 Annual Report to Congress vol. 2, 71–88 (*Researching the Causes of Noncompliance: An Overview of Upcoming Studies*) [hereinafter "2010 Proposal"]. Because the 2007 Review and the 2010 Proposal cite much of the literature discussing each of the relevant factors, this paper does not revisit that underlying literature or theoretical basis for the factors previously identified.

[2] See 2010 Proposal at 81 (Table 2.4.1, *Typology of Noncompliance and Potentially Operative Factor(s) Identified by the Literature*) (citing Robert Kidder and Craig McEwen, *Taxpaying Behavior in Social Context: A Tentative Typology of Tax Compliance and Noncompliance*, 2 Taxpayer Compliance 57, 56–62 (1989) and Leslie Book, *The Poor and Tax Compliance: One Size Does Not Fit All*, 5 Kans. L. Rev. 1, 23–33 (2003)).

TABLE 2. Factors Identified as Potentially Driving Compliance Behavior[3]

Factor	Description
Deterrence	"People comply when the potential sanction multiplied by the perceived likelihood of getting caught outweighs the economic gain from cheating." [However,] "the deterrence model is incomplete because it seems economically irrational for so many taxpayers to comply given the low probability of getting caught cheating."
Norms	"According to social norms and reciprocity theories, taxpayers who believe most other taxpayers comply are more likely to reciprocate by complying."
Tax morale	Taxpayers "who cheat may feel guilty when they break the norm if it has been adopted as the taxpayer's own tax morale." In addition, "those who trust the government and feel the tax laws and procedures are fair and fairly enforced may be more likely to feel a moral obligation to comply, even if the outcome of those procedures is unfavorable."
Trust	Taxpayers "may use unfair rules or procedures, unreasonable penalties, bad experiences with the IRS, or a lack of faith in government or the IRS to justify either reducing efforts to comply or active noncompliance."
Complexity and convenience	"Taxpayers who face complicated rules may be unable to comply, or may use complexity as a reason to justify noncompliance."
Preparers and other third parties	"Tax preparers may have a significant effect on tax compliance."

TAS also asked questions about demographics and affiliations. While these items may not directly affect compliance decisions, TAS included them because they may be correlated with or help explain the factors that do. For example, information about a person's affiliations could help TAS draw conclusions about the person's norms, tax morale, and related factors.[4]

TAS focused on sole proprietors because underreporting by sole proprietors represents the largest portion of the tax gap (*i.e.*, taxes not voluntarily and timely paid).[5] The IRS is unlikely to be able to detect or deter noncompliance by this segment through enforcement alone without expending significant resources because most sole proprietor income is not subject to third-party information reporting. Relatively inexpensive measures, such as document matching and correspondence examinations, cannot reliably detect such income. Thus, it is particularly important for the IRS to gain a better understanding of how to improve compliance among sole proprietors using levers other than economic deterrence.

TAS contracted with a consulting firm, Russell Research, to help conduct a telephone-based survey of two groups: a nationally representative sample of sole proprietors (the "National Survey"); and sole proprietors located in high- and low-compliance communities (the "Community Survey"). The discussion below describes the methodology and key preliminary results of both surveys.

Discussion

Methodology

TAS sorted taxpayers with DIF scores in the highest or lowest deciles into low- or high-compliance groups.

TAS relied on internal IRS statistical estimates of the likelihood that an audit would produce a significant adjustment (called a Discriminant Function or "DIF" score) as a proxy for a person's tax compliance, as described in the 2010 Proposal.[6] The IRS develops DIF scores for individual taxpayers in each "examination activity

[3] The factors and their descriptions come from the 2010 Proposal (pages 76–81), which synthesized them from tax compliance literature, including the 2007 Review.

[4] 2010 Proposal at 87.

[5] IR-2012-4, *IRS Releases New Tax Gap Estimates; Compliance Rates Remain Statistically Unchanged From Previous Study* (Jan. 6, 2012), available at http://www.irs.gov/uac/IRS-Releases-New-Tax-Gap-Estimates;-Compliance-Rates-Remain-Statistically-Unchanged-From-Previous-Study.

[6] The DIF is a computer algorithm that estimates the likelihood that an audit of a particular return would produce an adjustment. The DIF is based on data obtained and periodically updated from IRS National Research Program examinations. See 2010 Proposal at 86 n. 49 (and sources cited therein).

code" or EAC. For sole proprietors, EACs are defined by the taxpayer's total gross receipts (TGR) on Schedules C and F and total positive income (TPI), which is positive income from all sources before adjusting for deductions, exemptions, or negative income (*e.g.*, negative income from post-holiday returns). TAS excluded the EACs for low-income taxpayers claiming the earned income tax credit (EITC) because they may present a unique set of tax compliance issues.[7] For Tax Year 2009, six EACs included all sole proprietors residing in the United States who did not claim the EITC, as shown in the following table.

TABLE 3. Total Gross Receipts (TGR) and Total Positive Income (TPI) Limits for Certain Schedule C Examination Activity Codes (EACs)[8]

EAC	TGR	TPI
274	<$25,000	<$200,000
275	$25,000 - $99,999	<$200,000
276	$100,000 - $199,999	<$200,000
277	>$199,999	<$200,000
280	-	$200,000 - $999,999
281	-	>$999,999

For each of the six EACs,[9] TAS sorted all Tax Year 2009 returns by DIF score, then divided each group into ten deciles. Taxpayers with returns in the first and second deciles have the lowest DIF scores and are assumed to be the most compliant. Those in the ninth and tenth deciles have the highest DIF scores and are assumed to be the least compliant.[10] All other deciles are considered moderately compliant.

To minimize selection bias and the number of surveys required, TAS selected a random sample of taxpayers in high- and low-compliance groups in each EAC for the National Survey. However, TAS combined the four EACs with the fewest taxpayers into two groups (or "strata") with two EACs in each of these two strata. TAS also selected one group of taxpayers with medium levels of compliance from all EACs for comparison purposes. TAS received 3,306 responses to the National Survey, as shown on the following table.

TABLE 4. National Survey Responses by Strata and Population

National Sample Strata	Population	Responses
EAC 274 DIF Deciles 1 – 2	2,053,331	350
EAC 274 DIF Deciles 9 – 10	2,053,331	350
EAC 275 DIF Deciles 1 – 2	571,075	351
EAC 275 DIF Deciles 9 – 10	571,075	384
EACs 276, 277 DIF Deciles 1 – 2	268,565	359
EACs 276, 277 DIF Deciles 9 – 10	268,565	350
EACs 280, 281 DIF Deciles 1 – 2	256,306	383
EACs 280, 281 DIF Deciles 9 – 10	256,306	379
All EACs DIF Deciles 3 – 8	9,447,830	400
Total	15,745,384	3,306

[7] Because it is sometimes difficult to distinguish between a hobby and a real business, TAS considered, but ultimately rejected, the idea of excluding those in the lowest income ranges. Thus, the survey may have captured the views of some taxpayers who were conducting a hobby. However, the number of respondents likely to fall into that category was limited because TAS stratified the sample by EAC, as described below.

[8] IRS, Document 6209, *IRS Processing Codes and Information* 12–16 (Jan. 2012). Many parts of Document 6209 are designated as "official use only," but these EAC definitions are not.

[9] Each EAC was treated separately because the formulas that produce the DIF scores are unique to one EAC. So, for example, a specific DIF score in one EAC doesn't have the same meaning in a different EAC.

[10] As noted below, this assumption is a significant limitation of the study. TAS relied on DIF scores because taxpayers—particularly noncompliant taxpayers—might not respond accurately to questions about the extent of their own tax compliance. Although DIF scores are not perfect predictors of noncompliance, those in the top two deciles are likely to be less compliant than those in the bottom two deciles. Nonetheless, using this approach to classify people or geographic areas as "high-compliance" or "low-compliance" (as we do in this paper) must be viewed as very approximate—not definitive.

The national sample was large enough that we can be at least 95 percent confident that the results reflect the views of the universe of taxpayers in each stratum with a margin of error of 5 percent or less.

TAS identified communities with median DIF scores in the highest and lowest deciles as low- or high-compliance communities, but found few high-compliance communities.

TAS originally intended the Community Survey sample to have 2 strata with 350 respondents each, from high-compliance communities and low-compliance communities, cities, towns, and other geographic areas across the country identified by addresses with ZIP codes reported by the taxpayers on their returns. To be considered "high compliance," a community's residents must have a median DIF score in the bottom 30 percent (*i.e.*, the bottom three deciles). To be considered "low compliance" a community's residents must have a median DIF score in the top 30 percent (*i.e.*, the top three deciles). However, the location of taxpayers with high levels of compliance (or at least low DIF scores) was such that TAS could not identify enough high-compliance communities to generate 350 respondents. Simply put, there were few high-compliance communities. In particular, the criteria above yielded three U.S. geographic communities.[11] The distribution of high- and low-compliance taxpayers and communities was a significant discovery. In the end, the Community Survey had 535 respondents—362 from low-compliance communities and 173 from high-compliance communities.[12]

TAS developed survey questions and contracted with Russell Research to administer the survey.

TAS developed telephone-based survey questions to investigate the factors suggested by the tax compliance literature, as described above. TAS used the same questions for both the National and Community Surveys.

TAS contracted with Russell Research to refine the survey questions, administer the surveys, and compile summary statistics.[13] Russell Research conducted all interviews by telephone from January 3, 2012, to April 19, 2012. It contacted potential respondents up to four times. The response rate was 56 percent for the national sample and 54 percent for the community sample. This better-than-average response rate should help to minimize the likelihood that the survey results were affected by selection bias—the possibility that the views of nonrespondents are significantly different from the views of respondents.[14]

Important Assumptions and Limitations

TAS used DIF scores as a proxy for compliance by those in the top and bottom DIF deciles.

As discussed in the 2010 Proposal, it is difficult to measure actual compliance with perfect accuracy. Taxpayers are not likely to confess any noncompliance in response to a survey, and even detailed audits conducted by the IRS's National Research Program (NRP) are likely to contain errors. Even assuming that NRP audit results, as adjusted by IRS researchers, reflect actual compliance, the audit itself has an effect on the taxpayer's attitude about the tax system, potentially biasing the taxpayer's response to any subsequent survey. Thus, TAS decided not to survey taxpayers who had been subject to an NRP audit. While surveying taxpayers immediately before they were subject to an NRP audit might have been more productive, TAS deemed it overly deceptive. Thus, TAS opted to rely on DIF scores as an imperfect, but acceptable, measure of actual compliance, at least for those in the top and bottom DIF deciles.[15]

---

[11] In addition, the high-compliance criteria identified a military and a Native American community. This identification allows for future study, potentially observing mechanisms of authority and cohesion in those communities. For this phase of the study, however, the Army Post Office did not identify a geographic community as did other addresses, and the Native American community, with a quasi-sovereign history, had a fiduciary relation to the federal government (which was the subject of some survey questions).

[12] The DIF score for a particular survey respondent, however, may not correspond to the DIF score of the community. For example, the response of a taxpayer with a DIF score suggesting a high level of noncompliance could have been selected as a representative of a high-compliance community. The Community Survey was geographically limited. The goal was to look at the compliance norm in the geographic community and how institutions in that area might be influencing that norm; the sample was to be representative of the locality rather than another pool.

[13] The actual survey questions, along with a topline analysis by Russell Research, are reproduced in Appendices I and II, respectively, of National Taxpayer Advocate 2012 Annual Report to Congress vol. 2, 1–70 (*Factors Influencing Voluntary Compliance by Small Businesses: Preliminary Survey Results*), available at http://www.taxpayeradvocate.irs.gov/2012-Annual-Report/FY-2012-Annual-Report-To-Congress-Full-Report (last visited Aug. 8, 2013).

[14] See, e.g., Scott Keeter, et al., *Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey*, 70 Pub. Op. Quart. 759–79 (2006). It may also suggest that taxpayers were somewhat more interested in discussing their views about taxes than other subjects.

[15] Although some taxpayers in our sample had been subject to IRS examination or collection activity, we did not exclude them or place them into the noncompliant group, as the activity could have had an effect on their subsequent compliance behavior. Of course, any direct contact with the IRS could affect their views about the IRS and the survey results.

As noted above, this study assumes that taxpayers with high DIF scores (*i.e.*, in the ninth or tenth decile) are noncompliant and that those with low DIF scores (*i.e.*, in the first and second decile) are compliant. To the extent this assumption is incorrect, the survey results could be misleading. However, by excluding those in the middle deciles from the high- and low-compliance groups, TAS sought to improve the likelihood that taxpayers in those groups did, in fact, have high or low levels of reporting compliance.[16]

The Community and National Surveys are different in kind.

While the National Survey is statistically representative of the United States, the Community Survey of select communities relates to a different kind of data. The National Survey may reflect a response of the general population, but given the necessarily varied circumstances across the country, may not relate that response to any particular set of local conditions that could provide a deeper context. On the other hand, the Community Survey facilitates analysis of responses in relation to a more specific set of social circumstances. Thus, the Community and National Surveys are complementary.

TAS did not design the Community Survey sample for projection to any larger group of taxpayers. As set forth above, noncompliance (and by extension, compliance) may vary according to a typology. Focusing on a community permits identification of types, if any, that may not be nationally prevalent. The purpose of the Community Survey would not be to project an ideal type on other parts of the country, but rather to identify relevant factors or characteristics, such as trust in government or traditions of authority, that may occur outside the community context as well.[17] Similarly, in American studies, Middletown stands as a landmark (eponymously popularized by Public Broadcasting System (PBS) television) not because research on Muncie, Indiana (for which it was a pseudonym) was representative of the U.S. but because of the depth in which investigation of one community contextualized national trends.[18] In short, the Community Survey may be prototypical rather than typical.

TAS randomly selected taxpayers for the Community Survey from high- and low-compliance communities without excluding taxpayers with unrepresentative DIF scores. Accordingly, the high-compliance community sample includes taxpayers who are noncompliant and who would be included in the low-compliance group for purposes of the National Survey. Conversely, the low-compliance community sample includes taxpayers who are compliant and who would be included in the high-compliance group for purposes of the National Survey. Thus, responses from high-compliance communities may not be similar to responses from the high-compliance group, and responses from the low-compliance communities may not be similar to responses from the low-compliance group.

Key Findings of the National Survey[19]

The National Survey results are statistically representative of the views of Schedule C filers in the high- and low-compliance groups. As discussed in the 2007 Review and 2010 Proposal, a large body of research discusses the potential effect of various factors on tax compliance, but this study is the first to link survey responses to IRS estimates of the respondent's actual tax compliance. Thus, the National Survey results provide an unprecedented look at the differences between the views of the Schedule C filers that are the most and least compliant, at least according to IRS estimates.

---

[16] Per TAS research on 2006 NRP data, assessments for taxpayers in the top 20 percent of DIF scores were significantly higher on average than those for the lowest 20 percent.

[17] For example, one military community appeared to be highly compliant, but that is not the only U.S. military population, where others may be embedded in different contexts.

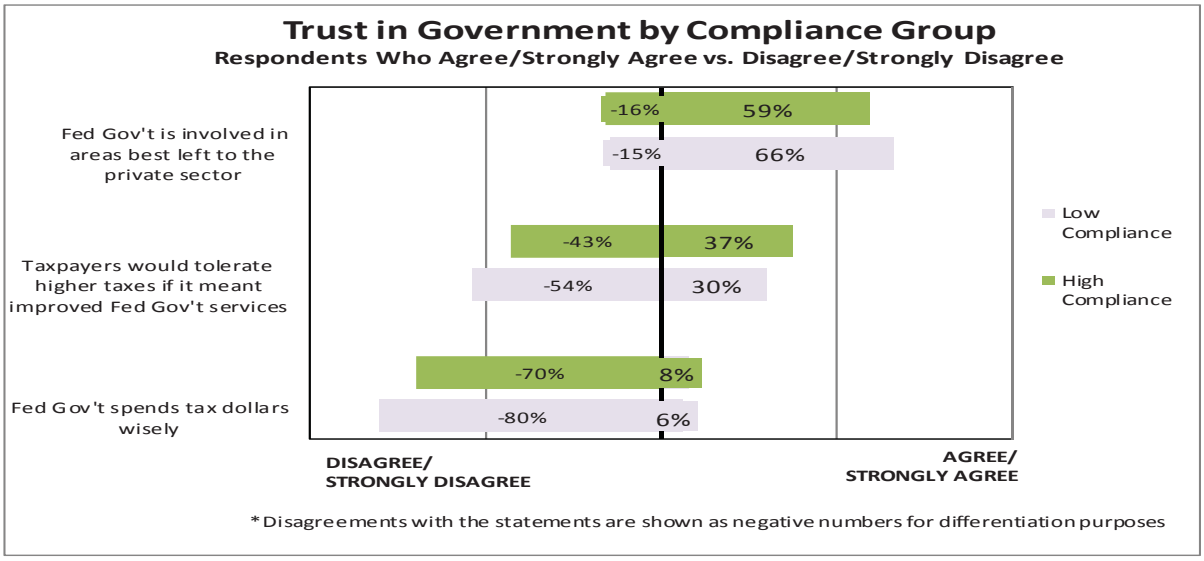
[18] Middletown (Muncie, Indiana) has been the subject of voluminous research on American social institutions. See *The First Measured Century* (PBS 2000); *Middletown* (PBS 1982); Theodore Caplow, et al., *All Faithful People: Change and Continuity in Middletown's Religion* (Minneapolis: Univ. of Minn. Press, 1983), *Middletown Families: Fifty Years of Change and Continuity* (Minneapolis: Univ. of Minn. Press, 1982); Rob't & Helen Lynd, *Middletown in Transition: A Study in Cultural Conflicts* (NY: Harcourt Brace, 1937), *Middletown: A Study in Modern American Culture* (NY: Harcourt Brace, 1929).

[19] This discussion of the National Survey cites aggregate figures that are weighted by EAC and DIF decile to reduce selection bias when projecting the summary statistics to the population of sole proprietors. For example, if five percent of the survey responses came from members of a stratum that made up 10 percent of the sole proprietor population, TAS gave the responses from that stratum more weight when computing summary statistics. Except as otherwise indicated, the discussion in this section generally focuses on findings where there are statistically significant differences (at a 95-percent level of confidence) between the high- and low-compliance groups.

Taxpayers in the high-compliance group were more likely to trust the government and the IRS.

Taxpayers in the high-compliance group were more likely to trust the government than those in the low-compliance group, potentially suggesting that negative views about the government promote symbolic noncompliance, as described in the typology (above). For example, those in the high-compliance group were less likely to agree that the government is involved in areas best left to the private sector (59 percent of the high-compliance group agreed vs. 66 percent of the low-compliance group), more likely to support higher taxes in exchange for improved government services (37 percent vs. 30 percent), and more likely to believe that the federal government spends tax dollars wisely (80 percent of the low-compliance group disagreed vs. 70 percent of the high-compliance group).[20] These results are generally consistent with research suggesting that trust in government has a positive effect on compliance.[21]

FIGURE 1. Trust in the Federal Government by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

Most taxpayers believe tax laws are unfair.

Only 15 percent of both groups agreed or strongly agreed that the tax laws are fair. Rather, most taxpayers believe that:

- Large businesses have loopholes to reduce their taxes that smaller businesses do not have;
- The wealthy have ways of minimizing their taxes that are not available to the average taxpayer;
- Not everyone pays his or her fair share; and
- The federal tax laws are unfair.

However, the low-compliance group was somewhat more likely to view the tax law as unfair than the high-compliance group (65 percent vs. 61 percent for the high-compliance group), which would be consistent with

[20] Most taxpayers in both groups (70 percent) also agreed or strongly agreed that taxes fund important government benefits and services.

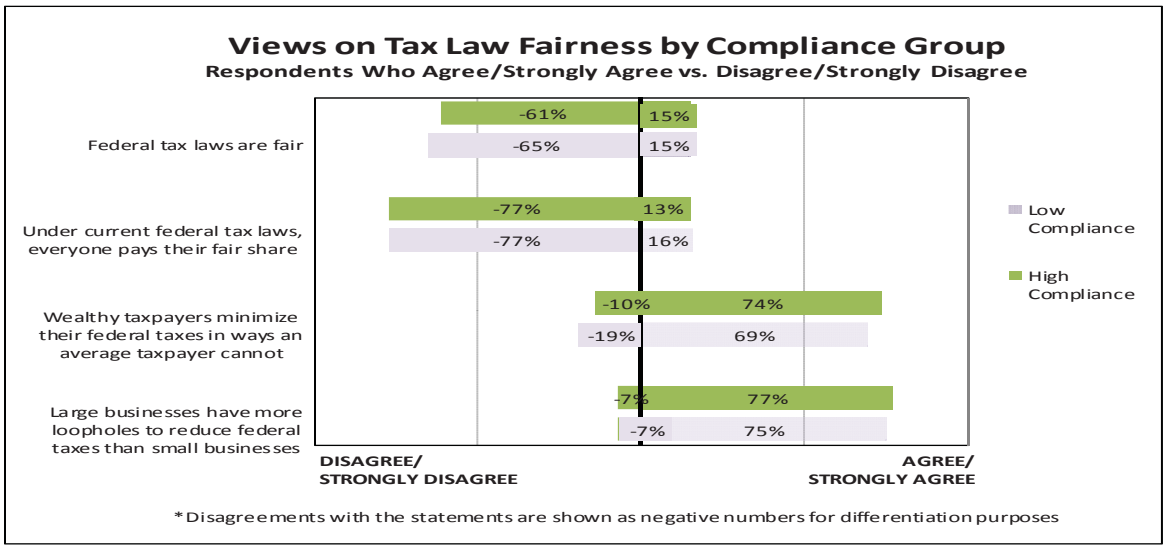
[21] See, e.g., Swedish Tax Agency, *Right From The Start, Research and Strategies* 6–7, 38-51 (Aug. 2005) (after surveying many papers from various disciplines, concluding that trust for tax agencies is an important determinant of voluntary compliance); Kristina Murphy, *The Role of Trust in Nurturing Compliance: A Study of Accused Tax Avoiders*, 28 Law and Human Behavior 187 (Apr. 2004) (finding that perceptions of procedural fairness and trust in the taxing authority had an impact on the motivation to comply); Tom R. Tyler, Why People Obey the Law 58–62 (Princeton Univ. Press 2006) (finding that "legitimacy" (defined as the perceived obligation to follow the law even if it is morally wrong, and respect and support for legal institutions, such as police and courts) has a significant positive impact on compliance after controlling for other variables). See also Joint Committee on Taxation, JCS-6-98, *General Explanation of Tax Legislation Enacted in 1998*, 19 (Nov. 24, 1998) (describing the 1998 IRS reorganization as needed to restore public confidence in the IRS, in large part, because "the Congress believed that most Americans are willing to pay their fair share of taxes, and that public confidence in the IRS is key to maintaining that willingness."); *Taxpayer Compliance, Volume 1: An Agenda for Research* 118 (Jeffrey A. Rother, John T. Scholtz, and Ann Dryden Witte eds., Univ. of Penn. Press 1989) (summarizing various studies that suggest commitment, attitudes toward the IRS, law, and government may have an impact on tax compliance).

symbolic noncompliance.[22] These views may lend support to calls for tax simplification as a way to increase tax compliance.

Those in the low-compliance group expressed less faith in the IRS.

The low-compliance group generally held more negative views about the IRS, potentially suggesting that negative views of the IRS promote symbolic noncompliance.[23] For example, those in the high-compliance group were more likely to believe that the IRS treats taxpayers fairly (47 percent of the high-compliance group agreed vs. 42 percent of the low-compliance group). The low-compliance group, by contrast, was more likely to report that the IRS is disrespectful (20 percent vs. 15 percent), and more concerned with collecting as much as it can than with collecting the correct amount of tax (42 percent vs. 25 percent—a 17-point difference!). Without adequate safeguards, the IRS's increasing use of automated procedures could give taxpayers this impression, which in turn, may contribute to noncompliance. These results may suggest the IRS could increase compliance by treating taxpayers fairly and publicly committing to initiatives promoting procedural justice and respect for taxpayers, thus promoting positive views about itself.[24]

FIGURE 2. Views About Tax Law Fairness by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

On the other hand, the low-compliance group was somewhat more likely to agree "that the IRS will work with you if you have difficulty paying your taxes," as 55 percent agreed vs. 49 percent for the high-compliance group. However, this view might have a greater effect on payment compliance than on reporting compliance. It could also reflect differences in knowledge about IRS procedures held by the groups. As noted below, the low-compliance group generally had more contact with the IRS.

---

[22] This difference is not statistically significant at a 95-percent level of confidence.

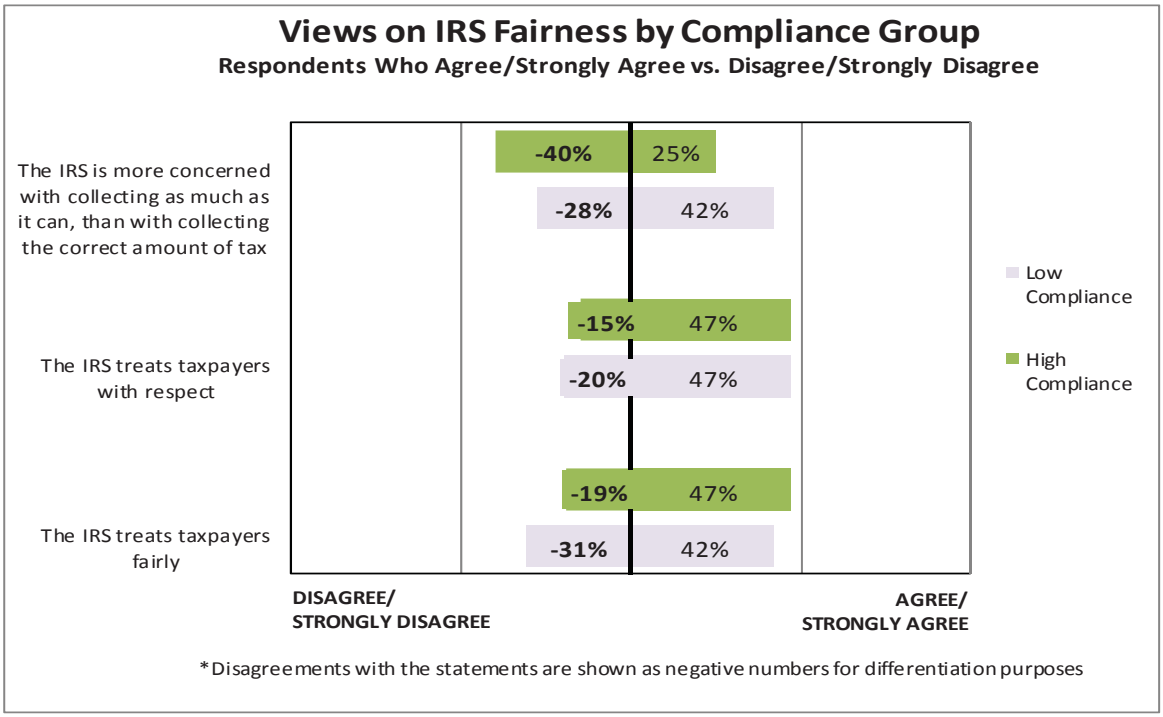
[23] This inference is consistent with previous research. See, e.g., *Taxpayer Compliance, Volume 1: An Agenda for Research* 93–96 (Jeffrey A. Rother, John T. Scholtz, and Ann Dryden Witte eds., Univ. of Penn. Press 1989) (discussing various studies).

[24] When IRS computers automatically propose adjustments and issue liens without reviewing all of the available information, the IRS appears more interested in collecting as much as possible than in collecting the correct amount. The National Taxpayer Advocate has suggested a wide range of steps the IRS could take to give taxpayers more confidence in the results of correspondence examinations, math error adjustments, and assessments against nonfilers. See, e.g., National Taxpayer Advocate 2011 Annual Report to Congress vol. 2, 63 (correspondence examination recommendations); National Taxpayer Advocate 2011 Annual Report to Congress 74 (math error recommendations); National Taxpayer Advocate 2011 Annual Report to Congress 93 (nonfiler recommendations). She has also recommended the IRS discontinue the practice of automatically filing the notice of federal tax lien (NFTL). See, e.g., National Taxpayer Advocate 2011 Annual Report to Congress 109, 128 (NFTL recommendations). Moreover, recent research suggests that collection alternatives (i.e., offers and installment agreements) are more closely associated with payment compliance than the automatic filing of a notice of federal tax lien. See *Investigating the Impact of Liens on Taxpayer Liabilities and Payment Behavior*, infra.

Those in the low-compliance group expressed less satisfaction with IRS services.

Those in the low-compliance group were more likely than those in the high-compliance group to report that the IRS does not offer the tax services they need (25 percent vs. 18 percent), that it is difficult to access the services the IRS provides (25 percent vs. 17 percent), and that they were more dissatisfied with the quality of the IRS services (27 percent vs. 21 percent). Thus, a lack of satisfaction with IRS services may contribute to noncompliance (*e.g.*, symbolic, procedural, lazy, or even unknowing noncompliance in terms of the typology), and the provision of better taxpayer services might increase tax revenue by improving compliance.

FIGURE 3. Views About IRS Fairness by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

Those in the low-compliance group expressed less trust in a preparer.

Those in the low-compliance group were more likely to report using a preparer than those in the high-compliance group (76 percent vs. 66 percent). This could be because the low-compliance group contained larger businesses (as noted below), facing larger tax preparation burdens. At least 90 percent of both groups reported that they always follow their preparer's advice, underscoring the importance of brokered compliance and noncompliance.[25] While both groups (79 percent and 80 percent of the low- and high-compliance groups, respectively) indicated they make sure they understand their return before signing, those in the high-compliance group were more likely to follow their preparer's advice than those in the low-compliance group (96 percent vs. 90 percent). This may suggest that preparers more often facilitate compliance instead of noncompliance.[26]

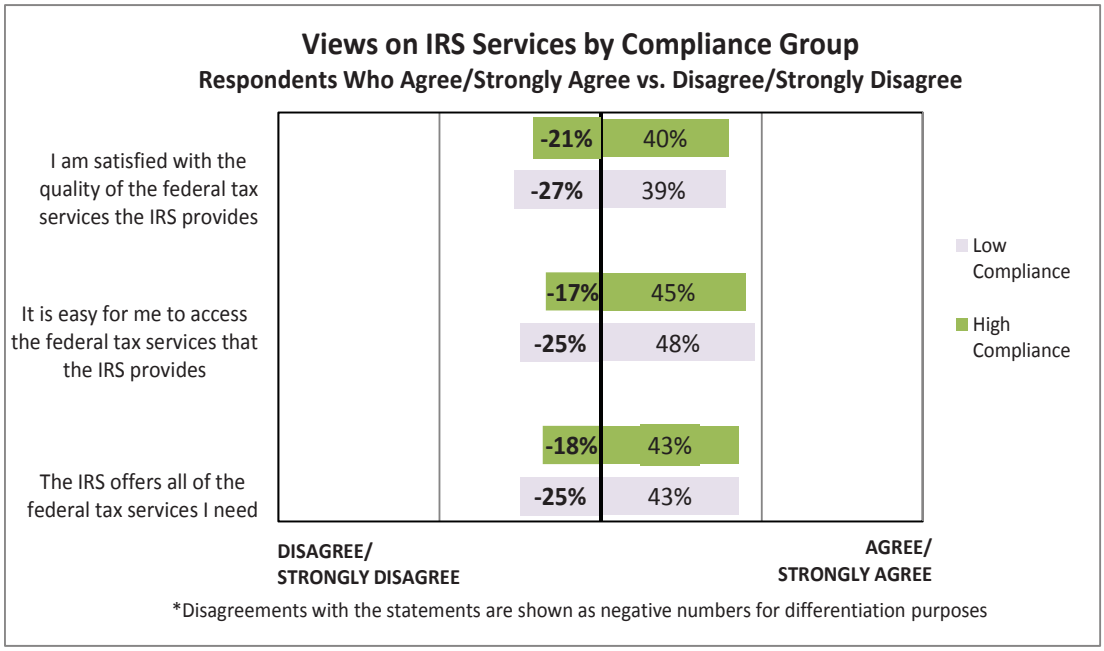
[25] As noted in the 2010 Proposal, the impact of the preparer on compliance probably depends on a combination of both the taxpayer's and the preparer's views toward compliance. Assume there are three types of preparers and taxpayers: (1) those who want to comply; (2) those who are willing to be more aggressive; and (3) those who are willing to cheat. Type one preparers may increase compliance by type two and type three taxpayers. Alternatively, those taxpayers may seek out type two or type three preparers. However, type two and type three preparers may reduce compliance by type one taxpayers unless those taxpayers either seek out type one preparers or are particularly resistant to the preparer's suggestions for tax savings. Similarly, type three taxpayers may pressure type one or type two preparers to be more aggressive than usual. See National Taxpayer Advocate 2008 Annual Report to Congress, vol. 2, § 3, 73 at 79-81 (Leslie Book, *The Need to Increase Preparer Responsibility, Visibility, and Competence*) (setting forth "The Types of Taxpayers and Preparers").

[26] Some research suggests preparers may improve compliance. See Steven Klepper, Mark Mazur, and Daniel Nagin, *Expert Intermediaries and Legal Compliance: The Case of Tax Preparers*, 34 J. L. and Econ. 205 (1991). See also Kim M. Bloomquist, Michael F. Albert, and Ronald L. Edgerton, *Evaluating Preparation Accuracy of Tax Practitioners: A Bootstrap Approach*, Proceedings of the 2007 IRS Research Conference 77 (2007) (finding preparers reduce math errors, but increase the incidence of potential misreporting). Other research suggests they do not reliably enhance compliance. See General Accounting Office (GAO),

Continued on next page.

Those in the high-compliance group were also more likely than those in the low-compliance group to indicate that the person who prepares their return finds creative ways to minimize their taxes (35 percent vs. 28 percent). Perhaps the groups had different views about what it means to find creative ways to minimize taxes, with the low-compliance group expecting the preparer to propose more aggressive positions. Another possibility is that those from the low-compliance group may view their preparers as part of the tax system, which they do not trust, as the IRS increasingly enlists preparers in its efforts to improve tax compliance (*e.g.*, by imposing due diligence requirements under Circular 230). Alternatively, those in the high-compliance group may simply seek out better preparers or at least have more meaningful conversations with them.

FIGURE 4. Views About IRS Services by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

Taxpayers in the low-compliance group were more likely to participate in local organizations and to report that other participants view the law and the IRS negatively.

Taxpayers in the high-compliance group were less likely than those in the low-compliance group to belong to a local business organization (11 percent vs. 16 percent), a local trade, labor, or other occupational organization (15 percent vs. 18 percent), or religious congregation (61 percent vs. 71 percent). To the extent association with these groups transmits local compliance norms, those norms appear to have a negative effect on compliance, rather than a positive one.

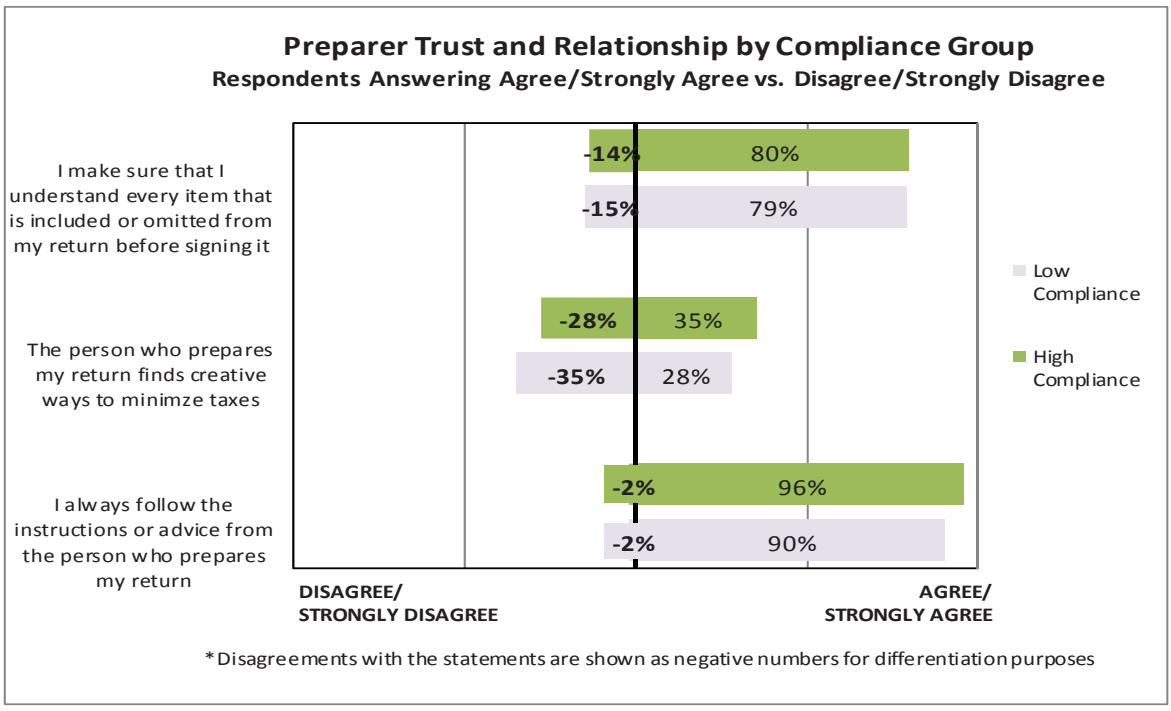
Continued from previous page.

GAO-02-509, *Tax Deductions: Further Estimates of Taxpayers Who May Have Overpaid Federal Taxes by Not Itemizing* (2002) (finding in 1998 about two million taxpayers overpaid their taxes by failing to itemize even though about half used a preparer); Treasury Inspector General for Tax Administration (TIGTA), *Analysis of Statistical Information for Returns with Potentially Unclaimed Additional Child Tax Credit* (2003) (finding about 230,000 returns filed by paid preparers in 2002 where taxpayers appeared eligible for Additional Child Tax Credits they did not claim); Janet Holtzblatt and Janet McCubbin, *Issues Affecting Low-Income Filers,* in The Crisis in Tax Administration 148, 159 (Henry Aaron and Joel Slemrod eds., 2004) (observing that about two-thirds of EITC returns, which have high levels of noncompliance, were prepared by paid preparers); Government Accountability Office (GAO), GAO-06-563T, *Paid Tax Return Preparers: In a Limited Study, Chain Preparers Made Serious Errors* 5, 23 (Apr. 4, 2006) (finding preparers made significant mistakes on 17 of the 19 returns prepared for GAO employees posing as taxpayers, including the omission of income on ten); TIGTA, Ref. No. 2008-40-171, *Most Tax Returns Prepared by a Limited Sample of Unenrolled Preparers Contained Significant Errors* 2 (Sept. 3, 2008) (finding preparers made mistakes on 17 of the 28 returns prepared for TIGTA employees posing as taxpayers, including six willful or reckless errors).

Those in the low-compliance group were more likely to participate in local organizations.

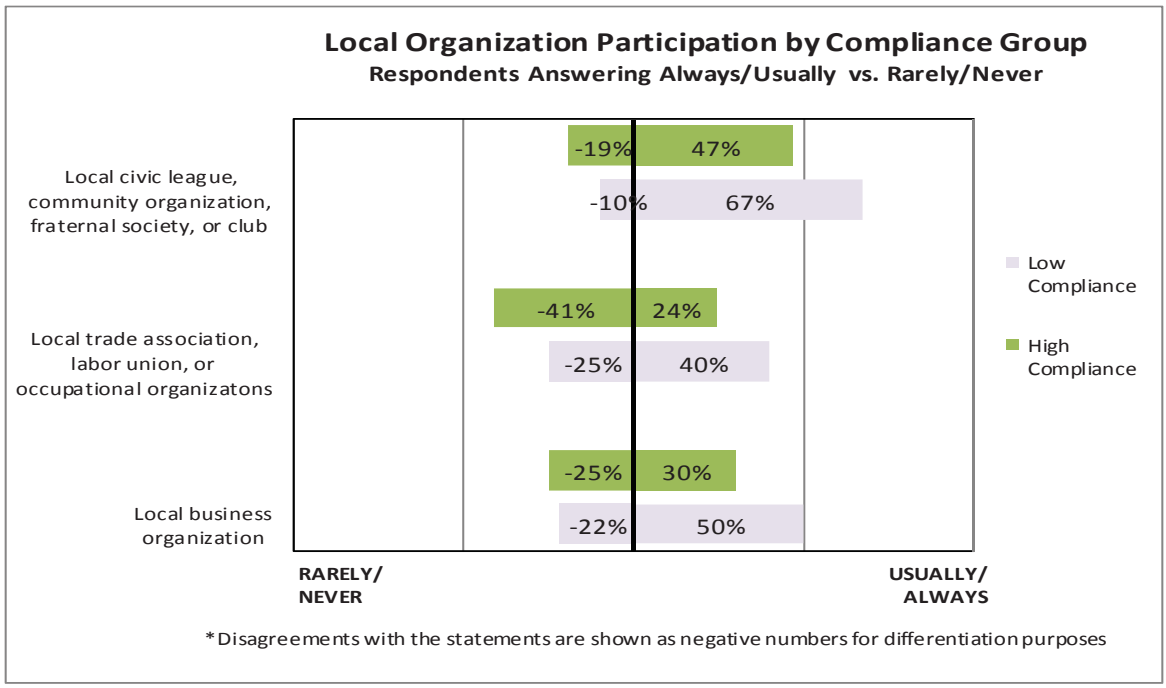
Among respondents who belong to local organizations, those in the low-compliance group were more likely to report that they usually participate. This was true for various organizations identified by the survey, including local business organizations (50 percent from the low-compliance group usually participate vs. 30 percent from the high-compliance group), local trade, labor, or occupational organizations (40 percent vs. 24 percent), and local civic, community, or fraternal organizations (67 percent vs. 47 percent). Thus, active participation in these groups appears to be negatively correlated with tax compliance, possibly promoting social noncompliance in terms of the typology. Perhaps those with a closer connection to local groups feel a weaker connection to the federal government, and a weaker obligation to comply with federal tax laws. They may also choose to associate with those who hold similarly negative views about the federal government and tax compliance, which reinforced their own views.

FIGURE 5. Preparer Trust and Relationship by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

FIGURE 6. Local Organization Participation by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

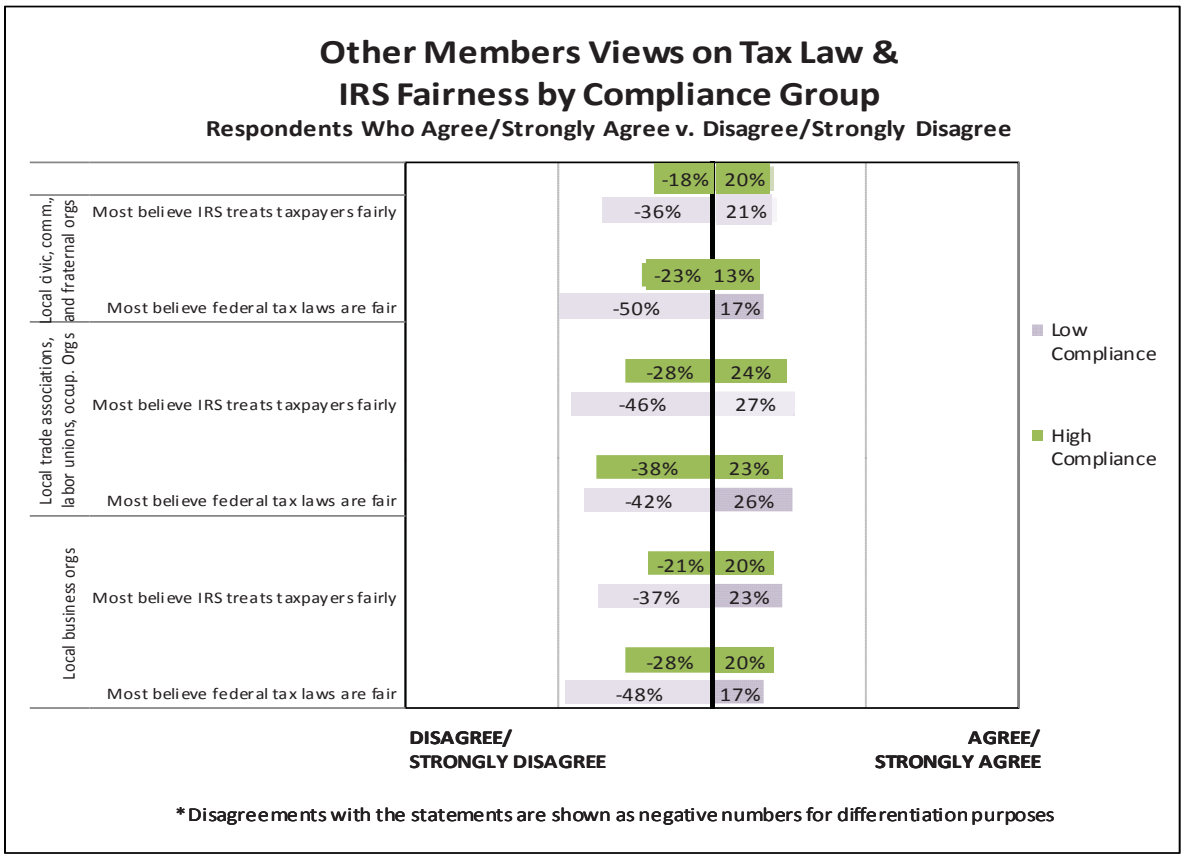
Those in the low-compliance group were more likely to report that other members of local organizations view tax laws and the IRS negatively.

Those in the low-compliance group were more likely than those in the high-compliance group to report that other members of local business organizations believe tax laws are unfair (48 percent of the low-compliance group vs. 28 percent of the high-compliance group) or that the IRS treats taxpayers unfairly (37 percent vs. 21 percent). They were somewhat more likely to report that other members of local trade, labor and occupational organizations believe tax laws are unfair (42 percent vs. 38 percent)[27] or that the IRS treats taxpayers unfairly (46 vs. 28 percent). They were also more likely to report that other members of local civic, community, and fraternal organizations believe the tax laws are unfair (50 percent vs. 23 percent) or that the IRS treats taxpayers unfairly (36 percent vs. 18 percent).[28] Participation in these organizations may have allowed taxpayers to learn that noncompliance is an acceptable norm among other participants, or perhaps they assumed that other participants shared their negative views. In any event, the differences in the responses to these questions by members of the high- and low-compliance groups may suggest that a person's perception about whether other participants in local organizations feel the tax law or the IRS is fair has an effect on their own compliance behavior (*e.g.*, social and symbolic noncompliance), perhaps eroding tax morale.

[27] This difference is not statistically significant at a 95-percent level of confidence.

[28] Curiously, 53 percent of those in the moderate-compliance group also disagreed or strongly disagreed that other members of these organizations believe the IRS treats taxpayers fairly, and the difference between their response and the average response of members of both other groups was statistically significant at a 95-percent level of confidence. Perhaps those in the moderate-compliance group are more concerned about any perceived mistreatment of others by the IRS because they are still trying to comply, whereas more of those in the low-compliance group are slightly less concerned about fairness because they have either given up on the IRS or are noncompliant for other reasons.

FIGURE 7. Other Members' Views About the Fairness of the Tax Law and the IRS by Compliance Group[29]



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence..

While most respondents reported that small businesses comply, those in the high-compliance group were more likely to report that their competitors do not.

According to social norms and reciprocity theories, taxpayers who believe most other taxpayers comply are more likely to reciprocate by complying.[30] However, the survey did not find that those in the high-compliance group were more likely to report that competitors were complying. Rather, those in the high-compliance group were less likely to do so—agreeing that most of their competitors report all of their income only 22 percent of the time as compared to 31 percent for the low-compliance group.[31] Moreover, there was no significant difference in the views of each group about whether many small businesses report all of their income (26 percent of the high-compliance group agree and 15 percent disagree, but 27 percent of the low-compliance group agree and 16 percent disagree).[32]

Most members of both groups also reported that they would be embarrassed if others found out they did not report all of their income. It is possible that the low-compliance group answered these questions defensively or feigned innocence—to avoid giving the impression that they were cheating or that the government

[29] Taxpayers were asked to provide a response with respect to members of the local organization(s) with which they most closely associate.

[30] See, e.g., Dan M. Kahan, *The Logic of Reciprocity: Trust, Collective Action, and Law*, 102 Mich. L. Rev. 71 (Oct. 2003).

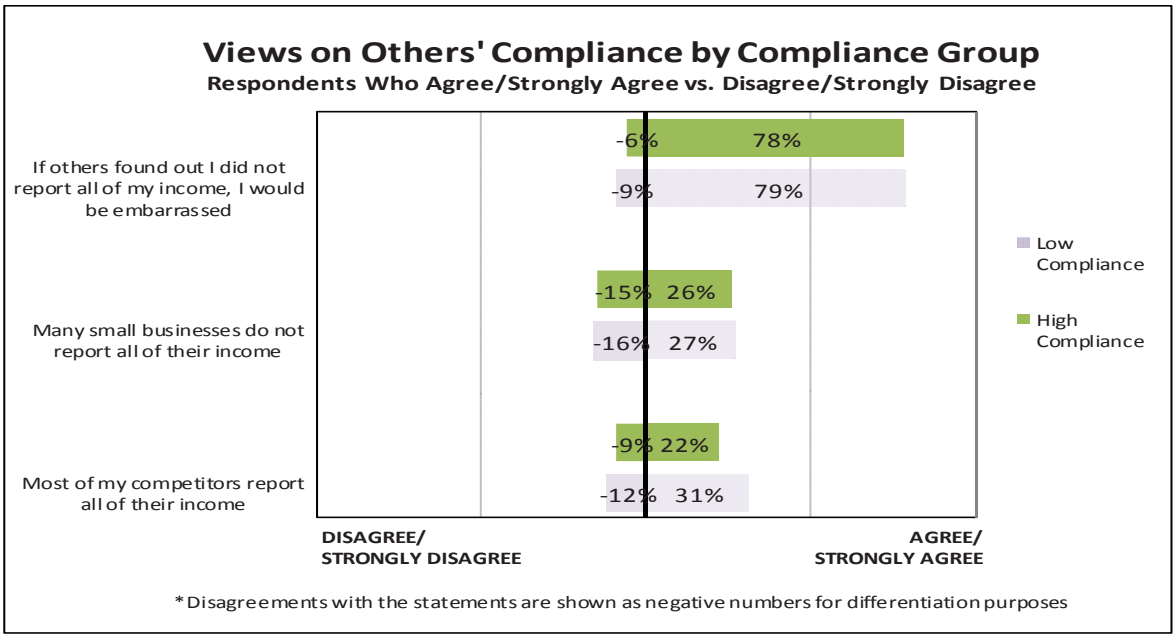
[31] However, most respondents (60 percent overall) were noncommittal, indicating they "don't know" or "neither agree or disagree."

[32] Both groups also reported that small businesses could survive even if they reported all of their income (54 percent of the high-compliance group vs. 56 percent of the low-compliance group). These differences are not statistically significant at a 95-percent level of confidence.

should do more to address noncompliance. Nonetheless, these results do not directly support the notion that social norms and reciprocity (at least among "most competitors") drive compliance decisions. As noted above, the views of other members of local organizations toward the IRS seemed to have a greater correlation with compliance than whether most competitors comply. Perhaps the norms of that peer group are more important than the norms of competitors, though both groups said they would be embarrassed if others learned they were noncompliant.

In other words, the tax compliance decision may be less about the views of others or economics, and more about how the business views itself in relation to the federal government. As noted above, those with negative views toward the federal government more often associate with like-minded individuals at the local level.

FIGURE 8. Views on Others' Compliance by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

Smaller businesses with local customers and those in professional or technical businesses were more often in the high-compliance group.

The low-compliance group had larger businesses. The low-compliance group had an average of about seven employees, as compared to about four for the high-compliance group. Similarly, the low-compliance group had average gross receipts of about $87,000, as compared to about $46,000 for the high-compliance group.[33] The low-compliance group also identified its customers as "primarily national" more frequently—19 percent of the time—as compared to 17 percent for the high-compliance group. While the difference is small, this finding is somewhat curious in light of the finding that those in the low-compliance group feel a closer connection to local organizations than national ones, though some local organizations may be local chapters of national ones. Of course, business owners may feel more of a connection to local organizations that they chose to associate with than to customers with whom they may not interact in this age of e-commerce.

Another explanation could be that as businesses grow, the economic benefit of noncompliance increases but the expected penalty does not—a finding consistent with economic deterrence theory. However, this explanation seems inconsistent with the notion that smaller businesses, which are more likely to have informal

[33] Concerned that some taxpayers might have lower DIF scores simply because they have less income that is not subject to information reporting, and thus less opportunity to cheat, TAS analyzed the sample further. TAS looked at all income sources (not just Schedule C income) and found that taxpayers in the high-compliance group from each EAC strata reported a significant amount of income that is not subject to information reporting. Thus, while income transparency likely affects reporting compliance, its effect on a person's DIF score, if any, does not always overshadow other factors.

accounting systems and deal in cash, are less likely to be compliant than larger ones that need to have formal financial accounting systems to prevent theft and to reflect any positive net income on those systems and their tax returns to obtain financing.[34] However, even businesses in the low-compliance group were relatively small, possibly small enough to retain informal accounting systems.

Taxpayers in construction-related and real estate-related industries appeared to be less compliant than those in other industries, as they each comprised 9 percent of the low-compliance group, but only 4 percent of the high-compliance group. By comparison, those in professional and technical service industries appeared to be more compliant, comprising 26 percent of the high-compliance group and 17 percent of the low-compliance group.[35] Perhaps information reporting, which generally promotes compliance, was more prevalent among professional and technical service industries than in construction and real estate. Industry-related norms, the type of noncompliance involved, or the type of taxpayers involved, as described below in our analysis of the Community Survey, could also have played a role.

Both high- and low-compliance groups professed a "moral" obligation to report income accurately.

Nearly all—96 percent of both groups—feel a moral obligation to report all of their income correctly. Moreover, those in the low-compliance group were more likely to say that everyone should correctly report all of their income—97 percent of the low-compliance group agree vs. 94 percent of the high-compliance group.[36] However, the low-compliance group may have answered these questions aspirationally (*e.g.*, they may not be living up to their aspirations because tax morale does not drive their tax compliance behavior) or defensively, to avoid making an admission.

Economic deterrence may not drive compliance decisions by those in either the high- or low-compliance groups.

Those in the low-compliance group were more likely than those in the high-compliance group to report that achieving financial success is important (88 percent vs. 85 percent) and that taking risks is necessary to achieve financial success (68 percent vs. 61 percent).[37] One might expect people who express financial concerns, as both groups did, to be motivated by economic deterrence.

However, the survey responses provide little support for the view that economic deterrence has an effect on reporting compliance. In terms of the typology, the survey did not reveal asocial noncompliance. On one hand, those in the low-compliance group were more likely to agree that hearing about people who were caught underreporting makes them more careful with their own taxes (66 percent vs. 61 percent for those in the high-compliance group), a response consistent with the notion that economic deterrence (or a lack thereof) has a stronger effect on their compliance decisions than on those of respondents in the high-compliance group.[38]

[34] As noted above, the accuracy of the DIF scoring algorithm could affect the results. For example, if the DIF overestimates the actual compliance of small cash businesses, which generate income that is difficult to detect, then the results would indicate that small businesses are more compliant than they actually are.

[35] Under "professional, scientific, and technical services," the North American Industry Classification System (NAICS) includes legal, accounting, engineering, design, computer, management, research, and advertising services.

[36] This difference was not statistically significant at a 95-percent level of confidence.

[37] Those in the low-compliance group were also more likely to be male. Males are often thought to have less aversion to risk. See, e.g., Alexandra Niessen and Stefan Ruenzi, *Sex Matters: Gender Differences in a Professional Setting*, Ctr. for Fin. Research, Working Paper No. 06-01, 14 (Feb. 2007), available at http://hdl.handle.net/10419/57738.

[38] This difference was not statistically significant at a 95-percent level of confidence. Of course, most people in both groups did acknowledge that such statements make them more careful, lending some support to economic deterrence theory.

FIGURE 9. National Survey Respondents' Industry by Compliance Group

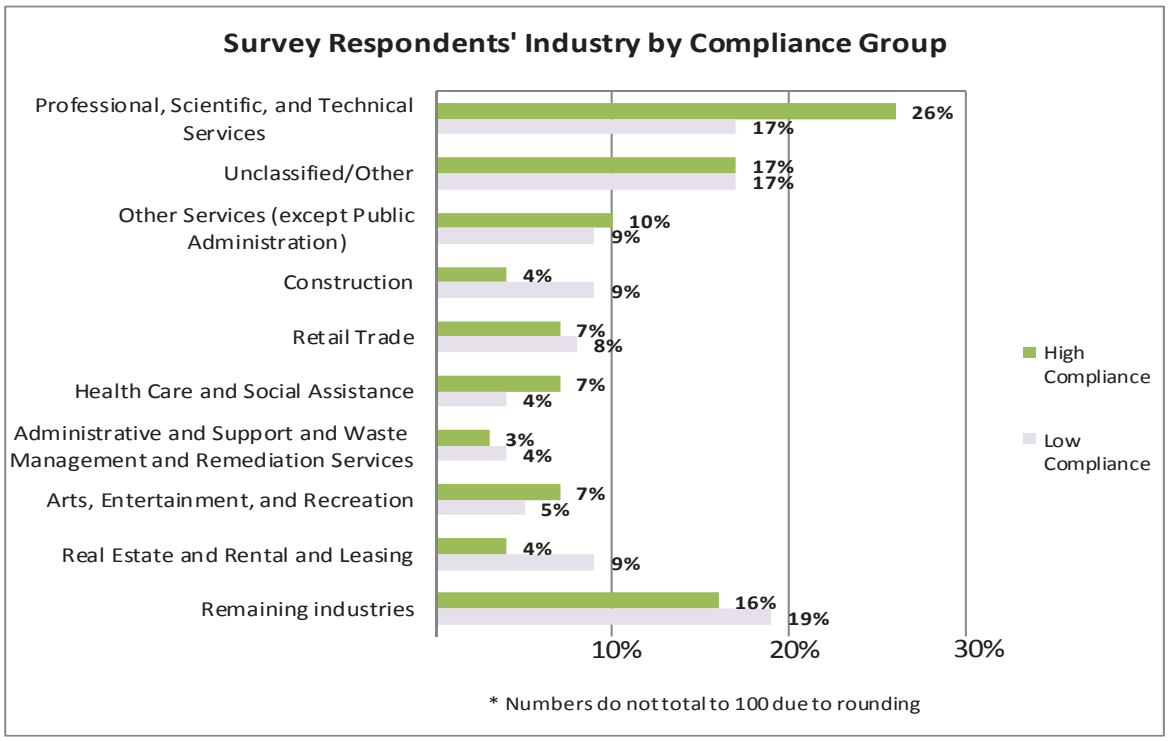
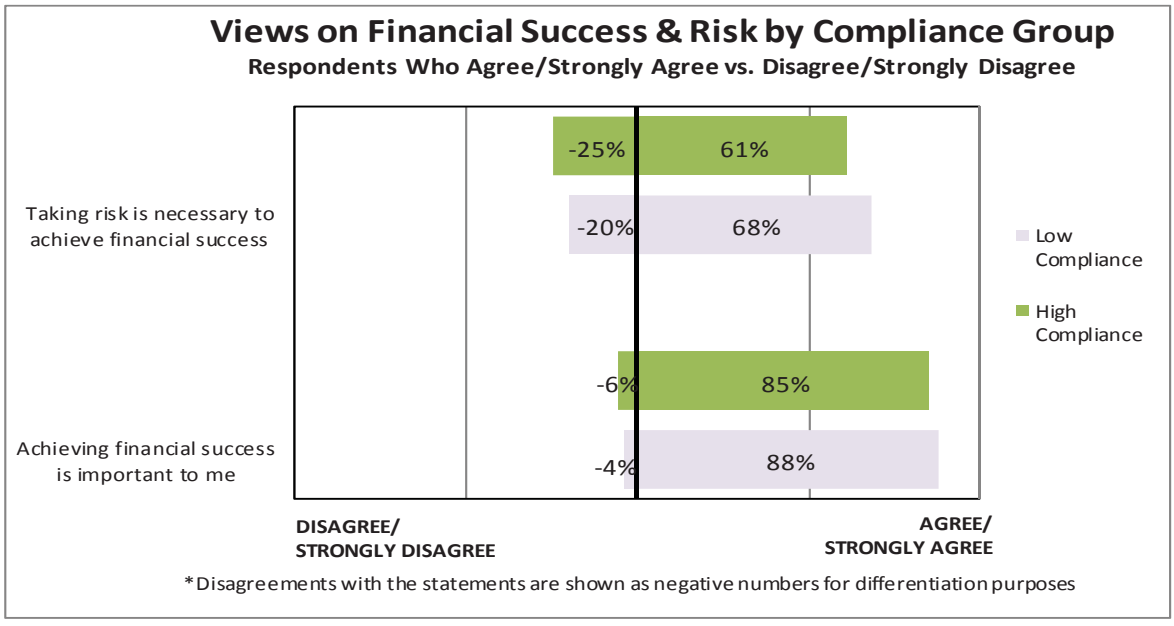


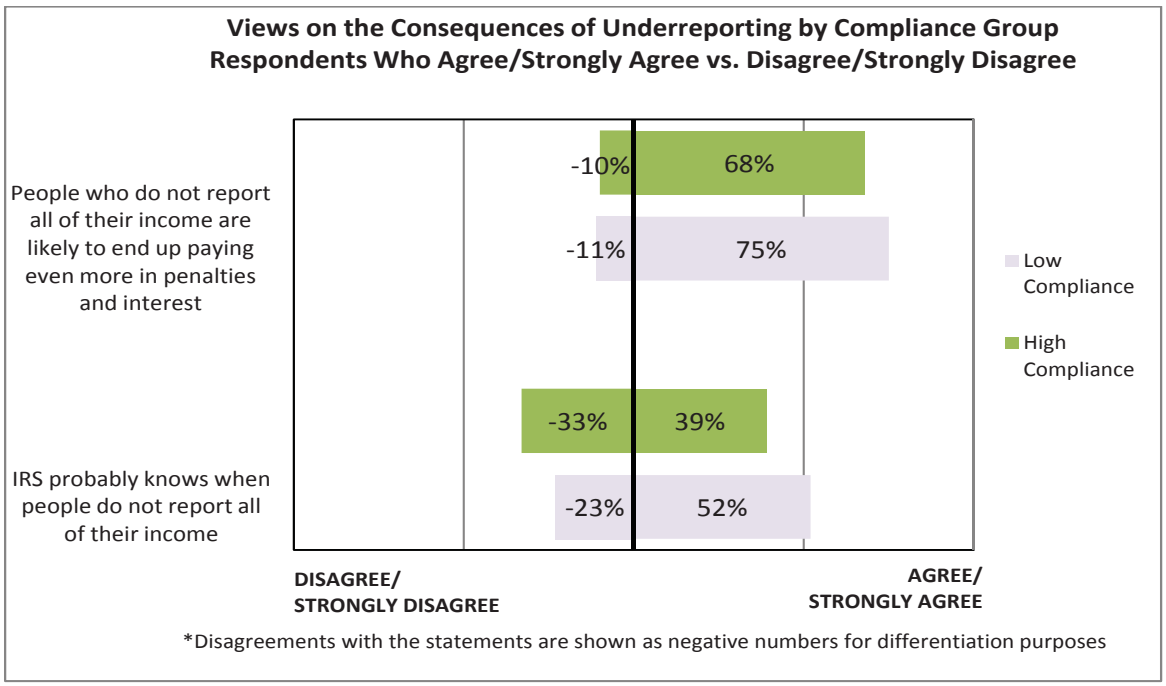
FIGURE 10. Views on Financial Success & Risk by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

On the other hand, those in the low-compliance group were also more likely to agree that the IRS probably knows when people do not report all of their income (52 percent vs. 39 percent for those in the high-compliance group); and that people who do not report all of their income are more likely to end up paying even more in penalties and interest (75 percent vs. 68 percent). If economic deterrence was a motivating factor for those in the low-compliance group, then (if answering truthfully) they might agree more often than those in the high-compliance group that it pays to cheat. They did not. Thus, the responses to these questions do not support the notion that a lack of economic deterrence drives noncompliance for those in the low-compliance group.[39]

FIGURE 11. Views on the Consequences of Underreporting by Compliance Group



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

Alternatively, even though those in the low-compliance group generally do not believe it pays to cheat, they have slightly larger businesses, slightly more employees and may be willing to take more risk on their taxes if necessary to expand their businesses or to meet payroll, particularly if the alternative is to discontinue operations.[40] These are the same reasons that small businesses sometimes fail to make employment tax deposits.[41]

Another possibility is that responses by the small subset of the low-compliance group that had actually been caught cheating affected the results. Those in the low-compliance group had been subject to IRS examination or collection contacts more often than those in the high-compliance group. Nine percent of the low-

[39] Of course it is possible that those in the low-compliance group answered these questions defensively—to avoid the implication that they may not have reported all of their income.

[40] Indeed, the low-compliance group was more likely to agree that you have to take risks to succeed.

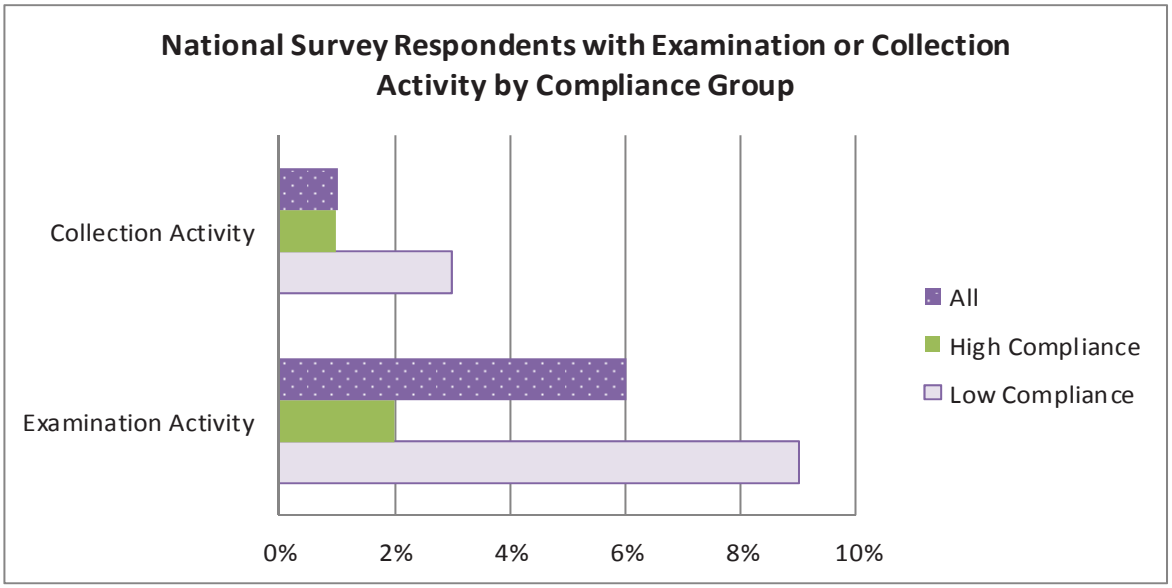
[41] See, e.g., SB/SE Research, 2009 Nationwide Tax Forum Focus Groups, DEN0116, *Your Clients and the Economy—How Can the IRS Help?* 3 (Jan. 2010) ("Because there is no money to pay expenses and meet obligations, participants stated that small business taxpayers are experiencing a number of secondary effects to include: falling behind on payments; not filing tax returns on time (or at all); going 'underground'; and 'burying their heads in the sand'.… [t]he IRS is not seen as a priority because small business taxpayers do not experience any immediate consequences of noncompliance. Therefore, payroll taxes and estimated taxes are last on the list."). Consistently, IRS research finds that taxpayers who owe a balance upon filing their returns are more likely than others to understate their tax liabilities. See Charles Christian, Phoenix District Office of Research and Analysis, *The Association Between Underwithholding and Noncompliance* 1–2 (July 14, 1995) (finding that "[o]n average, understated tax on balance due returns is ten times as large as understated tax on other returns.").

compliance group had been subject to an IRS examination, as compared to 2 percent of the high-compliance group.[42] Similarly, 3 percent of the low-compliance group had been subject to IRS collection activity as compared to 1 percent of the high-compliance group.[43]

Views about complexity were mixed, but most agreed the tax rules are so complicated it is very difficult to get a tax return exactly right.

Researchers have suggested that taxpayers who face complicated rules may be unable to comply, or may use complexity as a reason to justify noncompliance.[44] Survey responses about complexity were mixed and provide little insight about how complexity or burden affects compliance. On one hand, most taxpayers (more than 73 percent in both groups) agreed that their record-keeping system made it easy to compute their income tax. Most (about 64 percent overall) also agreed that the rules about what to report as income are clear. Thus, while complexity may have been a barrier to compliance for some, it was not a significant barrier for most respondents.

FIGURE 12. Percentage of National Survey Respondents Subject to IRS Examination or Collection Activity by Compliance Group



On the other hand, as noted above, most (70 percent of those who had tax preparation assistance) reported that they did not know the tax laws well enough to prepare their own returns. In addition, most agreed that the tax rules are so complicated that it is very difficult to get a tax return exactly right (56 percent overall agreed). However, taxpayers in the high-compliance group were more likely to agree with this statement than those in the low-compliance group (62 percent vs. 58 percent). Perhaps taxpayers in the high-compliance group were more concerned about making inadvertent errors than those in the low-compliance group.

Summary of the National Survey Results

Respondents from the low-compliance group were more likely to report that the government is too big and wastes tax dollars, that tax laws are unfair, and that the IRS is unfair (*e.g.*, often believing the IRS is more

[42] IRS, Compliance Data Warehouse (2012).

[43] Id.

[44] See, e.g., *Taxpayer Compliance, Volume 1: An Agenda for Research* 118, 128–129 (Jeffrey A. Rother, John T. Scholtz, and Ann Dryden Witte eds., Univ. of Penn. Press 1989).

concerned with collecting as much as possible instead of the correct amount, and indicating less satisfaction with IRS services). Members of the low-compliance group may have used these beliefs to justify noncompliance.

Surprisingly, respondents in the low-compliance group were more likely than those in the high-compliance group to believe that the IRS detects and penalizes noncompliance. This finding may seem inconsistent with the popular belief that small businesses cheat on their taxes because they do not think they will get caught (*i.e.*, insufficient economic deterrence).[45]

Both groups were idealistic, professing that it is morally wrong to cheat. Most members of both groups also reported that they would be embarrassed if others discovered they did not report all of their income. For those in the low-compliance group, however, other factors may have overshadowed these positive moral convictions and social pressures.

Those in the low-compliance group were more likely than those in the high-compliance group to participate in local organizations, which one might expect to be a source of positive tax compliance norms. However, they were more likely to report that other members of these organizations believe the law and the IRS are unfair, potentially countering the positive influence these affiliations might otherwise have had on tax compliance. Moreover, the closer association with local organizations by members of the low-compliance group could have undermined their connection with the nation and the national tax system as a whole.

The norms of competitors appeared to have little correlation with compliance. This may suggest that norms do not operate by reference to competitors. Rather, the views of other participants in local organizations may be more important.

Those in the low-compliance group operated slightly larger businesses and were somewhat more likely to use a preparer who could have persuaded them to comply or facilitated noncompliance—brokered compliance (or noncompliance) in the typology above. However, they were also less likely to follow the preparer's advice than those in the high-compliance group, potentially weakening any positive influence that the preparer sought to exert.

By contrast, respondents from the high-compliance group, while slightly smaller and less likely to use a preparer, were more likely to follow the preparer's advice. They were also less likely to participate in local organizations, suggesting that their compliance level was not social but motivated rather by morality, trust in government, trust in the IRS, or other internal factors contributing to high tax morale. It is unclear if affiliations and communications with those in local organizations who have little faith in government, federal tax law, and the IRS erodes the force of one's tax morale, or if those who become noncompliant seek to affiliate with those who would be more likely to feel that noncompliance was justified.

In either case, these results may suggest that the government could improve reporting compliance by improving the perceived fairness and efficiency of the government, the tax law, and the IRS; and by simplifying the tax code, improving procedural protections, and minimizing the IRS's reliance on procedures that may seem unfair (*e.g.*, excessive automation and lack of personal contact).[46] To address the perception by members of local groups that the tax law and the IRS are unfair, the IRS might retain a local presence and conduct outreach and education events for these groups, particularly in low-compliance communities (discussed below).[47]

---

[45] See, e.g., Susan Morse, Stewart Karlinsky, and Joseph Bankman, *Cash Businesses and Tax Evasion*, 20 Stan. L. & Pol'y Rev. 37 (2009) (discussing anecdotal accounts of cash businesses that did not expect the IRS to discover underreporting). This finding does not necessarily imply that taxpayers are economically irrational, particularly if they have no other source of financing and face the choice of either going out of business or underreporting. If a taxpayer could possibly use the temporary tax "savings" from underreporting to earn more than the likely tax, penalties and interest, which the IRS might collect later, then it may be rational for the taxpayer to underreport income even if he or she expects that the IRS will detect the noncompliance and impose penalties and interest. Moreover, other survey responses suggest that the low-compliance group was less risk averse than the high-compliance group.

[46] See, e.g., National Taxpayer Advocate 2009 Annual Report to Congress 3 (Most Serious Problem: *The Time for Tax Reform is Now*) (summarizing tax simplification proposals); *Complexity and the Tax Gap: Making Tax Compliance Easier and Collecting What's Due*, hearing before the S. Comm. on Finance (June 28, 2011) (testimony of Nina E. Olson, National Taxpayer Advocate) (same); National Taxpayer Advocate 2007 Annual Report to Congress 275 (Most Serious Problem: *The Accuracy-Related Penalty in the Automated Underreporter Units*) (recommending that IRS computers stop proposing negligence penalties); National Taxpayer Advocate 2008 Annual Report to Congress, vol. 2, at 2 (*A Framework for Reforming the Penalty Regime*) (proposing improvements to the penalty regime); National Taxpayer Advocate 2011 Annual Report to Congress 524 (recommending limits on expansion of IRS math error authority); *Options for Expanding the Remedies to Address Taxpayer Rights Violations*, supra (proposing remedies to strengthen procedural protections).

[47] See, e.g., National Taxpayer Advocate 2009 Annual Report to Congress 346–50 (legislative recommendation to require at least one appeals officer and one settlement officer in each state); National Taxpayer Advocate 2008 Annual Report to Congress 176, 192 (Most Serious Problem: *Local Compliance Initiatives Have Great Potential but Face Significant Challenges*) (recommending ways to enhance local compliance initiatives). TAS has at least one office in each state and Local Taxpayer Advocates routinely conduct outreach to local groups.

Key Findings of the Community Survey

According to the 2010 Proposal, the Community Survey was to address:

> What types of communities have homogeneous compliance attitudes? What local social practices, institutions (*e.g.*, volunteer, educational, and religious institutions), or attitudes increase or decrease compliance at the community level and why? Do taxpayers in communities with notably high or low levels of compliance identify more with the nation as a whole or the local community?[48]

One possibility was that the high-compliance communities would be homogeneous towns where residents have strong ties to local groups and institutions. This view could arise from the theory that social norms promote compliance. The Community Survey results offer a significantly different view. In short, like those in the low-compliance group, those in low-compliance communities appear to exhibit a stronger association with local institutions than national ones such as the federal government. Moreover, in constructing the Community Survey sample (described above), TAS discovered that taxpayers with high-compliance are not concentrated in homogeneous communities, at least not very many of them. Taxpayers in the low-compliance communities appeared in more concentrated geographic clusters across the country, especially in the South and West, as set forth below.

Taxpayers in the high-compliance communities were more geographically dispersed than those in the low-compliance communities.

As discussed above, to identify survey respondents who were sole proprietors, TAS used the DIF, an IRS indicator of the likelihood of underreported tax based on items reported on the return. While this measure may be imperfect, it is not geographically biased. Consequently, it was uncertain whether returns with similar compliance levels, as measured by DIF, would cluster geographically. From all areas, cities, and towns, in the U.S., those with median DIF scores in the top or bottom 30 percent constituted the low- or high-compliance communities, respectively.[49] As it turned out, populations ranging from 20,000 to 414,000 had measurably low compliance in 365 areas, cities and towns. At the same time, populations ranging from 22,000 to 60,000 had measurably high compliance in a few sites. The site selection process confirms a geographic aspect of tax compliance. In particular, low compliance levels clustered in geographic communities, while high compliance levels were more individually dispersed.

The map below shows that low-compliance communities appeared in 24 states. The map shows concentrations of low compliance, as measured, where it may become socially acceptable. The map reflects the locations of low-compliance communities but not their populations, some of which were larger than others. The site selection process was not an enforcement screen, lacking indicators of type or magnitude of noncompliance. Instead, the map helps visualize the social nature of noncompliance. The geographic observation raises issues about fostering communities of compliance given a social aspect to noncompliance.[50]

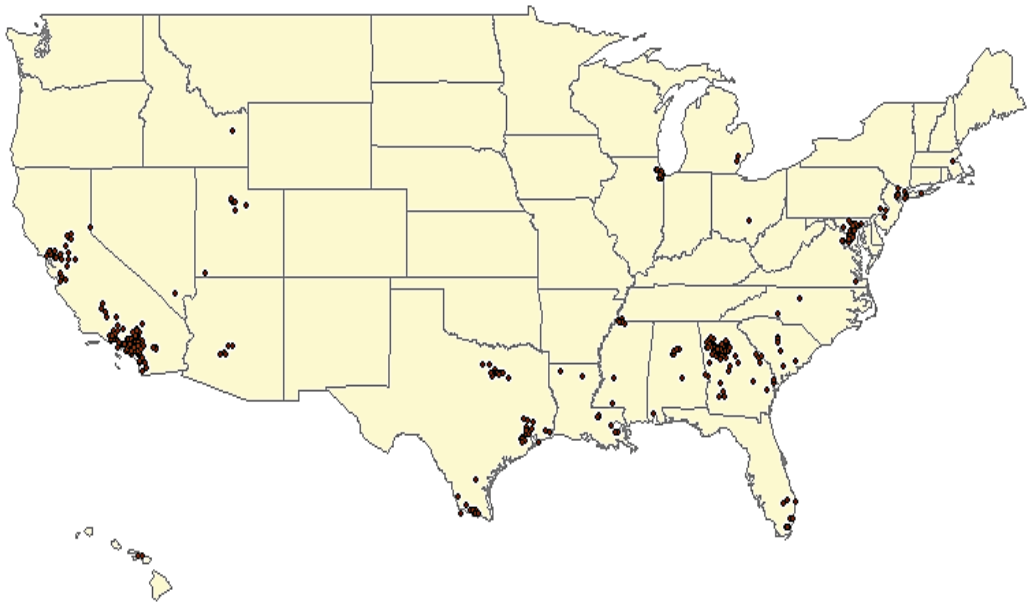
---

[48] 2010 Proposal at 86–87.

[49] TAS identified geographic communities from the addresses with ZIP codes reported by the taxpayers on their returns, generally cities, towns, or other distinct areas as denominated by the U.S. Postal Service.

[50] Geographers have classified regions of the U.S. based on local history, values, behavior, and culture. See Colin Woodard, AMERICAN NATIONS: A HIST. OF THE ELEVEN RIVAL REGIONAL CULTURES OF NO. AMER. (N.Y.: Viking, 2011); Joel Garreau, NINE NATIONS OF NO. AMER. (Boston: Houghton Mifflin, 1981); Raymond Gastil, CULTURAL REGIONS OF THE U.S. (Seattle: Univ. of Wash. Press, 1975); Wilbur Zelinsky, CULTURAL GEOGRAPHY OF THE U.S. (Prentice Hall, 1973).

FIGURE 13. Map of Low-Compliance Communities

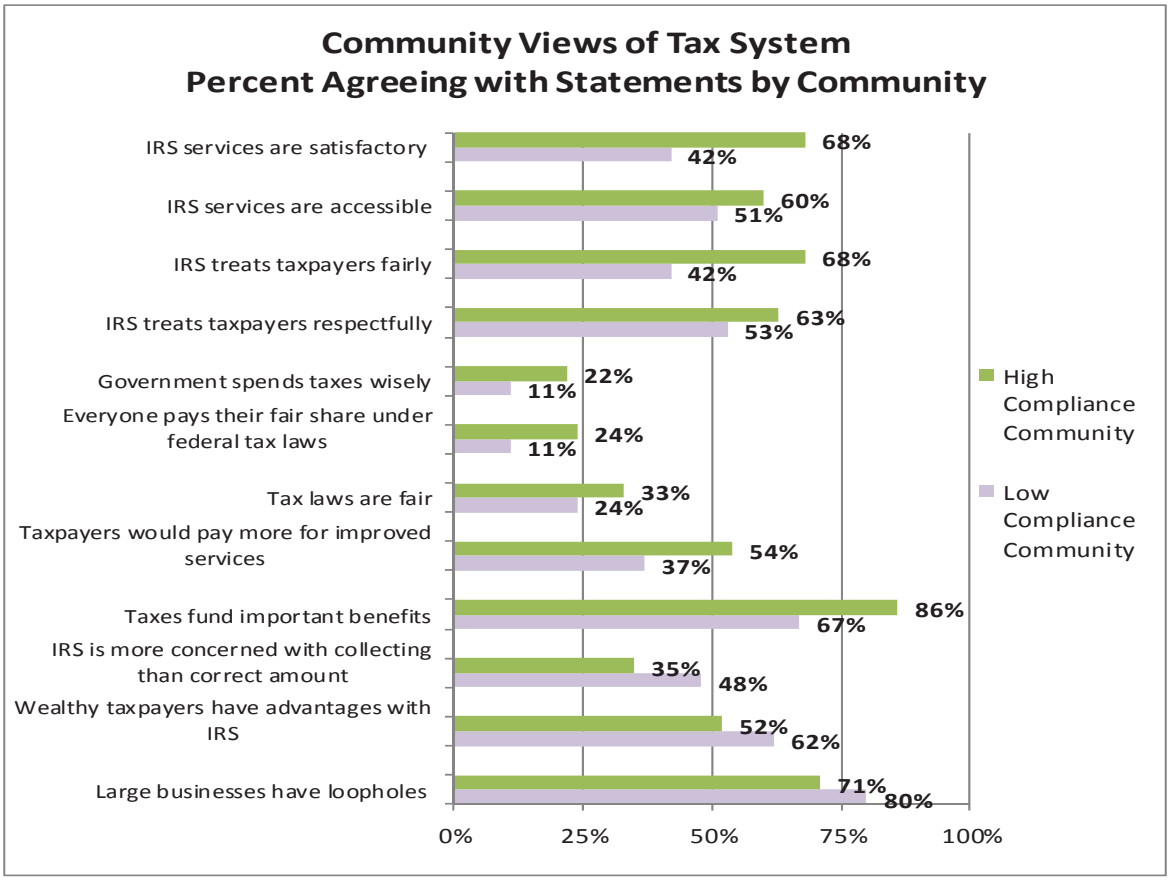


Respondents from low-compliance communities were suspicious of the tax system and its fairness, whereas those from high-compliance communities responded positively to government.

Respondents from low-compliance communities believed that large businesses and wealthy taxpayers have loopholes or advantages with the IRS (80 percent *vs*. 71 percent of those from the high-compliance communities, and 62 percent *vs*. 52 percent, respectively), which is more concerned with collecting as much as it can rather than the correct amount (48 percent *vs*. 35 percent). On the other hand, those from high-compliance communities felt that taxes fund important benefits (86 percent *vs*. 67 percent of those from low-compliance communities); taxpayers would pay more for improved services (54 percent *vs*. 37 percent); tax laws are fair (33 percent *vs*. 24 percent); everyone pays their fair share under federal tax laws (24 percent *vs*. 11 percent); and the government spends taxes wisely (22 percent *vs*. 11 percent). Similarly, those from high-compliance communities felt the IRS treats taxpayers respectfully (63 percent *vs*. 53 percent) and fairly (68 percent *vs*. 42 percent) with accessible (60 percent *vs*. 51 percent) and satisfactory services (68 percent *vs*. 42 percent).

While taxpayers in the low-compliance communities may tend to identify less with federal agencies, respondents from the high-compliance communities identified with the nation as a whole. In terms of the factors introduced above, respondents from high-compliance communities expressed trust in government, while the responses of the low-compliance group suggested a symbolic type of noncompliance.

FIGURE 14. Community Views of the Federal Tax System



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

Respondents from high-compliance communities were more likely to rely on preparers.

A substantial majority of the respondents from high-compliance communities used a third-party preparer (86 percent *vs.* 65 percent of low-compliance respondents) and always followed that person's advice (98 percent *vs.* 89 percent). Nevertheless, those from high-compliance communities made sure to understand the return before signing (91 percent *vs.* 84 percent from low-compliance communities).

Among business classifications, the biggest cluster in low-compliance communities was under "professional, scientific, or technical services"; in high-compliance communities, the "other" service industry.

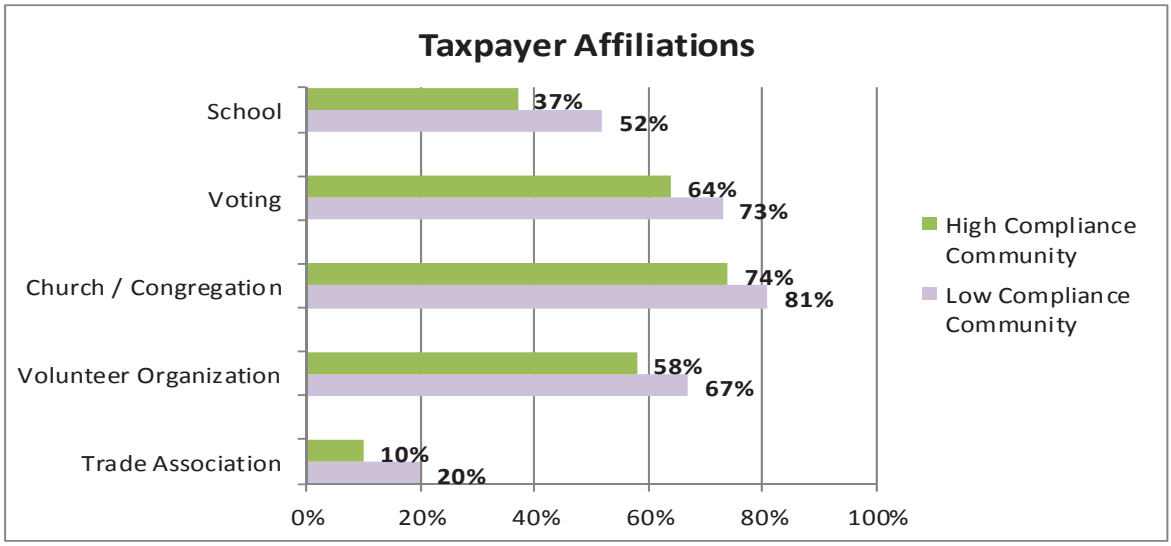
Respondents from the high-compliance communities most frequently clustered in "other services" (22 percent *vs.* 11 percent of low-compliance respondents), whereas those from the low-compliance communities most frequently clustered in "professional, scientific, or technical services" (22 percent *vs.* 11 percent from the high-compliance communities).[51] Those from the high-compliance communities were more than twice as likely to speak a language other than English at home (22 percent *vs.* 9 percent from the low-compliance communities). The Community Survey may have identified a unique type of "social" compliance related to a particular socio-economic experience, that of a linguistic minority employed in the service industry who expressed trust in government.

[51] Under "other" services, NAICS includes repair & maintenance, personal & laundry, civic & social, and private household services.

Low-compliance community respondents reported more participation in civic institutions than their high-compliance community counterparts.

Low-compliance community respondents were more likely than high-compliance community respondents to belong to a trade association (20 percent *vs.* 10 percent), volunteer organization (67 percent *vs.* 58 percent), or church or other religious congregation (81 percent *vs.* 74 percent),[52] and to vote (73 percent *vs.* 64 percent) or send children to local schools (52 percent *vs.* 37 percent).[53] Within those affiliations, those from low-compliance communities were more likely to disagree (or strongly disagree) with the propositions that most members believe the tax laws and IRS are fair (respectively, 29 percent *vs.* 18 percent and 25 percent *vs.* 15 percent for volunteer organizations; 32 percent *vs.* 16 percent and 26 percent *vs.* 13 percent for churches; and 29 percent *vs.* 14 percent and 20 percent *vs.* 9 percent for elected officials).[54] In other words, those from the low-compliance communities tend to belong to groups, which they believe share the view that taxes are unfair. In terms of the factors introduced above and the typology of noncompliance, set forth in Table 1, *Typology of Noncompliance*, above, these affiliations may be a form of social noncompliance.[55]

TABLE 15. Taxpayer Participation by Type of Association and Community



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

High-compliance community respondents were motivated by morals and deterrence.

High-compliance community respondents felt that tax reporting was a moral obligation (98 percent *vs.* 92 percent of those from low-compliance communities) and would be embarrassed if others found out they had under-reported (90 percent *vs.* 76 percent). Similarly, the high-compliance community respondents were risk-averse, more frequently agreeing that hearing about people who were caught underreporting would make them more careful (86 percent *vs.* 70 percent). Conversely, respondents from low-compliance communities evidently were not deterred despite their belief that the IRS probably knows when people underreport income (62 percent *vs.* 52 percent from the high-compliance communities). An inference could be made that deterrence efforts affect those predisposed to compliance.

---

[52] This difference is not statistically significant at a 95-percent level of confidence.

[53] This trend was generally consistent with that in the National Survey, except the frequency of voting among the low-compliance group was not higher than that of the high-compliance group.

[54] Additional comparisons, not statistically significant at a 95-percent level of confidence, were 55 percent vs. 47 percent and 46 percent vs. 36 percent for trade associations, and 42 percent vs. 15 percent and 37 percent vs. 9 percent for parents.

[55] If taxpayers from a low-compliance community feel that they have a support group in certain institutions—social noncompliance—then civic education addressing those institutions could leverage enforcement efforts. Civic education would mean not technical training on particular tax provisions, but "the cultivation of the virtues, knowledge, and skills necessary for political participation." Amy Gutmann, DEMOCRATIC EDUCATION (Princeton Univ. Press, 1987) 287. Maintaining a low level of tax compliance may be a form of political non-participation motivated by a skepticism of fairness in taxation—symbolic noncompliance—as described in the typology of noncompliance. Thus, popular dissemination of information about the institutions that ensure fairness, e.g., the checks and balances created by an independent judiciary and Congressional oversight, could be a responsive form of civic education.

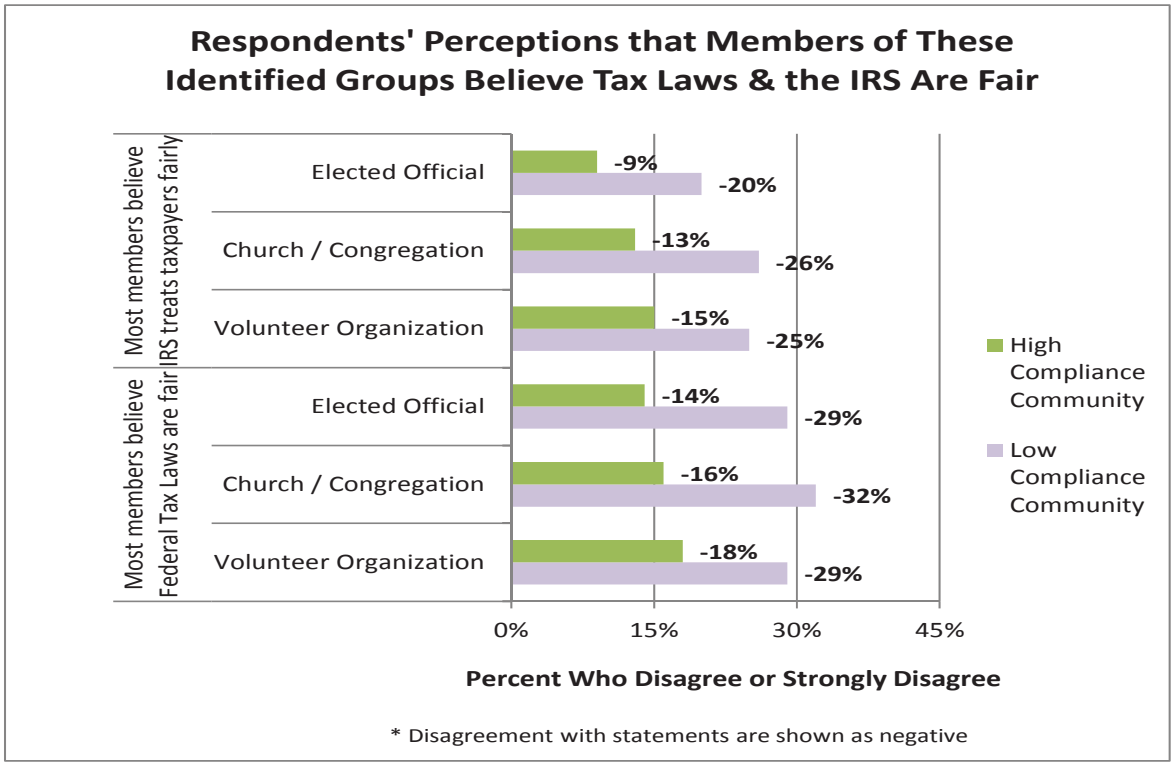
The effect on compliance of financial concerns by those in high- or low-compliance communities was unclear.

Paradoxically, respondents from high-compliance communities were more likely to feel that most small businesses could not survive if they reported all of their income (23 percent *vs.* 16 percent of low-compliance community respondents). However, the National Survey did not reproduce this result (12 percent *vs.* 15 percent of the low-compliance group).[56] In terms of the factors discussed above, it is unclear that deterrence motivated compliance. In terms of the typology introduced above, "asocial" noncompliance due to financial incentives did not appear as a major force in the Community Survey.

Those in the high- and low-compliance communities responded similarly to questions addressing complexity.

Both groups responded without significant difference to questions about how complicated the tax rules are (64 percent of the highly-compliant *vs.* 63 percent of low-compliance respondents) and the clarity of income reporting rules (73 percent *vs.* 68 percent). Consequently, the Community Survey did not reveal significant procedural, "lazy," or unknowing noncompliance.

FIGURE 16. Respondents' Perceptions that Members of These Identified Groups Believe Tax Laws & the IRS Are Fair



Except as otherwise noted, all of the differences between the two groups are statistically significant at the 95-percent level of confidence.

[56] This difference is not statistically significant at a 95-percent level of confidence.

Summary of the Community Survey Results

The Community Survey focused on areas that did not represent the nation as a whole but did represent locales where tax compliance levels were markedly low or high. The site selection process identified hundreds of low-compliance communities, indicating a social aspect of noncompliance consistent with questionnaire responses showing a high degree of social affiliations. On the other hand, there were so few high-compliance communities as to make them prototypical rather than typical. The high-compliance communities may have had unique experiences with government contributing to an ideal type of "social compliance" that could be the obverse of social noncompliance typology. Additionally, the high-compliance communities responded positively to morals and preparers as well as deterrence. By contrast, the low-compliance communities evidently were not deterred even though they believed that the IRS could detect under-reporting. The combination of risk tolerance and geographic concentration of low-compliance communities could form the basis for targeted innovation in tax administration that would go beyond deterrence toward the social and moral factors underlying compliance.

Preliminary Observations

As reflected in the 2007 Review, social norms and related factors may help explain tax compliance. As discussed above, TAS designed a survey questionnaire to probe into norms and related factors. While this survey elicited direct responses from taxpayers, the "social" nature of norms should be observable even beyond these responses, potentially by observing characteristics of the high- and low-compliance communities or regions. Future research could build upon the survey results by investigating social noncompliance and social compliance in sites where they occur.[57] While tax reporting may be a private decision, compliance levels appear to depend on values that are shared or at least commonly understood.[58]

Commonly understood values are social or geographic, and therefore susceptible to study through market research or public sources beyond the questionnaire. What is the "means of communicating these learned beliefs, memories, perceptions, traditions, and attitudes that serves to shape behavior"?[59] Not all behavior stems from local interaction.[60] Yet geographically-dispersed populations, such as diasporas with common origins in the past, or virtual communities on the Internet, may be exceptions that prove the rule.[61] Like tax administration as a whole, compliance research could advance by meeting taxpayers where they are, in geographic locations where they build communities around common behavior.

In this study, tax compliance has turned out to be "retail." Clusters of measurably similar compliance levels may lie in a cultural region. While individual predispositions like risk aversion are factors, they are expressed within regional norms.[62] Seemingly nongeographic behavior may exhibit regional effects that researchers have isolated using statistical techniques such as "regression analysis"—analysis used to understand how a "dependent variable" (*e.g.*, legal compliance) changes when any one of the "independent variables" (*e.g.*, location or other relevant factor) changes.[63]

---

[57] Whereas the survey method may be consistent with "[m]ost theories in social science today" which "are based on the assumption that individuals are atomistic and thus independent of one another," this assumption "leaves unresolved the problem of accounting for the order one finds in society." James Duncan, *The Superorganic in American Cultural Geography*, 70 ANNALS OF ASSOC'N OF AMER. GEOGRAPHERS 181, 183 (1980).

[58] "Culture is public, because meaning is." Clifford Geertz, *Thick Description: Toward an Interpretive Theory of Culture*, INTERPRETATION OF CULTURES (N.Y.: Basic Books, 1973) 12.

[59] Mona Domosh, Terry Jordan-Bychkov, et al. THE HUMAN MOSAIC: A THEMATIC INTRO. TO CULTURAL GEOGRAPHY 12th ed. (N.Y.: W.H. Freeman & Co., 2012).

[60] "As a cerebral entity, a culture may flourish, move about, and propagate itself solely within the heads of a number of footloose individuals. Such extreme cases do occur, of course, but normally the facts of location and the processes of interaction with other localized or spatially structured phenomena do matter greatly." Wilbur Zelinsky, CULTURAL GEOGRAPHY OF THE U.S. (Prentice Hall, 1973) 76.

[61] Even Internet use depends on users' geographic location. See, e.g., Eric Gilbert, Karrie Karahalios & Christian Sandvig, *Network in the Garden: An Empirical Analysis of Social Media in Rural Life*, Conf. on Computer-Human Interaction of Assoc'n for Computing Machinery, Florence (2008).

[62] Why would tax compliance, among other characteristics, be part of cultural geography? "Imagine someone who is, among other things, a Czech-American Lutheran plumber, a member of the VFW, an ardent Cleveland Indian fan, a radio ham, a regular patron of a particular bar, and a member of a car pool, the local draft board, the Book-of-the-Month Club, and the Republican party, and a parent whose son attends a particular college. Each of these subcultures will tend to have its own array of gear and physical arrangements, spectrum of economic and social beliefs and practices, cluster of abstract concepts, and, not least important for our purpose, distributional spread in physical space." Zelinsky, CULTURAL GEOGRAPHY at 74.

[63] "While differences in standard demographic or economic variables such as age composition, median education, or median income account for a good deal of the variance among sections of the country" in particular social statistics, "there is a significant remainder that may be related" solely to geographic characteristics. Raymond Gastil, CULTURAL REGIONS OF THE U.S. (Seattle: Univ. of Wash. Press, 1975) 116.

By identifying high- and low-compliance communities, this survey prepares the way for potential research in particular geographic areas. Relevant aspects of high compliance, such as trust in government or respect for authority, may occur throughout the country, even if not in a high-compliance community. Future research could go beyond the factors underlying compliance to test how tax administration may respond to regional traditions with certain services in particular geographic regions.[64]

Future research could also use more sophisticated tools to analyze the National Survey data. While this preliminary analysis identified important correlations between (estimated) tax compliance and responses to the survey questions, applying "regression analysis" (described above) to the data might provide further insight into which responses (or other observable factors) have the greatest effect on (estimated) tax compliance. For example, while this preliminary analysis reveals a correlation between estimated tax compliance and attitudes about the government, the law, and the IRS, a regression analysis might reveal the relative importance of these attitudes after controlling for the effect of other factors.

Alternatively, data mining techniques (*e.g.*, "segmentation" or "cluster" analysis) could identify groups of survey responses that are most frequently associated with each other. Such analysis might enable researchers to identify various distinct types of noncompliance. For example, this analysis might find a particular segment of the low-compliance group for whom complexity presents a barrier to compliance. It might distinguish this segment from another for whom complexity is not a barrier, but justifies noncompliance on the basis of negative views about the IRS, the law, and the government. Such analysis might help to inform policymakers about how to tailor an effective approach to address different types of noncompliance and different segments of the population.

Conclusion

In conclusion, the TAS survey has helped to identify which factors significantly influence compliance. In turn, analyses of the factors and related data point to the operative types of noncompliance. Finally, knowledge of these factors and types of noncompliance can inform service and enforcement programs.

The results of both surveys suggest that norms and distrust of the national government, the law, and the IRS may promote noncompliance. Respondents from both the low-compliance groups and from low-compliance communities held negative views about government and the IRS and were more likely to participate in local organizations. They were also more likely to believe that other members of those organizations held similarly negative views, which appeared to reinforce their own views, though they generally professed that noncompliance was morally wrong. In other words, they affiliated with others who reinforced noncompliance norms at the local level, and probably feel a closer connection to a local collective than to the national collective. In terms of the typology discussed above, this tendency to affiliate where distrust of government is the norm may be a form of social and symbolic noncompliance.

Consistently, the results also suggest that tax morale and trust in government, the law, the IRS, and preparers may promote compliance. Respondents from the high-compliance group and the high-compliance communities were less likely to participate in local organizations, suggesting that their compliance level was not social but motivated rather by morality, trust in government, trust in the IRS, or other internal factors contributing to high tax morale.

Those in both the high- and low-compliance groups also expressed a high level of trust in a preparer, but those in the high-compliance group expressed more trust. Similarly, a greater reliance on preparers by respondents from the high-compliance communities suggested a type of "brokered compliance." Thus, the survey

[64] See Most Serious Problem: *The IRS Is Substantially Reducing Both the Amount and Scope of Its Direct Education and Outreach to Taxpayers and Does Not Measure the Effectiveness of Its Remaining Outreach Activities, Thereby Risking Increased Noncompliance*, supra; National Taxpayer Advocate 2009 Annual Report to Congress 346 (Legislative Recommendation: *Strengthen the Independence of the IRS Office of Appeals and Require at Least One Appeals Officer and Settlement Officer in Each State)*; National Taxpayer Advocate 2007 Annual Report to Congress 162 (Most Serious Problem: *Service at Taxpayer Assistance Centers*); National Taxpayer Advocate 2003 Annual Report to Congress 145 (Most Serious Problem: *Taxpayer Assistance Centers*).

results suggest that "brokered compliance" may be a potential benefit of a well-regulated preparer profession, which the National Taxpayer Advocate has long championed.[65]

The survey results did not reveal as much about the effect of deterrence and complexity on reporting compliance. With respect to deterrence, this may have been because some respondents answered defensively—to avoid explicitly implicating themselves in noncompliance. With respect to complexity, the survey did not specifically identify procedural, "lazy," or unknowing noncompliance as posing major concerns. Nonetheless, complexity likely promotes the view that the government, the law, and the IRS are unfair or cannot be trusted, and the survey responses suggest that these views may reduce compliance.

Given the emergence of social and symbolic noncompliance as the primary types of noncompliance among small businesses, treatments that promote trust in government, the law, and the IRS may be most effective. As a practical matter, this might include tax simplification, an expansion of taxpayer protections and remedies, and taxpayer education.[66] This kind of education would be normative, relating to trust in government, rather than technical.[67] Traditional enforcement measures designed to deter could be ineffective, both because those likely to respond may be predisposed to comply and because the survey results did not suggest that asocial behavior (*i.e.*, behavior that may be addressed by increasing deterrence) is prevalent. [68]

[65] See 2010 Proposal at 81 (Table 2.4.1, *Typology of Noncompliance and Potentially Operative Factor(s) Identified by the Literature*); National Taxpayer Advocate 2008 Annual Report to Congress 423 (Legislative Recommendation: *The Time Has Come to Regulate Federal Tax Return Preparers*); National Taxpayer Advocate 2004 Annual Report to Congress 67 (Most Serious Problem: *Oversight of Unenrolled Return Preparers*); National Taxpayer Advocate 2003 Annual Report to Congress 270 (Legislative Recommendation: *Federal Tax Return Preparers Oversight and Compliance*); National Taxpayer Advocate 2002 Annual Report to Congress 216 (Legislative Recommendation: *Regulation of Federal Tax Return Preparers*).

[66] For a discussion of procedural protections that could improve trust in government, see National Taxpayer Advocate 2011 Annual Report to Congress 493–518 (Legislative Recommendation: *Enact the Recommendations of the National Taxpayer Advocate to Protect Taxpayer Rights*) and National Taxpayer Advocate 2007 Annual Report to Congress 478–489 (Legislative Recommendation: *Taxpayer Bill of Rights and De Minimis "Apology" Payments*). For a summary of the National Taxpayer Advocate's simplification proposals, see Complexity and the Tax Gap: Making Tax Compliance Easier and Collecting What's Due, *Hearing Before the S. Comm. on Finance* (June 28, 2011) (statement of Nina E. Olson, National Taxpayer Advocate).

[67] For a discussion of the types of education that might be effective, see 2007 Review at 162–170.

[68] On deterrence, TAS has initiated further research on the effect—if any—of audits on subsequent reporting compliance as measured by the DIF.

Transfer Pricing: Strategies, Practices, and Tax Minimization

Kenneth Klassen, University of Waterloo; Petro Lisowsky, University of Illinois of Urbana-Champaign; and Devan Mescall, University of Saskatchewan[1]

In cooperation with the Tax Executives Institute (TEI), we conducted an extensive field survey of 219 multinational company tax directors to provide a detailed account from within multinational corporations of differences in transfer pricing practices and strategies, and their role in tax minimization. Transfer pricing in a tax setting is a topic shared across accounting, economics, and law disciplines. Within this broad literature, the complexities of transfer pricing are often simplified while transfer prices themselves are seen as an important tool for multinational firms to reduce global taxes. For example, Hassett and Newmark (2008) state that one mechanism for income shifting is tax-motivated transfer pricing, defining this phenomenon as "the practice of multinational corporations of arranging intrafirm sales such that most of the profit is made in a low-tax country" (p. 208). Media accounts reinforce this perception and blame opportunistic transfer pricing for the low amount of taxes paid by large companies in countries such as the U.K., including specific mentions of Google, eBay, and Starbucks (Bergin 2012 and Milne 2012). However, broader views of tax strategy (e.g., Scholes and Wolfson 1992) suggest such one-dimensional analyses ignore other uses of the transfer prices, e.g., to support decentralization and coordination (Baldenius et al. 2004).

Due to the proprietary nature of transfer prices, documenting evidence of whether and how much transfer pricing facilitates tax avoidance is challenging. Researchers are generally left with data that provide only an external, indirect view of firms' operations and internal transactions (Gordon and Hines 2002; Jacob 1996). In addition, there are many challenges in identifying international tax planning from financial statements (see Donohoe et al. 2012), leaving researchers to use less conventional methods and data sources. For example, Blouin et al. (2012) use proprietary data from the Bureau of Economic Analysis (BEA) to examine transfer pricing within firms. We complement these studies by using our TEI survey participants to directly evaluate the links between transfer pricing and corporate tax minimization.

The goal of this study is to provide a detailed analysis of the role of tax departments in the setting of transfer prices. We undertake this analysis to inform tax policy setters, practitioners, advisors, and researchers studying this growing area of activity. Global tax authorities have raised concern about the loss of tax revenues that may be the result of aggressive transfer pricing practices (OECD 2013). In 2010, the U.S. Internal Revenue Service (IRS) announced a greater focus on transfer pricing to address potential revenue loss including plans to increase staffing by creating a new Director of Transfer Pricing position, which was filled in 2011 (Ossi and Shepherd 2010). The global professional services firm Ernst & Young (E&Y) began surveying tax authorities in 1995, and in their most recent survey of 48 tax authorities in 2012, they conclude:

> "One constant runs through the history of the survey: tax authorities continue to add staff devoted to transfer pricing. In a climate of budget freezes for many government agencies, tax authorities appear to have made the cost/benefit calculation to incur additional staffing costs in order to investigate transfer pricing. The penalty burden is also increasing." (Ernst & Young 2012, p. 7)

Given the tax enforcement environment, aggressively using transfer prices to reduce taxes may not be optimal for all companies. In fact, the same E&Y report also advises that "[t]axpayers should not be complacent about their transfer pricing risk" and recommends that "[c]ompanies should pursue tax certainty" (p. 5). This sentiment is echoed by Alvarez and Marsal Taxand's 2012 survey of chief financial officers who identify transfer

[1] University of Waterloo, University of Illinois at Urbana-Champaign, and University of Saskatchewan, respectively. The authors thank Tax Executives Institute (TEI) for distributing our survey and TEI members for their participation. Special thanks for the cooperation and support of Timothy McCormally and Mary Fahey. The authors thank Marc Alms, Andrew Bauer, Muris Dujsic, Christy MacDonald, Paula Moore, and Leslie Robinson for assistance in developing our survey instrument; workshop participants at the University of Iowa for helpful comments; and Stephen Powers for excellent research assistance. Finally, Kenneth Klassen acknowledges the generous support of the Robert Harding Research Leadership Fellowship at the University of Waterloo, and Petro Lisowsky acknowledges the generous support of the PricewaterhouseCoopers Faculty Fellowship at the University of Illinois.

pricing as the largest tax risk facing multinational companies. In the academic literature, Towery (2013) finds that the second-largest area of corporate income tax uncertainty as disclosed on the IRS's Schedule UTP in 2010 is related to international transfer pricing. Also, Mescall and Klassen (2013) find that increased tax risk related to strict transfer pricing enforcement adversely impacts cross-border merger and acquisition premia. Given the growing focus on minimizing tax risk, it is therefore unclear for which multinational corporations transfer pricing is a useful tool for tax minimization, and for which corporations the operational and enforcement costs are too great to risk implementing aggressive transfer pricing strategies.

We provide a rich, detailed, and direct account of transfer pricing for tax purposes, as reported by multinational corporations' tax departments. Our survey shows that a larger proportion of multinational firms assess their transfer pricing practices on *compliance-based* measures than on *tax minimization* measures, contrary to stereotypes on the (near) ubiquitous exploitation of transfer prices by multinational firms to reduce their tax burdens.[2] We also find that, among transfer pricing strategies, nonmanufacturing firms are more likely to adopt a goal of tax minimization than manufacturing firms. In addition, firms with less intense internationalization are more likely to pursue a goal of tax compliance as opposed to minimization. These results are consistent with the transfer pricing literature that uses indirect evidence. We also find that transfer pricing consumes, on average, a larger proportion of tax budgets in nonmanufacturing firms and firms with more intense internationalization. Similarly, firms that assess their transfer pricing success using a goal of tax minimization spend a higher percentage of their transfer pricing resources on tax planning compared to firms that assess their transfer pricing success using a goal of tax compliance.

Our data reveal that firms assessing transfer pricing success using the goal of tax minimization have more experienced personnel, more well-funded internal tax departments, and more resources devoted to tax planning. Further, the frequency of managers citing a goal of tax minimization increases when the firms' transfer pricing objectives focus on reducing cash taxes paid. In multivariate analyses, we estimate that focusing on cash taxes as a goal of transfer pricing reduces effective tax rates (ETRs) by approximately 3.7 percentage points. Yet firms that focus on tax compliance to assess the success of their transfer pricing practices report ETRs that are 5.4 percentage points higher than firms claiming neither goal.

Finally, our respondents report that over the past decade, transfer pricing has become a larger portion of firms' overall tax budgets. Over half of our respondents indicated that their transfer pricing functions are underfunded, although we find no evidence that compliance burdens have resulted in tax planning being inaccessible to smaller firms. Overall, our unique survey data allow us to refine our understanding of when transfer pricing results (and does not result) in tax reduction. We highlight that transfer pricing opportunities do not result in uniform tax minimization across all multinationals because firm-specific strategies and practices differ substantially across these firms.

In sum, our study is a first step towards providing new and direct evidence on the role of transfer pricing in tax minimization. The results should interest researchers, practitioners, and tax authorities interested in the growing area of transfer pricing implementation and compliance. Although our evidence shows that transfer pricing is a material tax minimization tool and that a significant number of firms are evaluating their transfer pricing success based on tax minimization, we also provide strong evidence that even more frequent are multinationals' goal to comply with international tax laws.

[2] Although it is possible that the respondents are more reluctant to suggest that their goals lean more toward tax minimization rather than tax compliance, our multivariate regression tests show that their perception of how they are measured is indeed strongly linked to the actual effective tax rates reported by these firms. Thus, to the extent the respondents distort how their transfer pricing practices are evaluated, it would only work against us finding significant results when testing their link to GAAP effective tax rates.

References

Alvarez, and Marsal, (2012). "CFO Matters: Tax Perspectives Survey Highlights." Available online at: http://www.alvarezandmarsal.com/en/about/action_matters/documents/CFOSurvey_FINAL_002.pdf.

Baldenius, T., N. Melumad, and S. Reichelstein. (2004). "Integrating Managerial and Tax Objectives in Transfer Pricing." *The Accounting Review* 79 (3): 591–615.

Bergin, T. (2012) "How Starbucks Avoids Paying Taxes in The U.K." National Post (October 17, 2012). Available online at: http://business.financialpost.com/2012/10/17/how-starbucks-avoids-paying-tax-in-the-u-k/.

Blouin, J., L. Robinson, and J. Seidman. (2012). "Conflicting Transfer Pricing Incentives and the Role of Coordination," Working paper.

Donohoe, M., G. McGill, and E. Outslay (2012). "Through a glass darkly: What can we learn about a U.S. multinational corporation's international operations from its financial statement disclosures?" *National Tax Journal*, 65 (4): 961–984.

Ernst & Young (E&Y). (2012). "2012 Global Transfer Pricing Tax Authority Survey" *EYGM Limited*. Available online at: http://www.ey.com/Publication/vwLUAssets/2012-Global-TP-Tax-Authority-Survey/$FILE/2012-Global-TP-Tax-Authority-Survey.pdf.

Gordon, R.H., and J.R. Hines, Jr. (2002). "International Taxation." In: *Handbook of Public Economics* vol. 4, A. Auerbach and M. Feldstein. (Eds.), Amsterdam: North-Holland.

Hassett, K., and K. Newmark. (2008). "Taxation and Business Behavior: A Review of the Recent Literature." In: *Fundamental Tax Reform: Issues, Choices and Implications*. Diamond, J., Zodrow, G. (Eds.), MIT Press, Cambridge: 191–214.

Jacob, John. (1996). "Taxes and Transfer Pricing: Income Shifting and the Volume of Intrafirm Transfers." *Journal of Accounting Research* 34 (2): 301–312.

Mescall, D. and K. Klassen. (2013). "How does transfer pricing risk affect premia in cross-border mergers and acquisitions?" Working paper.

Milne S. (2012). "A Roll Call of Corporate Rogues Who Are Milking The Country" *The Guardian*. Available online at: http://www.guardian.co.uk/commentisfree/2012/ oct/30/roll-call-corporate-rogues-tax.

OECD. (2013). *Addressing Base Erosion and Profit Shifting*, OECD Publishing. Available at: http://dx.doi.org/10.1787/9789264192744-en.

Ossi, G. and M. Shepherd. (2010). "The IRS's Emphasis on Transfer Pricing." *Corporate Taxation* 37 (4): 3–6.

Scholes, M., and M. Wolfson. (1992). *Taxes and Business Strategy: A Planning Approach*. Englewood Cliffs, NJ: Prentice-Hall.

Towery, E. (2013). "How Do Disclosures of Tax Aggressiveness to Tax Authorities Affect Reporting Decisions? Evidence from Schedule UTP" Working paper.

Demand for Aggressive Tax Planning[1]

Peter Bickers, Tracey Lloyd, Bhaskaran Nair, and Michael Slyuzberg,
Inland Revenue, New Zealand

1. Introduction

This research was undertaken to inform New Zealand Inland Revenue's current and future initiatives aimed at reducing Aggressive Tax Planning (ATP). An ATP scheme is generally understood to mean any scheme where the purpose or benefit of the scheme appears to be the reduction of taxable income or inflation of deductible expenditure, and the tax advantage sought is not clearly sanctioned by the tax laws. In other words, ATP involves those schemes that may follow the letter of the law but not its spirit (Nash & Pross, 2009).

ATP is a global phenomenon of huge importance to tax administrations. A handful of the largest ATP transactions can result in hundreds of millions of dollars in lost tax revenue. ATP can also be used to artificially increase financial entitlements, and to reduce income-related liabilities, such as accident insurance levies, child support payments, and student loan repayments.

A worrying consequence of ATP is its effect on the public's trust in the tax system. In particular, ATP creates the public perception that many 'rich' people pay far less than their share of tax.

Demand for ATP in New Zealand has changed considerably over the last two decades. The 1990s saw a lot of mass-marketed ATP schemes. However, the popularity of these schemes largely ended with the introduction of the deferred deduction rule and promoter penalties in the early 2000s.

In 2001, there was a marked increase in customized ATP arrangements when the highest personal tax rate rose to 39 percent, while the corporate and trustee rates were at 33 percent. Tax agents provided a range of structures to avoid the personal services attribution rules and to exploit the progressive marginal rate system. The result was a number of self-employed people having company and trust structures set up, often with family members involved, to split income and limit exposure to the 39 percent rate.

It is envisaged that the recent flattening of the marginal tax rates is likely to reduce demand for ATP. However, there will still be arrangements designed to avoid reporting income or to produce deductions or losses for tax purposes. What we can deduce from the last 20 years or so is that taxpayers' behaviour is influenced by perceived inequities in the tax system, and the opportunity or incentive to engage in ATP activity.

Inland Revenue has recently won a number of tax avoidance cases, including some high-profile wins in the Supreme Court. This has changed the landscape markedly and the business community is now more cautious about tax planning.

2. Objective of this Research

Inland Revenue's Compliance Planning and Development Group, jointly with the ATP Portfolio (Investigations and Advice), has embarked on a strategic approach to ATP involving understanding demand and supply of ATP. This will enable Inland Revenue to adapt risk management strategies and identify successful legislative and administrative responses to ATP.

The aim of this research was to investigate the drivers behind customers' demand for ATP.

[1] We gratefully acknowledge the research fieldwork, analysis, and reporting of Litmus, an independent market research company, that conducted the in-depth interviews with tax agents and small and medium enterprises owners.

3. Methodology

A mixed method approach was used to understand ATP behaviour and attitudes among Individual and Business customers. The quantitative analysis incorporated profiling and two distinct segmentation methodologies in order to fully explore customer characteristics. In addition, qualitative interviews were conducted to provide direction for the quantitative profiling, and to gather perceptions directly from ATP customers, tax agents, and Inland Revenue staff. The compositions of the samples are shown in Table 1.

TABLE 1. Composition of Qualitative Interviews with ATP Risk Customers, Tax Agents and IR Staff

Composition of sample			
	Customers (n=22)	Tax agents (n=24)	IR staff (n=8)
Main distinguishing feature	Income situation* Established entrepreneurs (multiple businesses and investments) (6) Starting out entrepreneurs (multiple businesses, often change in business) (5) Long-term small businesses (professionals, tradespeople and farmers) (7) Long-term sole traders (professionals and tradespeople) (4)	Business situation In an accountancy firm (16) Sole trader (8)	Experience in private accountancy firm prior to working at IR Previously worked in an accountancy firm (7) Not previously worked in an accountancy firm (1)
Locations	Auckland (14) Wellington (4) Palmerston North (4)	Auckland (12) Wellington (6) Palmerston North (6)	Auckland (1) Wellington (3) Hamilton (3) Whangarei (1)

*All customers interviewed had complex income situations. For instance, all had investments in property and/or shares. Some (e.g., some new entrepreneurs and small business owners) were also moving in and out of paid employment as a result of fluctuations in their career and business success.

3.1 Profiling and Segmentation

The profiles of customers at risk of ATP[2] were compared with wider, general population profiles in order to assess the characteristics that discriminate between ATP risk and non-ATP risk customers. Population parameters were derived from a number of sources, including Inland Revenue's administrative data and Statistics New Zealand, for demographic profiles and industry classification. Customer segmentation was conducted on the data, and three clusters emerged: 'ATP-risk,' 'Opportunist,' and 'non-ATP risk.'

We also carried out in-depth, qualitative interviews with:

- 22 Inland Revenue customers from an 'at-risk-of-ATP sample list;'
- 24 tax agents with clients on the ATP risk sample list; and
- Eight Inland Revenue staff with in-depth understanding of the relationship between customers and their tax agents.

The interviews were up to one hour in duration, and were conducted from 20 April to 14 June 2012.

3.2 Software

A package called 'poLCA', implemented in the Inland Revenue environment, was used to estimate the latent class regression model for polytomous outcome variables. The latent class regression model enables us to estimate the effects of covariates on predicting latent class membership.

For the canonical discriminant analysis, the SAS procedure, 'Proc Disc,' was used to develop classification

[2] 'Customers at risk of ATP' in this study are 'customers who Inland Revenue suspects may be involved in ATP.' The criteria for this ATP risk group were that: (i) as a result of an initial check by Inland Revenue investigations staff, they had been identified as having indications of potential ATP risk that warranted an audit, and (ii) they had been audited by Inland Revenue sometime between 2001 and 2010, and a discrepancy was found.

criteria to assign group membership. The classification criteria in Proc Disc takes into account the prior probabilities of groups as a linear function.

3.3 Limitations

This analysis covers only two customer groups: Small and Medium Enterprises (SME) and High Wealth Individuals (HWI). Other customer groups, such as Large Enterprises and Not for Profit Organisations, are out of scope for this exercise.

This research relies on ATP risk categorisation provided by the Investigations and Advice Group and did not re-categorise the supplied ATP risk cases. As stated in the methodology, the ATP risk group in this research was based on customers '*suspected* of ATP' rather than *proven* as ATP.

4. Key Findings

This section synthesises the findings from the four data collection methods used for this project. There were three quantitative analysis methodologies to profile and segment ATP risk customers, and there were in-depth interviews with Individual and Business customers at risk of ATP, tax agents, and Inland Revenue staff. The findings cover customer demographics and discriminating characteristics related to ATP, perceptions of tax and ATP, and the relationship between tax agents and customers regarding ATP.

TABLE 2. Individual ATP Risk Customer Profile Comparison

Characteristic	Approximate Percentage of Individual Populations	
	ATP risk customers	Wider population
Male	80%	50%
Over 45 years old	90%	43%
Live in Auckland	60%	35%
Live in Palmerston North	11%	5%
Self-employed	64%	8%
Salary/wage earners	8%	82%
Professional/ technical	20%	7%
Financial/insurance service	20%	3%
Rental/hiring/real estate	19%	1%
Construction	1%	6%
Filing compliance rate[a]	50%	60%
Payment compliance rate[a]	60%	85%
Linked to a tax agent	94%	24%
In tax debt	9%	16%
Had tax debt of over $10,000	5%	1%
Income over $60,000	63%	10%
Paid donations	30%	11%
Paid donations under $1,000	9%	4%
Had tax credit claims over $1,000	68%	22%
Had median expenses over $10,000[b]	24%	1%
Had audit discrepancy higher than $25,000[c]	63%	24%
Had assessed shortfall penalties of over $10,000[d]	30%	5%

[a]This relates to customers' on-time filing and payment compliance between 2001 and 2010.

[b]This relates to the period from 2001 to 2010.

[c]This relates to the audited population between 2007 and 2011.

[d] This refers to the audited population between 2007 and 2011.

4.1 Profiling Individual and Business Customers at Risk of ATP

Profiles were produced for Individual and Business customers at risk of ATP and compared with profiles of the wider customer population. The following tables show the major differences in characteristics discovered for each profile.

TABLE 3. Business ATP Risk Customer Profile Comparison

Category	Approximate Percentage of Business Customers	
	ATP risk customers	Wider population
Operating for 10 to 20 years[a]	67%	29%
Operating in Auckland	60%	34%
Have been struck off	15%	3%
In liquidation	5%	< 1%
Close companies	40%	35%
Loss attributing qualifying companies (LAQCs)	28%	1%
Consolidated companies	8%	< 1%
Finance/insurance services	22%	22%
Rental/hiring/real estate	18%	1%
Professional/scientific/technical	17%	7%
Filing compliance rate[b]	66%	72%
Payment compliance rate[b]	82%	84%
Linkage to tax agent	89%	61%
Had business turnover > $500,000	29%	3%
Had liabilities > $500,000	34%	13%
Had tax debt > $10,000	9%	3%
Had a median profit > $60,000[c]	30%	6%
Had a median profit > $500,000[c]	11%	<1%
Had expenses > $250,000[d]	32%	22%
Travel expenses claims > $10,000	10%	4%
Audit discrepancy amount > $100,000[e]	48%	15%
Had assessed shortfall penalty[f]	21%	13%
Had assessed shortfall penalty > $10,000[f]	13%	4%

[a]Organisation commencement date is used for calculating business age.

[b]Customers' overall compliance behaviour in the period from 2001 to 2010 was used to derive filing and payment compliance.

[c]Median profit pertains to the period from 2001 to 2010.

[d]This relates to the period from 2001 to 2010.

[e]This relates to the audited population between 2007 and 2011. They were compared with ATP cases for the same period.

[f]This refers to the audited population between 2007 and 2011.

In the qualitative interviews, Inland Revenue staff had very consistent perceptions of ATP risk customers that strongly matched what was found in the above profiling. For instance, Inland Revenue staff described the ATP risk customers they had encountered as most commonly male, aged 30-55 years, high-income earners, and Auckland-based.

Further, Inland Revenue staff described Business ATP risk customers as likely to be: in industries such as property development and car dealerships; self-employed; multinational companies; starting or growing a business; changing the structure of a business; and in complex income situations (e.g., with multiple income sources and/or complex expense calculations).

4.2 Tax Agent Role

The role of tax agents in ATP has received considerable attention, as they are in a position to act either as an exploiter or an enforcer of tax law. In the quantitative analysis, we found that 94 percent of Individual ATP risk customers were linked to tax agents, and 89 percent of Business ATP risk customers were linked to tax agents.[3]

Interestingly, ATP risk customers with tax agents were more likely to be located in Auckland, Wellington, and Palmerston North. Fifty-four percent were in Auckland, 18 percent were in Wellington, and 10 percent were in Palmerston North, compared to 33, 10, and 6 percent respectively for the wider population. This finding is directly comparable to the finding that Individual and Business ATP risk customers are also more likely to be in Auckland, Wellington, and Palmerston North.[4]

Tax agents with clients in the ATP risk group generally had around 1 to 5 ATP risk Individual customers (amongst their total client base of perhaps 100+) and 1 to 5 ATP risk Business customers as clients.

In the qualitative interviews, ATP risk customers, tax agents, and Inland Revenue staff all talked at length about the influence of tax agents on how customers approach ATP risk. Customers expect tax agents to advise them on any risks associated with their tax affairs, including any likelihood of their tax planning actions being unacceptable to Inland Revenue. In some cases, this involves agents presenting their clients with two or three tax planning options, and the tax risks associated with each. Unfortunately, tax agents themselves sometimes find it difficult to be sure about the 'grey areas' of tax, and how Inland Revenue will judge different tax planning options.

> *"The current tax law is over the heads of most clients and also for many accountants."*
> *(Tax agent)*

The participants' descriptions in qualitative research of the ATP customer-tax agent relationships they had witnessed indicated there were two clear dimensions to those relationships:

1. Whether it is the customer or the agent who is the catalyst for the ATP activity; and
2. The level of equality between the customer and agent (for example, mutual trust and a free flow of information between them).

These two dimensions can be used to describe four possible types of ATP customer-tax agent relationships. The views expressed in the qualitative interviews illustrated all four of these relationship types.[5]

The four relationship types are as follows:

Type 1. Customer-prompted ATP, unequal relationship.

The customer is wanting to engage in ATP, but may not be willing to share ATP-relevant information with the agent or Inland Revenue. In addition, the relationship between customer and agent can become hostile if tax disputes with Inland Revenue arise.

> *"They don't want to reveal information to the agent, won't give the facts, not transparent."*
> *(Inland Revenue staff)*

[3] It should be noted that many of the businesses not linked to a tax agent would have internal employees who are tax professionals.

[4] See findings for 'Region' in the Profiling section.

[5] Further research would be required to confirm these suggested typologies.

Type 2. Customer-prompted ATP, equal relationship.

The ATP is prompted by the customer and the agent supplies the necessary tax expertise. The two parties work together with a mutual understanding and benefit.

> *"[The customer] wanted 10 percent off the tax bill, they didn't care how."*
> *(Inland Revenue staff)*[6]

Type 3. Agent-prompted ATP, unequal relationship.

In this example, the agent prompts the ATP and runs the tax planning with little input from the customer. When issues arise, the customer feels it is unfair that they are targeted by Inland Revenue as they do not believe they were fully informed about the compliance risks by their tax agent. This opinion was shared by a large proportion of the customers interviewed in this research.

> *"Inland Revenue defines this as tax avoidance and thinks we should know better, but we trusted our accountant."*
> *(Customer)*

Type 4. Agent-prompted ATP, equal relationship.

In this relationship, the ATP risk activity is prompted by the agent as a way to gain or keep the customer as a client. The customer is aware of the tax risks, but expects the agent to steer the way and win any tax interpretation arguments with Inland Revenue.

> *"Agents consider avoidance and come up with two or three commercial reasons and say 'we should get over the line'."*
> *(Inland Revenue staff)*

In some cases, Inland Revenue staff believe the customer is fully aware that it's 'too good to be true,' but it can be difficult to confront these customers directly as the agent often interposes himself/herself if Inland Revenue investigates the case.

4.3 Segmentation

Customer segmentation was used to find the characteristics that define ATP risk customers and non-ATP risk customers. A dataset consisting of a mixture of ATP risk customers and non-ATP risk customers was segmented into discrete customer groups that share similar characteristics. Individual and Business customers were segmented separately. Three clear clusters are visible in both the Individual customer segmentation (Figure 1), and the Business customer segmentation (Figure 2), with the Business customer segmentation being the most clearly defined. The three clusters in each case were:

- cluster 1: 'Non-ATP risk' customers
- cluster 2: 'Opportunists'
- cluster 3: 'ATP risk' customers.

The 'Opportunists' cluster was so named to emphasise its similarity to the 'ATP risk' cluster in both segmentations.

Classification error calculations[7] showed that the model tested was robust. For Individuals, 96 percent of ATP risk cases were correctly classified as ATP (exclusively comprised of ATP risk cases), and for Business customers 72 percent were correctly classified as ATP (see Table 4).

[6] In this example, the Inland Revenue staff member was talking about an experience they had when working in a private tax agency as a tax agent.

[7] Classification error rate indicates the efficiency of the model in correctly classifying cases. This is estimated by applying the model results to a new dataset that was not part of the data used in the model.

Reading Figures 1 and 2:

—Canonical variable 1: Best linear combination of independent variables.

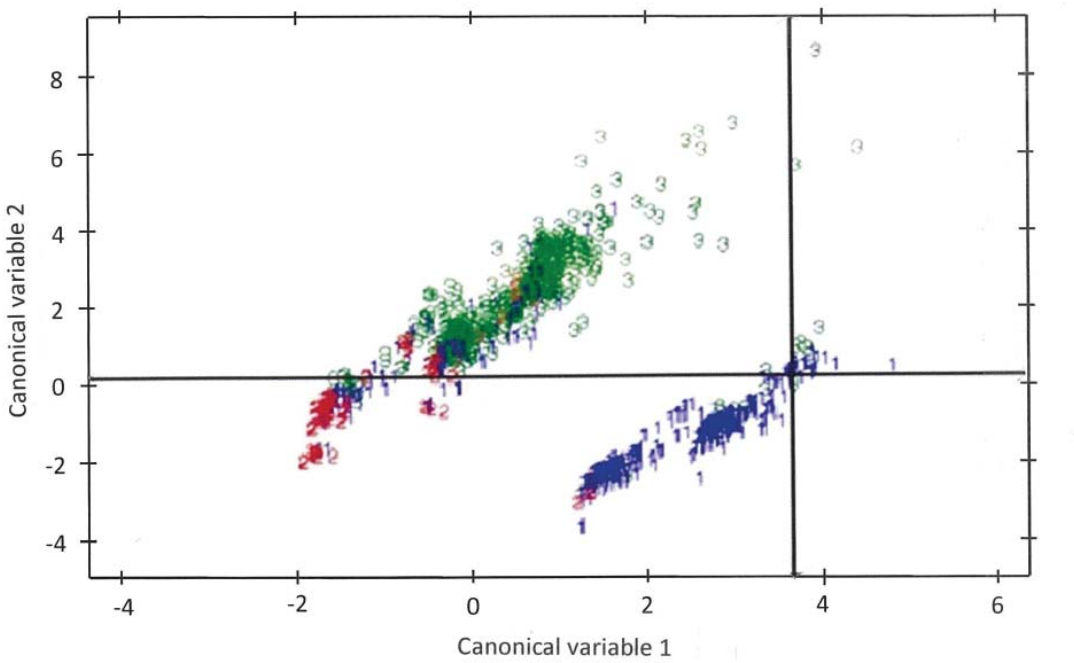
—Canonical variable 2: Next best linear combination of variables independent of canonical variable 1.

—The horizontal and vertical lines shown in the graph help identify the optimal cut-off points in the canonical variables that separate the groups.

The scatter plots and the error rate calculations show that the model is efficient in predicting ATP risk and non-ATP risk cases, although as expected with this form of modelling, some cases are misclassified. This model can therefore be used to identify potential ATP risk customers from among the population as long as there is no assumption that the prediction is 100 percent accurate.

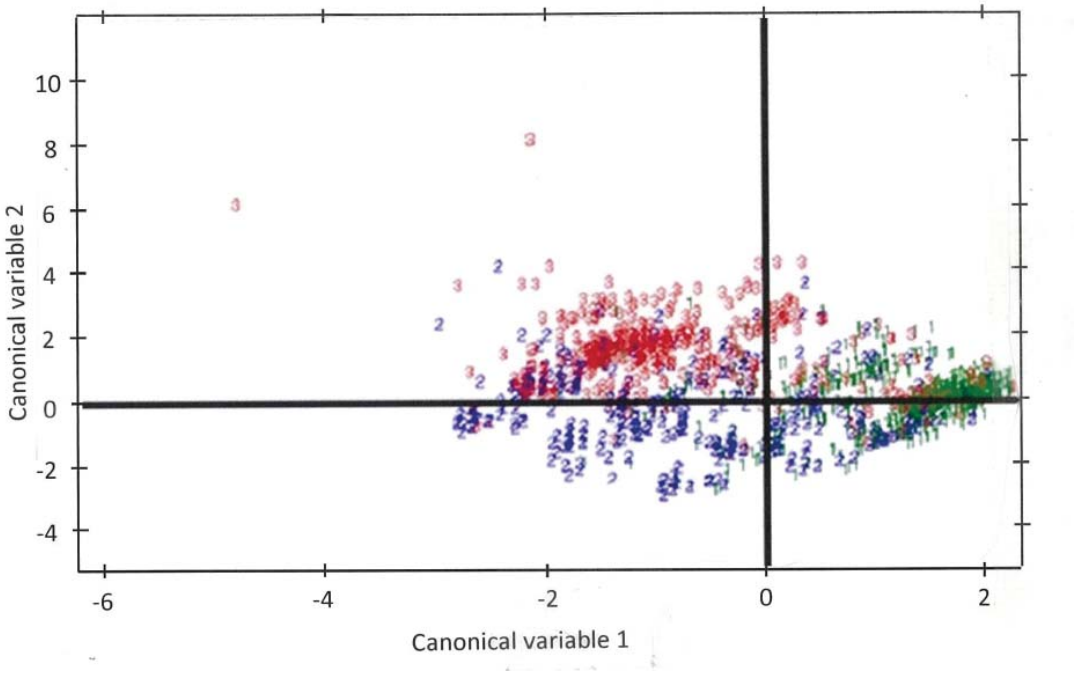
The qualitative research echoed the above quantitative segmentations as it explored perceptions of ATP. It showed that people describe a full spectrum of tax planning from universally acceptable tax planning, through to tax planning that can potentially be viewed as overly aggressive, and tax planning that most customers and tax agents would agree is unacceptable (see Figure 3).

FIGURE 1. ATP Risk Customer Segmentation—Individual Customers



Key: The codes in the graph refer to the clusters as follows; 1 is the 'Non-ATP risk' cluster, 2 is the 'Opportunist' cluster, and **3** is the 'ATP-risk' cluster.

FIGURE 2. ATP Risk Customer Segmentation—Business Customers



Key: The codes in the graph refer to the clusters as follows; 1 is the 'Non-ATP risk' cluster, 2 is the 'Opportunist' cluster, and **3** is the 'ATP-risk' cluster.

TABLE 4. Classification Rates for Individual and Business Customers

Percentage correctly classified		
	Individual customers	Business customers
ATP (Cluster 3)	96%	72%
Opportunists (Cluster 2)	79%	77%
Non-ATP (Cluster 1)	97%	91%
Error rate	*(10%)*	*(19%)*[8]

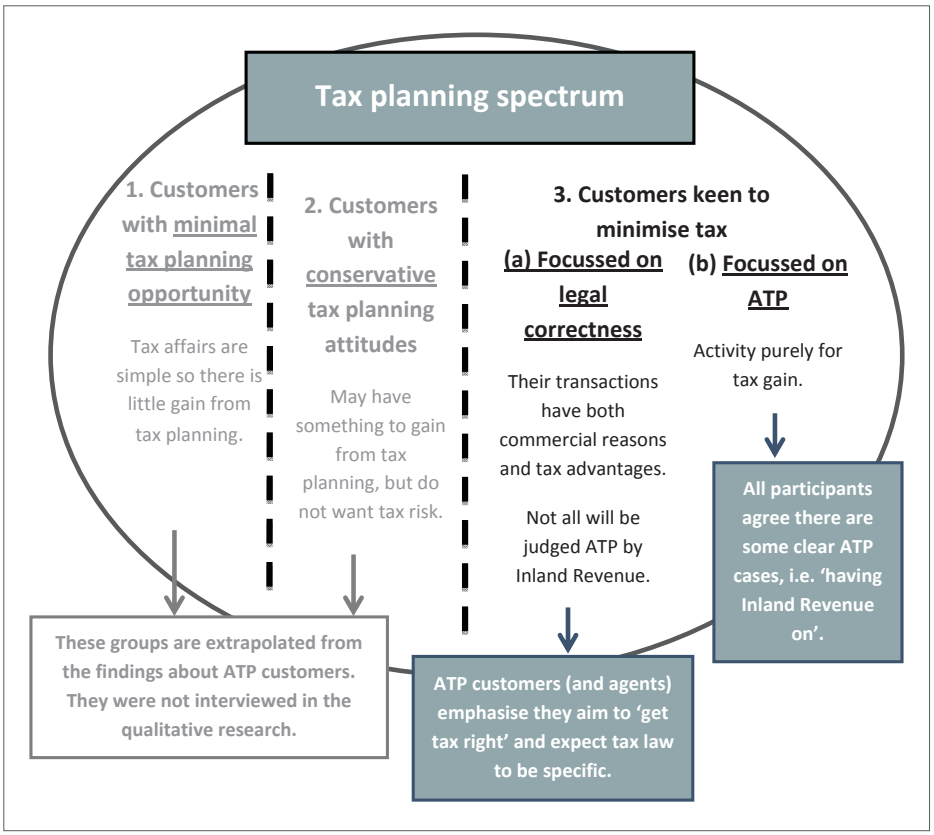
The core aspect of this qualitative categorisation is how customers decide what is 'acceptable' and 'unacceptable' tax planning. The customers and agents interviewed[9] felt that their tax planning activity was legal and a necessary part of business success (hence part 3a of the above spectrum).

> *"I don't believe I should be paying more tax than I have to. I have successful businesses that pay a lot of tax. Why should I pay more?"*
> *(Customer)*

[8] While there is more classification error for Business customers than Individual customers, the error rate of 19% is acceptable given the complexity of the data used in the analysis.

[9] It is important to note that the customers and agents interviewed in this research were drawn from a sample of customers which Inland Revenue had identified as 'at risk of ATP,' but not proven to be engaged in ATP.

FIGURE 3. The Tax Planning Spectrum



Nevertheless, the ATP customers and tax agents interviewed also accepted that 'some people' take tax planning too far. Some participants relayed personal experiences of other people who they felt were clearly overly aggressive in their tax planning (labelled 3b in the above spectrum).

> *"I see wealthy acquaintances that keep making losses but continue claiming these losses against their income so they don't have to pay tax. It's a deliberate system to avoid paying tax motivated by having more money."*
> *(Customer)*

The motivations described in this spectrum may be seen as the likely motivations underpinning the three clusters (ATP, Opportunists, and non-ATP) in the segmentation graphs (Figures 1 and 2). The 'Non-ATP' cluster would align with both parts 1 and 2 of the spectrum diagram. The 'Opportunist' and 'ATP' clusters are closely placed in the segmentation, and hence are difficult to distinguish without further research. Nevertheless, these two clusters would be expected to align with parts 3a and 3b of the spectrum diagram.

4.4 Discriminators of ATP risk

Before looking at the quantitative findings for the discriminators of ATP, it is useful to first look at the qualitative findings for how ATP risk customers view tax. These show that, in general, ATP risk customers are actually fully supportive of paying tax as a social responsibility.

> *"It's essentially not our money, it's the government's money."*
> *(Customer)*

However, ATP risk customers (and their agents) are less agreeable to the concept of being obliged to meet the 'intent' of the tax law, and place far more emphasis on meeting the 'letter' of the law.

> *"Every New Zealander uses the roads, schools and hospitals and it is our obligation to promote the value of people paying their fair share, but no more, within the law."*
> *(Tax agent)*

Further, the Inland Revenue staff interviewed stated that the most aggressive tax planners they have dealt with showed no interest even on the first aspect of 'social responsibility to pay tax.'

> *"It's a sense of entitlement. They don't expect to pay tax on a $10M profit, and see that as normal."*
> *(Inland Revenue staff)*

Along with the above views of tax, it is also valuable to consider the customers' context of operating a business. Business ATP risk customers placed a strong emphasis on growth of the business. However, some Inland Revenue staff believe that customers are sometimes led by their focus on growth into overly aggressive tax planning.

> *"The entrepreneurial type, they have two or three businesses that didn't take off. They want to use those losses [for ATP] …. they play around with shareholding and intellectual property."*
> *(Inland Revenue staff)*

The quantitative analysis looked for the key discriminators (or drivers) of ATP risk. The discriminators for the Individual and Business customer segmentations show certain similarities, particularly administrative compliance, the role of tax agents, and liabilities (see Table 5).

TABLE 5. Contribution of Variables into Cluster Class: Individuals and Businesses

	Discriminators		
	Primary Discriminators	Secondary Discriminators	Minor Discriminators
Individual customers	- Administrative compliance[10] - Linked to tax agents	- Annual income after expenses - Individual age - Liabilities[11] - Professional services	- Real estate and financial services - Tax credit claims - Donations paid
Business customers	- Administrative compliance - Business growth[12] - Linked to tax agents	- Business age - Liabilities	- Travel expenses claim - Auckland - Finance and insurance services - Annual turnover

The strongest discriminator of ATP risk for both Individual and Business customers is administrative compliance. Individual customers at risk of ATP have *markedly* lower administrative compliance than non-ATP risk Individual customers. Similarly, Business customers at risk of ATP have *slightly* lower administrative compliance than non-ATP risk Business customers.

Interestingly, there is a dual finding regarding administrative compliance. The quantitative analysis showed that lower administrative compliance is a discriminator of ATP risk. However, the qualitative interviews indicated that some 'smart' ATP risk customers maintain a high level of administrative compliance with the apparent intention of hiding their ATP activity.

[10] Administrative compliance refers to on-time filing and payment of tax.

[11] Liabilities emerged as a secondary driver in the modelling stage but was not notable in the profiling of Individual ATP customers (i.e., 2% of Individual ATP customers had liabilities of over $500,000, which was not clearly distinguishable from the 1.3% of the wider population who had liabilities of over $500,000).

[12] Business growth emerged as a primary driver for Business ATP customers at the modelling stage but was not investigated at the earlier profiling stage of this study. We analysed business growth for Business ATP customers only, not for the wider Business customer population. In this calculation, growth is calculated as change in annual turnover from one year to the next and relates to the period from 2001 to 2010. For Business ATP customers: 37% had 'low' growth (<30% growth), 10% had 'medium' growth (30% to <60% growth) and 7.5% had 'high' growth (60% growth and above). The rest had either no growth or the information was not available.

> *"They do anything to avoid tax, but they do file their taxes and look like complying."*
> *(Inland Revenue staff)*

This latter finding means that there will be some challenge in how administrative compliance can be used as an indicator of ATP risk.

The use of tax agents is another primary discriminator for both customer types, particularly so for Individual ATP risk customers. The tax agents in question were more likely to be those with higher client density (more than 100 clients).

Another result that emphasised the role of tax agents is that, for Individual ATP risk customers, 70 percent of ATP risk customers using tax agents had audit discrepancies higher than $25,000, whereas 50 percent of ATP risk customers with no tax agents had discrepancies in that range.

For ATP risk Business customers, approximately 88 percent of those who used tax agents had audit discrepancies higher than $25,000, whereas 63 percent of those without tax agents had discrepancies in that range.

And finally, as part of considering the discriminators of ATP, it is notable that there was a strong view that ATP activity in New Zealand has been declining in the last 10 years.

> *"We used to boast about how many companies we had and how many trusts, and now we boast about how few. In the past, the accountant used to structure everything so it was tax effective and I ended up with so many company structures and accountancy fees I'd have to phone the accountant to find out which company was which. People are sick of paying fees and a lot of loopholes have been shut down by Inland Revenue."*
> *(Customer)*

This reduction in ATP activity was seen as due to people's expectations of financial gain from ATP being greatly reduced by three environmental factors:

1. New Zealand's reduced tax rates, plus the reduction in the gap between New Zealand's top personal tax rate and the company tax rate;
2. Inland Revenue's tougher approach alongside stronger ATP legislation that is more 'generalised' (relying less on specific examples); and
3. The global economic crisis (which has both positive and negative effects on ATP).

Regarding New Zealand's tax rates, the general view was that lowering the rates has markedly reduced ATP activity.

> *"When the personal tax rate was 60 percent, everyone thought it was legalised robbery, so it became a big incentive for people to look for a way to pay less tax."*
> *(Customer)*

Nevertheless, some Inland Revenue staff believe that some customers will maintain their ATP even when the expected gains are small.

> *"When the Individual's tax rate was 39 percent, the natural tendency was to go to a Company or Trust for 33 percent."*
> *(Inland Revenue staff)*

Regarding Inland Revenue's tougher approach, the general view amongst Inland Revenue staff is that the implementation of penalties for the promotion of ATP was vitally important, but that an overall increase in the intensity of Inland Revenue's approach to ATP was also effective.

> *"Inland Revenue is smarter now, taking it on, more aggressive, going back to things from four years ago."*
> *(Inland Revenue staff)*

Alongside this, there is a view that, with New Zealand's anti-avoidance tax legislation becoming more generalised in recent times, Inland Revenue has been far more successful in challenging ATP risk activity. This

shift away from having to write specific legislation for every existing and future type of ATP activity towards generalised legislation is seen as an effective deterrent of ATP.

Nevertheless, ATP risk customers and their agents, and some Inland Revenue staff, also feel that the shift has left customers vulnerable to being penalised by Inland Revenue for only 'moderate' tax planning actions.

> *"There has been a shift in recent case law by Inland Revenue, almost taking the view that everything is avoidance unless you can prove otherwise."*
> *(Tax agent)*

And regarding the third environmental change, the global economic crisis, some participants believe that people's appetite for business-related risk in general has been reduced.

> *"I'm probably more risk averse and have been since the crash in 1987. Back in those days, I thought I could walk on water."*
> *(Customer)*

Nevertheless, some Inland Revenue staff believe that there may be some customers whose appetite for ATP has been increased as a result of their financial losses.

5. Discussion

Tax planning is an important part of managing finances for Business customers and Individual customers with complex financial income or expense situations. Legislating for all the tax situations that currently exist and will exist in future is a monumental task, and needs to take into account customers and tax agents as well as the tax authority. This needs to include awareness of customers' and agents' 'honest' intentions and errors, versus 'taking advantage' of grey areas of legislation. Getting this right will serve customers and tax authorities alike as certainty is a valuable cost-saving goal for all concerned. It will also help maintain the integrity of the tax system in the view of the wider population.

5.1 Spirit of the Law

The research has highlighted that ATP risk customers generally expect their tax planning behaviour to be judged according to written tax legislation. The concept of following the 'spirit' of the tax legislation is contrary to this expectation.

It will certainly be important for Inland Revenue to continue working to make customers and tax agents aware of the need to follow the spirit of the law. One inherent difficulty will be that customers and agents will naturally ask for examples to illustrate what is meant by the spirit of the law. There may be downsides to providing such examples as they may add support to the perceived need for concrete examples in tax legislation rather than broader conceptual thinking such as 'spirit of the law.'

Nevertheless, it is pleasing to note that even ATP risk customers see value in tax authorities addressing certain ATP cases, albeit at the most extreme end of the tax planning spectrum. In fact, some participants made extremely positive statements about the importance of tax. What is different for them is that they have opportunities for tax planning that the wider population does not have, and it is unsurprising that, for them, tax can be both a social responsibility and a 'cost to be minimised.'

However, this research has indicated that some customers are in no way interested in the social responsibility of tax. There are a number of theories on compliance that state an 'illegitimate profit seeker' weighs up their expected additional earnings against their expected legal penalties when deciding how aggressive to be in their tax planning (Grasmick & Green, 1980; Kagan & Scholz, 1984, p. 69; Kirchler & Maciejovsky, 2001; Murphy, 2002). This implies that ATP risk customers are less motivated to pay tax than non-ATP risk customers. However, this requires more study.

5.2 Strong Findings for Characteristics

This research has highlighted that ATP risk customers are clearly distinguishable; along with a shadow group of customers with similar characteristics (we have called these 'Opportunists'). The quantitative profiling of ATP risk customers produced some remarkably strong findings. The differences for Individual customers in characteristics such as gender, location, age, entity class, occupation, income and expenses are obvious, as are the differences for Business customers in characteristics such as age of business, location, entity class, nature of business, turnover and expenses. Inland Revenue staff also identified these characteristics on personal experience only.

Due to the strength of these findings about characteristics, Inland Revenue will be better able to discriminate between high and low ATP risk customers, which means better targeting and tailoring of ATP initiatives. In addition, the similarity of the 'ATP' and 'Opportunist' clusters means that there is value in targeting 'Opportunists' as well as ATP risk customers, as they also appear to have an elevated risk of ATP.

However, it is always worthwhile noting that there are still ATP cases that do not match the above characteristics, and there will no doubt be non-ATP cases that do match these characteristics.

5.3 Agents' Role and ATP Demand

In this research, we were specifically looking at customer demand for ATP rather than supply from the tax agent. However, we were aware that the closeness of the agent-client relationship would make it difficult to unravel the demand and supply link, hence the inclusion of tax agent interviews in this research. As expected, the demand story centres on the nature of the agent-customer relationship. This was emphasised even further by the finding that ATP risk customers are more likely than the wider population to be linked to a tax agent.

This latter finding cannot be construed as indicating agent linkage is causal for ATP. Indeed, this research has shown that the appetite for ATP risk can just as easily come from customers as from tax agents, or it may be a combination. Many ATP studies have highlighted the risk nexus between customers and tax agents (Murphy & Byng, 2002). A key aspect of this relationship is their appetite for tax risk. Sakurai & Braithwaite (2001) discuss how customers who are open to low-risk tax minimisation strategies often find themselves with tax agents who serve taxpayers who are open to high-risk minimisation strategies.

One aspect of the customer-tax agent relationship that is problematic for any ATP initiatives is that a large number of customers put considerable onus on their tax agent to deal with their tax matters, including managing their risk relating to ATP. This means that customers will not always feel responsible for their ATP risk. Agents also will not appreciate being targeted with ATP initiatives when the responsibility for compliance does not rest with them. Therefore, any initiatives addressing ATP demand will need to focus on the decision-makers in ATP activities, and take into account the varying levels of communication, equality, and responsibility that comprise agent-client relationships.

6. Conclusion

This study increased our understanding of customer demand for ATP by (i) showing that high-and low-risk ATP customers can be identified using Inland Revenue's administrative data, and (ii) describing their characteristics and underlying motivations regarding ATP.

6.1 Profiling ATP Customers

These results show that customers at risk of ATP can be profiled using their primary demographic and financial characteristics. Individual and Business customers at risk of ATP have separate profiles as they have some different characteristics, but there are similarities such as in the use of tax agents, nature of business and location.

ATP profiling can provide Inland Revenue with a tool for identifying future cases of customers being at risk of ATP.

Both the qualitative and quantitative elements of the research found that tax agents are a major influence on ATP risk customers. Of most concern is that agents are described as driving or prompting ATP risk activity with some customers. Further study of this interaction of 'supply and demand' for ATP is required.

6.2 Segmenting ATP Customers

The segmentation modelling approach used in this study reliably discriminated Individual and Business ATP risk customers into three latent classes: non-ATP risk customers; ATP risk customers; and Opportunists.

The segmentation had a very low classification error rate for both Individual and Business customers, showing the classification method was robust. The segmentations for the Individual customers and Business customers were both clear cut, with the Individual customer segmentation being the clearer of the two.

The qualitative interviews echoed the Individual and Business customer segmentations in that they indicated a full spectrum of views regarding what is acceptable and unacceptable tax planning. At the 'aggressive' end of the spectrum, there appears to be a small number of customers with very combative ATP attitudes. The majority of the ATP customers appear to have more moderate views of ATP, but are keen to 'not pay more tax than they need to'.

The segmentation showed that about 10 percent of Individual and 20 percent of Business customers at risk of ATP did not depend on an agent. This requires further investigation as this indicates a group of customers with a unique form of ATP risk.

Having a reliable basis for identifying customers at risk of ATP means Inland Revenue will be able to better target its ATP-related audits and other interventions.

6.3 ATP Drivers

This research has shown that, even though ATP risk customers and tax agents agree with the necessity of tax, they strive to minimise tax as far as legally possible. The overall impression is that customers at risk of ATP are generally inclined towards low-end ATP behaviour, but they may fluctuate between 'acceptable' and 'aggressive' tax planning. Hence, ATP risk customers are illustrating precisely the concept of ATP in that they are concentrating on the 'letter of the law but not the intent'.

The quantitative analysis of Inland Revenue's data on ATP risk customers has shown that there are both similarities and contrasts in the ATP drivers for Individual and Business ATP risk customers. Both are heavy users of tax agents and yet both have lower administrative compliance which seems surprising as agent use is normally associated with higher compliance. Both groups have pronounced financial activity with their high income/turnover and liabilities.

What is interesting about Individual ATP risk customers is that they stand out from the wider population of Individual customers with their notably higher use of agents, being 'professionals,' and their specific use of travel expense claims, tax credit claims, and donations.

Business ATP risk customers, on the other hand, are harder to distinguish from the wider Business customer population. Their administrative compliance and use of tax agents are only slightly different to those of the wider Business customer population.

The data and the characteristics used in this research are limited and the outcome of this study can be used as a guideline for future research to explore the behaviour of ATP risk customers more closely. For example, tax research literature has identified that Effective Tax Rate (ETR) has an important role in identifying ATP risk from non-ATP risk customers. Any future work should focus on this factor more closely to come up with a pragmatic solution.

Some positives can be seen in the overall environment surrounding ATP risk activity. In the last 10-12 years, New Zealand has seen business and personal tax rates reduced, stronger tax legislation, a tougher ap-

proach from Inland Revenue, and a reduced appetite for risk due to the global economic crisis. However, ATP risk activity has not ceased, and Inland Revenue has a long way to go to move ATP risk customers and agents from concentrating on legal specifics to concentrating on the overall intent of the tax legislation.

Further work needs to be done on a sample of customers who are actually known to be involved at different levels of ATP schemes including high-, medium-and low-risk scenarios.

Finally, much is still needed to be done to understand the complex nexus existing between ATP risk customers and tax agents.

References

Dyreng, S., Hanlon, M., & Mayde, E. (2007). *The Effects of Managers on Corporate Tax Avoidance.* Working Paper, University of North Carolina.

Grasmick, H. G., & Green, D. E. (1980). *Legal Punishment, Social disapproval and Internalisation as Inhibitors of Illegal Behaviour*. Journal of Criminal Law and Criminology, 71(3):325–335.

Greggi, M. (2008). *Avoidance and Abus de Droit*. eJournal of Tax Research, vol 6, no.1, pp. 23–44.

Kagan, R. A., & Scholz, J. T. (1984). The Criminology of the Corporation and Regulatory Enforcement Strategies. In K. Hawkins and J. M. Thomas (eds) Enforcing Regulation (p. 69). Boston: Kluwer-Nijhoff Publishing.

Kirchler, E., & Maciejovsky, B. (2001). *Tax Compliance Within the Context of Gain and Loss Situations, Expected and Current Asset Position, and Profession*. Journal of Economic Psychology, 22:173–194.

Murphy, K (2002). *Aggressive Tax Planning: Differentiating Those Playing the Game from Those Who Don't.* Centre for Tax Systems Integrity. Canberra: The Australian National University.

Murphy, K., & Byng, K. (2002). *Preliminary Findings from the Australian Tax System Survey of Tax Scheme Investors*. Centre for Tax Systems Integrity Working Paper No. 40. Canberra: The Australian National University.

Nash, J., & Pross, A. (2009). The OECD *Work on Aggressive Tax Planning*. Tax Notes International, October 12, pp. 117–119.

Robinson, J.R., Sikes, S.A., & Weaver, C.D. (2008). The Impact of Evaluating the Tax Function as a Profit Centre on Effective Tax Rates. Working Paper, University of Texas.

Sakurai, Y., & Braithwaite, V. (2001). *Taxpayers' Perceptions of the Ideal Tax Adviser: Playing Safe or Saving Dollars?* Centre for Tax Systems Integrity Working Paper No. 5. Canberra: The Australian National University.

Vanistendael, F. (1997). *Judicial Interpretation and the Role of Anti-Abuse Provisions in Tax Law*, in G.S.Cooper (ed.), Tax Avoidance and the Rule of Law, Amsterdam, p. 153.

3

Corporation Income Tax Enforcement

Macias ◆ Wang

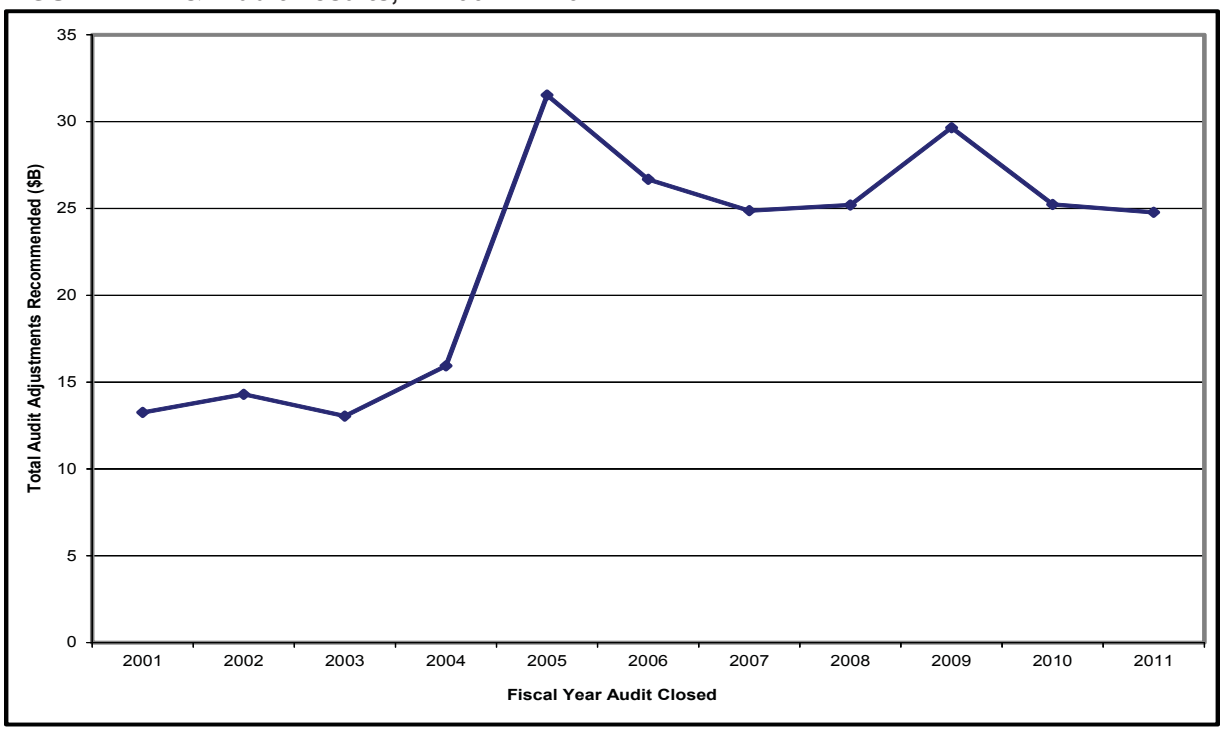
Howard

Analysis of Ten-Year Trends in Large Business Examination Results (Fiscal Years 2001–2011)

David Macias and Kimmy Wang,
Large Business and International Division, Internal Revenue Service[1]

The Large and Mid-Size Business ("LMSB") division, now called the Large Business and International ("LB&I") division, produced dollar examination results in Fiscal Year (FY) 2005 that were double those in FY 2004 and earlier.[2] Examination results remained at a higher level in FY 2006 through FY 2011.[3]

FIGURE 1. LB&I Audit Results, FY2001–FY2011



This paper analyzes the reasons for the increase in examination results in FY 2005 and later through research and analysis of various IRS databases, review of related financial information from SEC filings, and interviews with selected examination teams.

[1] The authors wish to acknowledge comments and assistance from Charles Boynton, and Thomas Brandt. Any errors are solely the responsibility of the authors. The opinions expressed are those of the authors and do not necessarily represent positions of the U.S. Department of the Treasury or the Internal Revenue Service.

[2] This paper analyzes the magnitude of the additional tax and penalties recommended by auditors at the conclusion of the audits. Some of this "recommended" adjustment is often arbitrated in administrative appeal and/or litigation, resulting in a smaller amount being formally assessed. And, although it is normally not a problem among corporations, the amount that is formally assessed as legally due may not be paid in full. Nonetheless, this paper does not address the assessed or collected amounts that correspond to these recommended amounts—largely because there are often long lags of time between the recommendation and the eventual assessments and collections.

[3] For simplicity, this paper will use the current name of Large Business and International (LB&I) for discussing any year within the Fiscal Year (FY) 2001-2011 period. The U.S. federal government uses a fiscal year of October through September. Fiscal Year 2001 is the fiscal year ending September 30, 2001.

The results of this study will provide insights into factors that may have contributed to increased exam results and can be used by LB&I in understanding prior results and for consideration in future decision making.

Examination Closures

When the IRS discusses examination results for a particular year, they mean any examination that closed during that fiscal year. For example, the results for any examination completed in FY 2005 would be included in that year's results. It does not matter as to the tax year of the return involved. The results for a fiscal year include the audit of any return of any tax year that closed in that fiscal year.

IRS Databases

The authors used the Audit Information Management System ("AIMS") and Audit Computer Information System ("ACIS") to analyze examination results from FY 2001 through FY 2011. The Issue Management System ("IMS") was used to analyze specific issues raised by examination field Revenue Agents from FY 2007 through FY 2011. Prior to that, the Coordinated Examination Management Information System for Large Cases ("CEMIS") was used to analyze specific issues from FY 2001 through FY 2006. The CEMIS database is more restrictive in the type of information that can be accessed. CEMIS captured only the top ten issues raised on an examination whereas IMS captures all of the issues raised during an examination. The results are further broken down by LB&I sub-industry (taxpayers with similar principal business activities and a common Standard Industrial Classification (SIC) code) to analyze their contribution to the overall increase in examination results and to determine strategic issues that need the attention of top management to foster voluntary compliance.

Issues to address

1. Why did LB&I field examination results (in dollars) double in 2005 compared with prior years? And why did they remain constant thereafter?
2. Was there a particular sub-industry or Internal Revenue Code Section issue that drove the increase?
3. Were the recommended dollars generated from a small group of entities?
4. What actions should LB&I management take as result of this study to foster increased voluntary compliance?

Findings

1. The FY 2004 LB&I Special Initiative to Improve Business Results by decreasing cycle time on both Industry Cases ("IC") and Coordinated Industry Cases ("CIC") Examinations appears to have been successful.[4]
2. LB&I introduced issue tiering in FY 2006, which did help in focusing IRS resources on the most significant issues.
3. The exam recommended dollars in FY 2005 and thereafter were mostly driven by a small number of sub-industries. These sub-industries are Utilities, High Technology, Petroleum, Commercial Banking, Securities and Financial Services and Telecommunications.
4. The major issues that drove the increase in examination dollars appear to be transfer pricing and capitalization.
5. The increase in recommended dollars appeared to be concentrated in a small group of CIC cases. We identified 147 CIC examinations that drove examination results from FY 2001 through FY 2011.

[4] LB&I cases fall into two categories, Coordinated Industry Cases (CIC) or Industry Cases (IC). A CIC is a taxpayer and its effectively controlled entities that warrant the application of team examination procedures and meet the required point criteria. A case qualifies as a CIC if after using the point criteria the case totals 12 or more points as outlined in IRM Exhibit 4.46.2-2. An IC is a taxpayer and its effectively controlled entities that warrant the application of Case examination procedures but do not meet the definition of a CIC.

6. Also considered were the fluctuations of Revenue Agent (non-supervisory, bargaining unit employees) resources during the same time period. The number of revenue agents available to perform examinations fluctuated from a high of 4,440 in 2001 to a low of 3,778 in 2003. There was no correlation between revenue agent resources and the number of case closures.

LB&I Special Initiative To Improve Business Results

LB&I set new currency goals for FY 2004 for both IC and CIC cases. The currency goals centered on the average number of months that IC and CIC examinations take. The new goal was 12.7 months for IC cases and 30 months for CIC cases. The initiative also set a goal that all in-process IC cases be closed by April 30, 2004 and all in-process CIC cases by closed by March 31, 2005. LB&I did meet these targets by FY 2005, which resulted in more focused issue examinations. As a result of the tightened cycle time, examiners focused on the most material issues present in the examination. This freed up examination resources to focus on increasing coverage and to focus on areas of greatest compliance risk.

Figure 2 shows the downward trend in cycle times for both IC and CIC examinations due to both the currency initiative and issue focus. In FY 2004, the cycle time for CIC cases was 29.5 months. This amount decreased significantly by FY 2011 to 18.9 months, or a 27 percent decrease in cycle time. For IC cases, the drop in cycle time from FY 2004 to FY 2011 was also 29 percent.

FIGURE 2. LB&I Audit Cycle Time (From Status 12 to Closure), FY01–FY11[5]

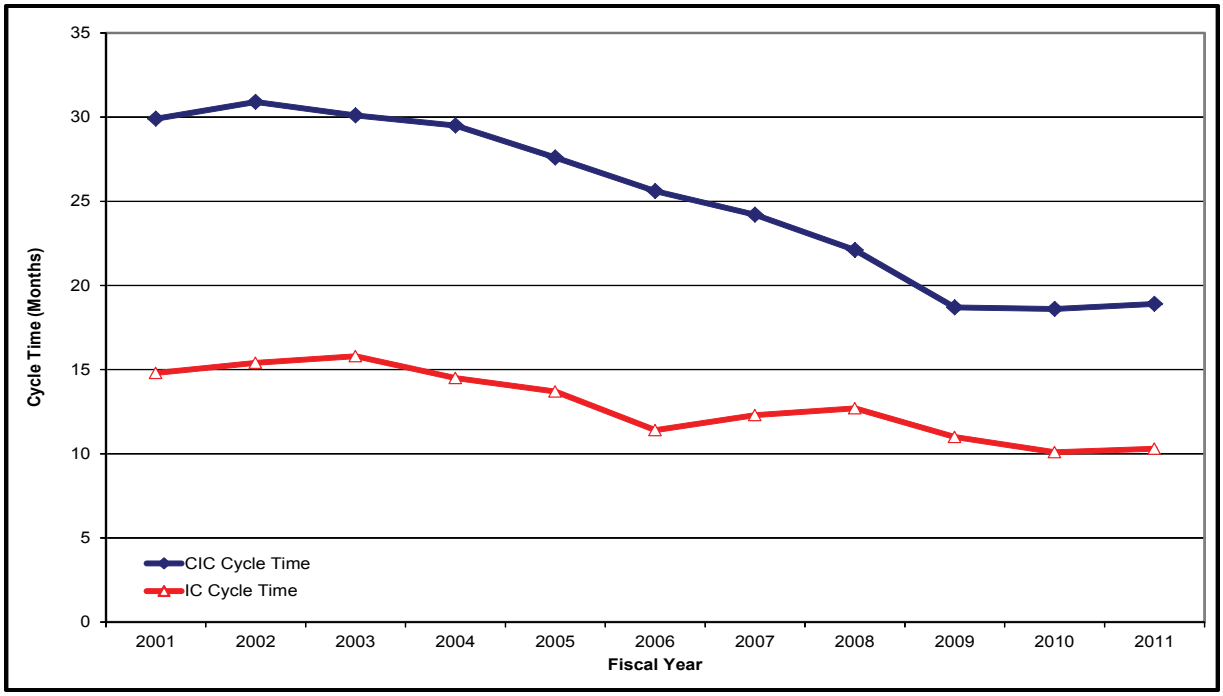


Figure 3 shows that the currency initiative worked for LB&I. Total CIC return closures reached a high in FY 2005 of 5,760. The number of closures dropped off thereafter as a result of LB&I becoming more current in the years under examination. We will later see that the recommended dollars were the highest in FY 2005 and gradually dropped in succeeding years.

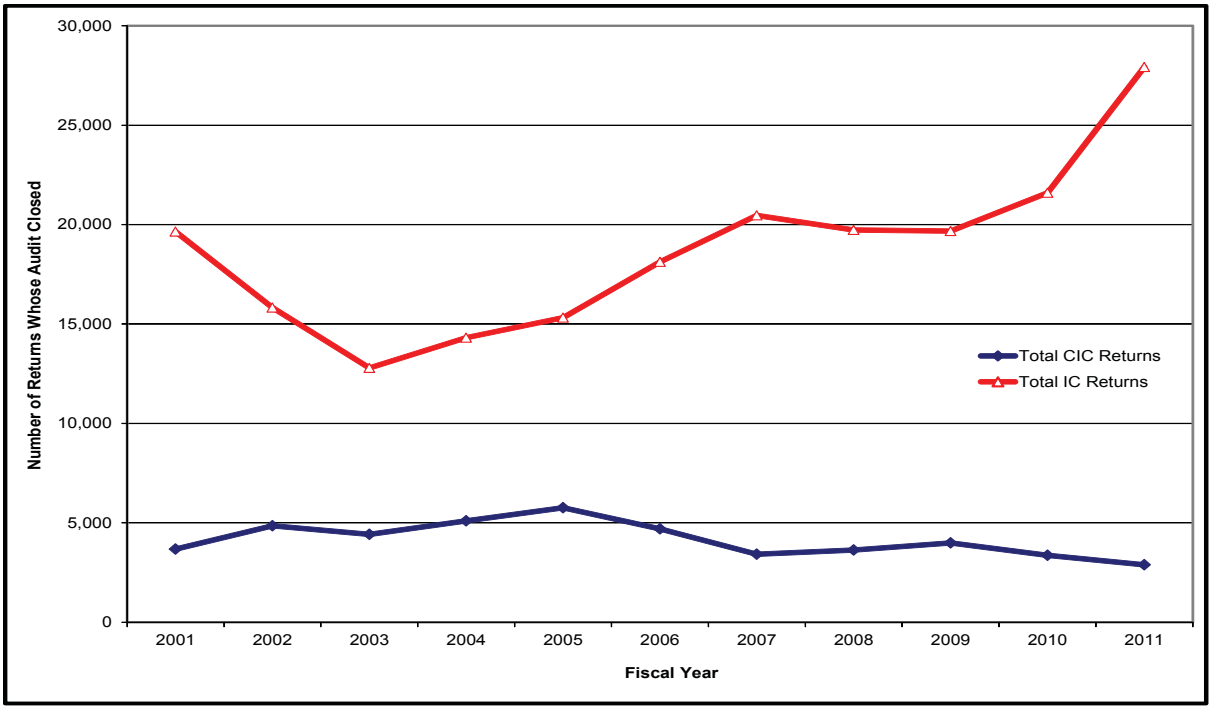
One of the goals LB&I set in FY 2006 was to "improve ability to deploy resources based on risk assessment."[6] This goal resulted in more attention to issues with high compliance risk and more examination of IC

[5] Status 12 indicates that the tax return is in Examination Status.

[6] Fiscal Year 2006 LMSB Field Focus Guide

cases. The number of IC case closures increased from 15,318 in FY 2005 to 27,926 in FY 2011, or an increase of 82.3 percent. Even though IC cases do not generate as many recommended dollars as CIC cases, the coverage rate was increased for compliance coverage of these taxpayers. Many IC taxpayers, although not as large as the CIC taxpayers, still have very large asset size and operate in a multi-national level.

FIGURE 3. Number of LB&I Return Closures, FY01-FY11



Issue Tiering

In FY 2006, LB&I introduced Issue Tiering as an element of their Issue Management Strategy.[7] It was designed to prioritize issues in a coordinated manner that provided consistency of treatment among taxpayers. Tier I issues were defined as issues of "high strategic importance" that have a significant impact on one or more industries. Tier I issues included:

- Transfer of Intangibles / Offshore Cost Sharing
- Foreign Tax Credit Generator
- Research Credit Claims
- All Recognized and Listed transactions (Reportable Transactions on Form 8886)
- Section 199, Domestic production Deduction

Data and Analysis

The cumulative dollars recommended by LB&I revenue agents was between $13.25 billion dollars in FY 2001 and $15.93 billion dollars in FY 2004. The recommended dollars roughly doubled in FY 2005 to $31.52 billion and has stayed at a higher level through FY 2011.

[7] Tiering issues were discontinued on August 27, 2012

TABLE 1. Audit Recommended Dollars for FY 2001 to FY 2011

LB&I Asset Groupings	FY01	FY02	FY03	FY04	FY05	FY06	FY07	FY08	FY09	FY10	FY11
Total	**$13.25**	**$14.30**	**$13.04**	**$15.93**	**$31.52**	**$26.68**	**$24.87**	**$25.20**	**$29.65**	**$25.23**	**$24.78**
$10-$250M in Assets	$0.49	$0.64	$0.78	$0.73	$1.40	$1.17	$0.73	$0.90	$1.10	$1.49	$0.93
Assets over $250M	$12.76	$13.66	$12.26	$15.20	$30.12	$25.51	$24.14	$24.30	$28.55	$23.74	$23.84

Source: ACIS

Table 1 also illustrates that the dollars doubled in FY 2005 among both the largest and the smallest corporations, but the largest dollar increase occurred among the largest taxpayers (assets over $250 million dollars). The recommended dollars in this group grew from $15.2 billion in FY 2004 to $30.12 billion in FY 2005. This group of taxpayers accounted for over 94 percent of the total recommended dollars between FY 2005 through FY 2011. By comparison, the smaller taxpayers do not significantly contribute to the overall dollars recommended by LB&I. The increase in recommended dollars in FY 2005 and FY 2006 may be attributed to the large number of CIC case closures of 3,500 and 2,396 respectively. Many of these cases were closed in FY 2005 and FY 2006 due to the currency initiative. These cases also have been in process for many months (35 months in FY 2005 and 36.8 months in FY 2006), which may indicate that the dollar amount of the issues were large and complex. The number of months the case was in process from Status 12 to closure started dropping off in FY 2007 through FY 2011. The decrease in cycle time is the result of examinations being more current and issue focused.

Table 2 shows dollars recommended from LB&I CIC examinations. Table 2 also illustrates that the increase in examination recommended dollars occurred in the unagreed category. The overall unagreed amount ranged between $7.16 billion dollars and 9.16 billion dollars between FY 2001 and FY 2004, averaging $8.33 billion. The average then increased in FY 2005 through FY 2011 to $17.65 billion, which is a 112 percent increase.

TABLE 2. LB&I Audit Recommended Dollars for CIC by Disposal Code (in $ Billions)[8]

Fiscal Year	Disposal Code			
	3 Agreed	7 (Unagreed)	Other Disposal Codes	Total
2001	$0.48	$9.13	$0.39	$9.99
2002	$1.03	$9.16	$0.34	$10.54
2003	-$0.19	$7.88	$0.62	$8.31
2004	$0.36	$7.16	$1.66	$9.19
2005	$2.17	$19.69	$1.00	$22.85
2006	-$1.83	$16.93	$0.22	$15.31
2007	$0.64	$16.37	$0.24	$17.26
2008	-$0.41	$14.88	$0.39	$14.86
2009	$2.34	$20.53	$0.46	$23.34
2010	-$2.35	$15.79	$1.05	$14.49
2011	$0.45	$19.36	$0.63	$20.44
Total	**$2.69**	**$156.89**	**$7.01**	**$166.59**

Figure 4 plots LB&I examination recommended dollars from FY 2001 through FY 2011 in total and by corporation asset size. The large increase in recommended dollars occurred between FY 2004 and FY 2005, from 15.93 billion dollars to 31.52 billion dollars, or a 97.87 percent increase.

[8] Disposal codes are used to indicate the disposition of an examination. For example disposal code 3 is agreed cases.

Figure 5 plots types of LB&I examination closures. The unagreed dollars (Disposal Code 07) mimics the results for all disposal codes (both agreed and unagreed). Figure 5 illustrates that the majority of examination recommended dollars are unagreed.

Figure 6 tracks only LB&I examination recommended dollars that resulted in an overall additional tax adjustment for the taxpayer. Refund or no-change cases have been eliminated from the data. This graph clearly illustrates that increases in the overall examination recommended dollars mirror unagreed recommended dollars in every year, from FY 2001 through FY 2011. It also shows that unagreed examination recommended dollars are driving total examination results in LB&I.

FIGURE 4. Audit Recommended Dollars by Corporate Assets

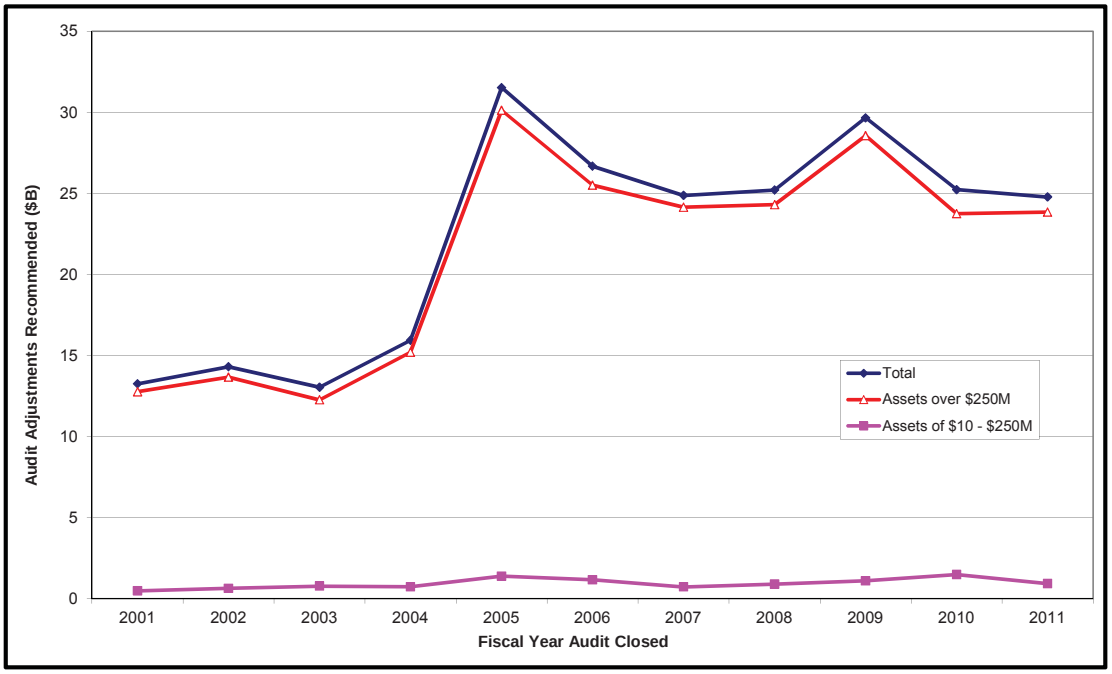


FIGURE 5. CIC Taxpayers All Audit Recommended Dollars by Disposal Code

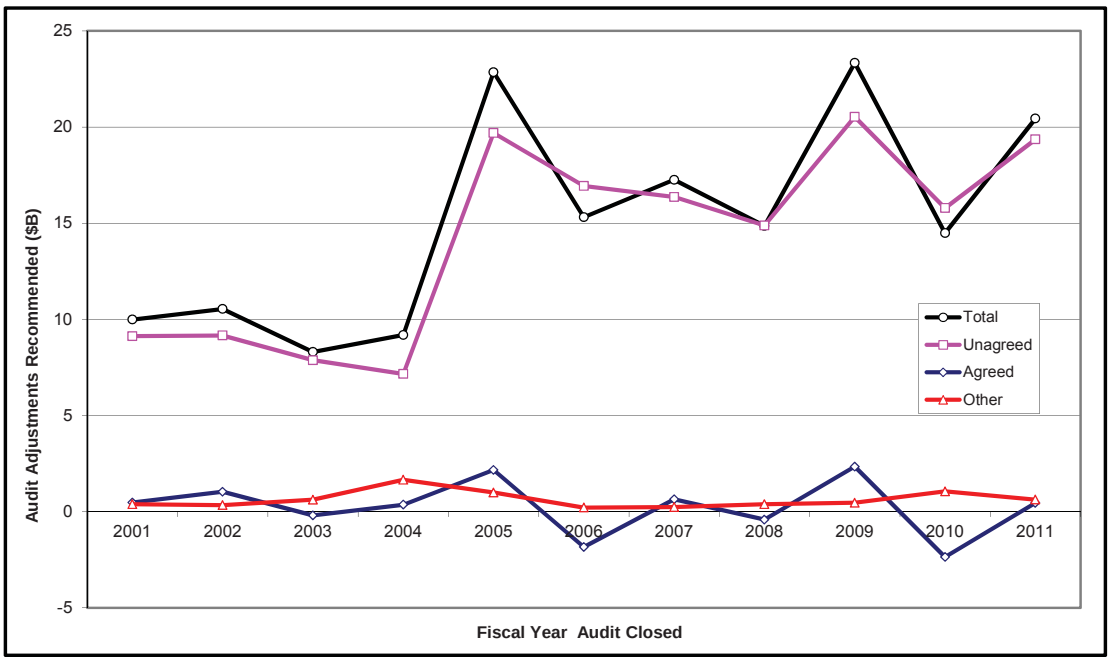


FIGURE 6. Audit Recommended Positive Dollars by Disposal Code, CIC Taxpayers

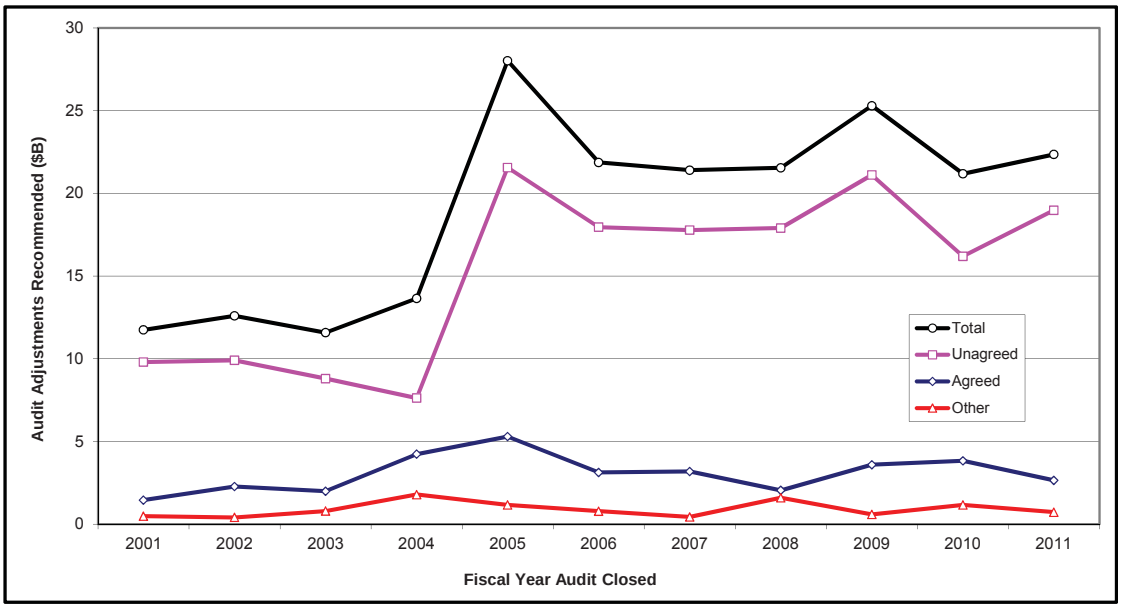
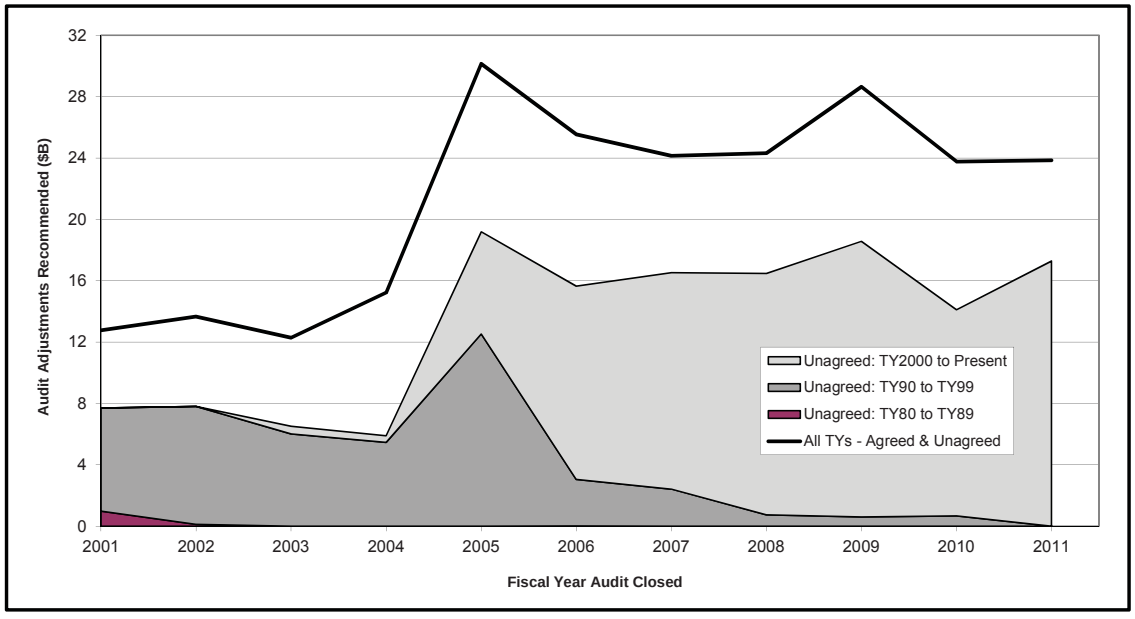


Figure 7 illustrates the mix of tax years among unagreed cases having aggregated adjustments of at least $100 million during the period, and compares that with all recommended adjustments. For example, the majority of unagreed recommendations closing in FY 2005 were from the Tax Years 1990 through 1999. Note also that the majority of total adjustments in each year are associated with unagreed cases in which the aggregated recommended adjustment is at least $100 million.

FIGURE 7. Allocation of Unagreed Recommended Adjustments of $100M or More Among LB&I Audits By Return Tax Year and Fiscal Year of Closure, Compared with All LB&I Audit Adjustments



Sample Selection

We have previously shown that the majority of examination dollars recommended pertain to CIC unagreed adjustments. We have selected 147 taxpayers with the largest unagreed dollars recommended (unagreed audit amounts exceeding 250 million in total during 2001 to 2011 grouped by Taxpayer Identification Number). See Table 3. The 147 taxpayers accounted for 71 percent of the total dollars recommended, but only 2 percent of the number of taxpayers examined.

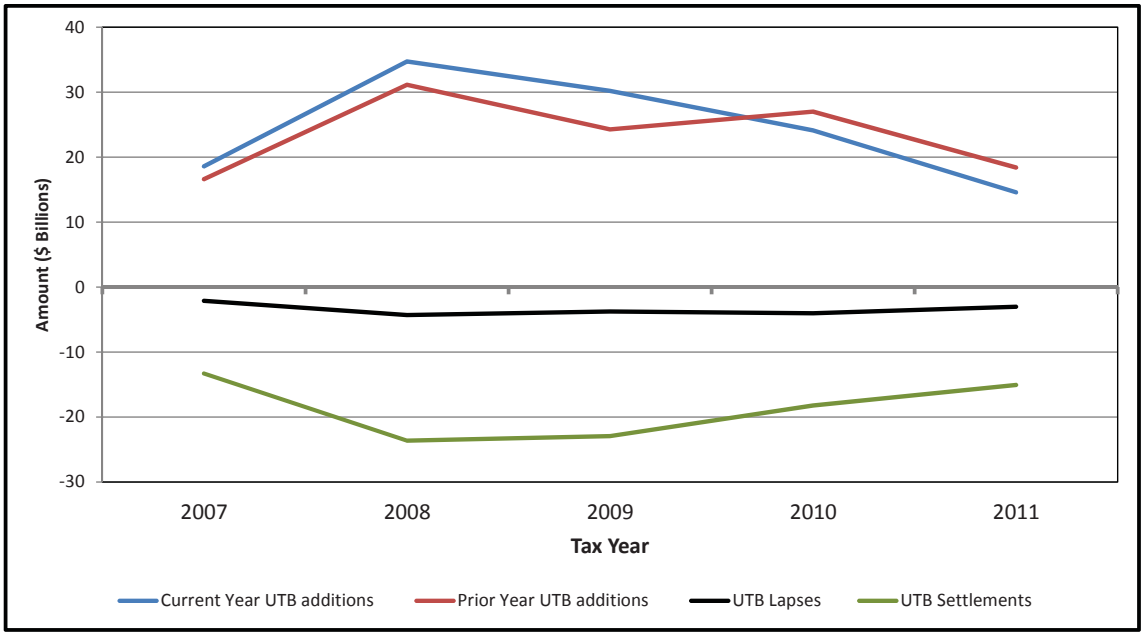
TABLE 3. Comparison of Selected CIC Taxpayers With Total LB&I Population, FY2001-FY2011 Closures

Attributes	Selected CIC Returns	Total LB&I Population	Selected Returns as a Share of the Total
Number of Audited Taxpayers	147	8,443	2%
Number of Audited Tax Returns	6,778	41,873	16%
Recommended Audit Adjustment ($ Millions)	$127,203	$179,886	71%

IRS Audits and Uncertain Tax Benefit (UTB) Financial Reporting

We analyzed public CIC taxpayer financial statements for evidence of the impact of IRS audit exam results on the financial statement reporting of examined taxpayers. Figure 8 reflects financial statement disclosures of 500 of the largest CIC taxpayers' Uncertain Tax Benefit (UTB) movements from TY 2007 through TY 2011.

FIGURE 8. UTB Movement by Tax Year, 500 Largest CIC Taxpayers



In Figure 8, the Current Year UTB addition and Prior Year UTB addition reflect a similar pattern over the 5-year period. Both current year and prior year UTB balances spiked in TY 2008. Recall that Figure 3 shows spikes in FY 2005 and FY2009. During an examination, the notices of proposed adjustment (NOPA) usually are issued throughout the audit for smaller and more agreeable issues. Those proposed audit adjustments for the largest, more complex and contentious issues are routinely issued toward the end of the audit cycle. This is mainly due to the time it takes to collect and analyze all the facts relevant to the issues. The final proposed tax increase probably won't be settled until the end of the audit cycle.

For the exam case closures in FY 2009 in Figure 3, it is reasonable to assume that the audit adjustments were issued and known to taxpayers before FY2009, such as in FY 2008. The spike of the FY 2009 exam result matched a similar spike in financial statement UTB disclosure in TY 2008. That is, the taxpayers knew of the proposed adjustments prior to filing their 2008 financial statements and appropriately increased their income tax reserve taking into account their current information on the status of the LB&I examination.

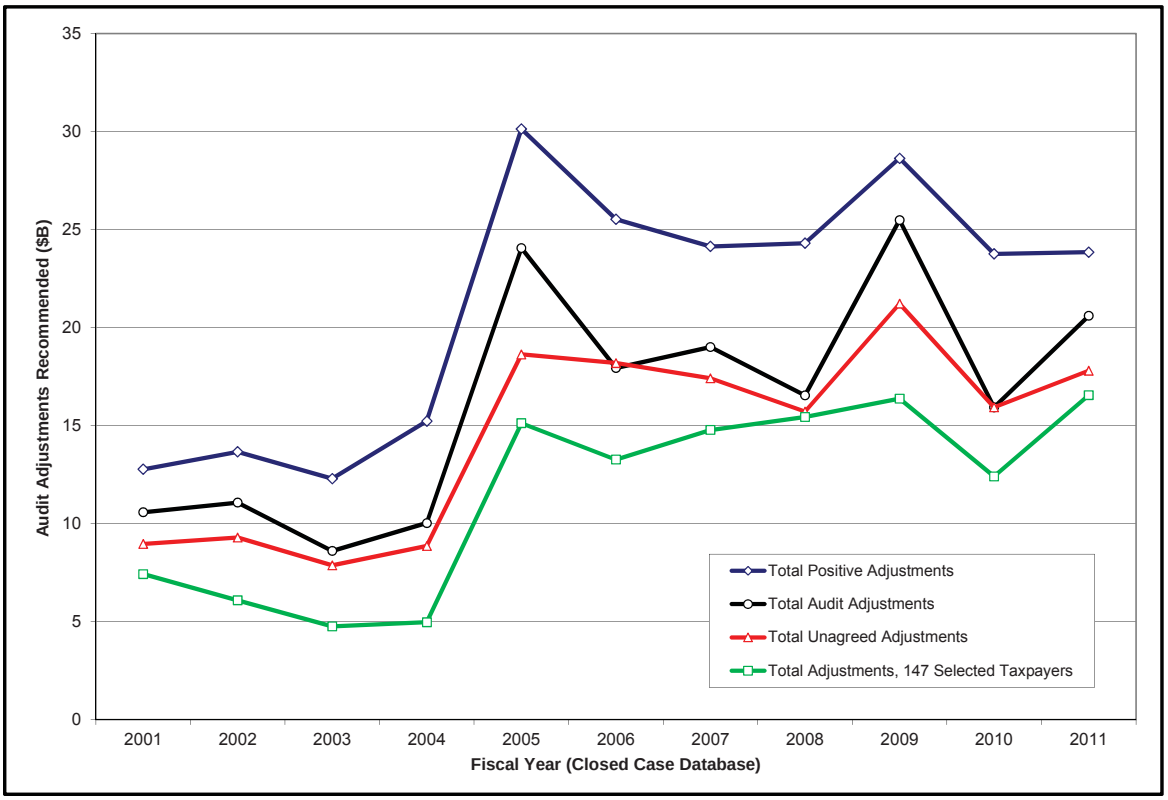
Also interesting is a similar upward movement of the current year UTB and prior year UTB in TY 2008 with a similar downward trend in the years thereafter. There were some internal case studies that demonstrated that taxpayers make a financial statement income tax reserve based on proposed audit adjustments. This explains the correlation of prior year, anticipatory, UTB movements with LB&I exam audit results. The internal case studies, although limited at this point, do support the data pattern between exam audit results and the financial statement reserve movement.

The Schedule UTP filing data for TY 2010 and TY 2011 reflects that the most reported IRC section is Section 41 (Research Credit) and Section 482 (Transfer Pricing), these IRC sections have a recurring effect on the taxpayer's tax return reporting for years to come. This may explain in part why the current year UTB addition has a similar moving pattern as the prior year UTB addition.

Selected Taxpayer Characteristics

Figure 9 breaks down LB&I examination recommended dollars by fiscal year by total positive adjustments (ACIS Table 37) , total adjustments, total unagreed adjustments, and the total adjustments for the 147 selected taxpayers.[9]

FIGURE 9. Audit Adjustments of Selected CIC Taxpayers Compared With All of LB&I



[9] ACIS table 37 refers to unagreed audit recommended dollars without negative adjustment.

The overall LB&I recommended dollars are heavily influenced by the examination outcome of the 147 taxpayers in our study. As the recommended dollars for the 147 taxpayers fell between FY 2001 – FY 2004, the overall examination results also fell. When the recommended dollars for the 147 taxpayers increased in FY 2005, so did the overall results.

Selected Taxpayers—Top Issues Raised by Examination

Figure 10 shows that the average recommended examination dollars for our selected population of 147 taxpayers increased from $6 billion dollars per year from 2001–2004 to $15 billion dollars per year in 2005–2011.

FIGURE 10. Average Examination Recommended Amount (in $ Billions), Selected CIC Taxpayers for Two Periods

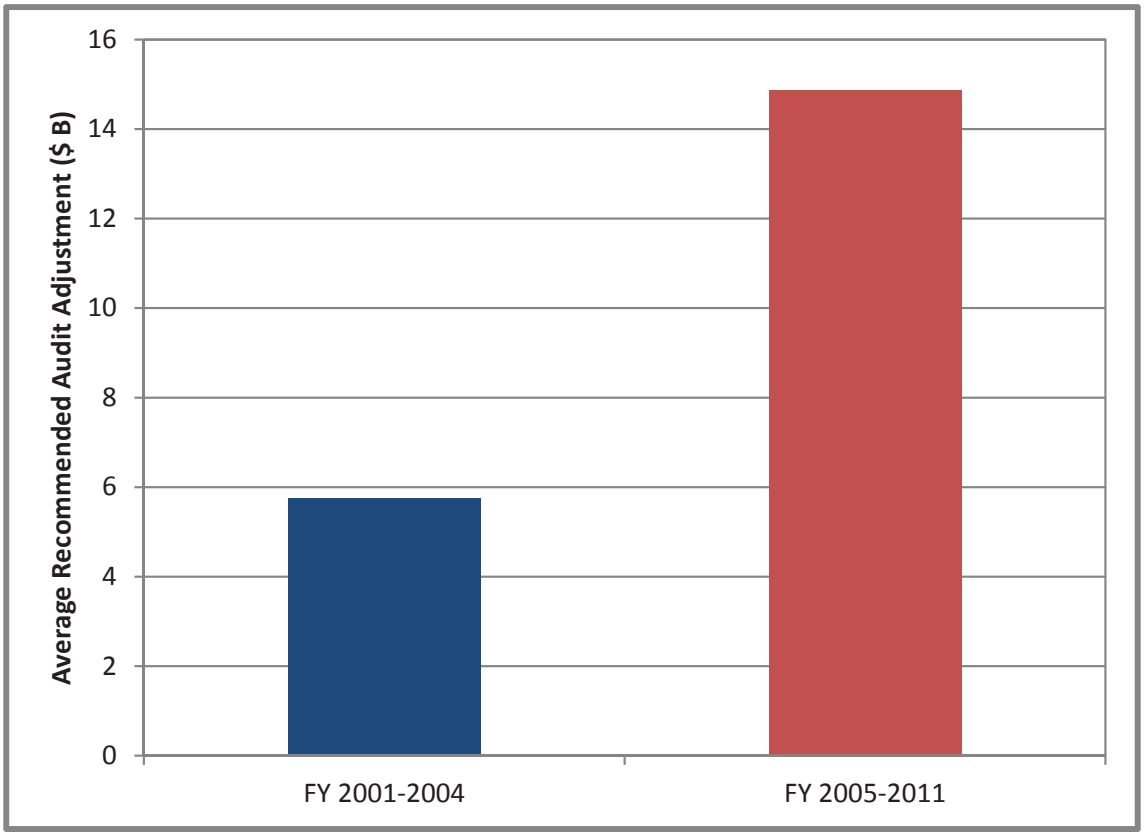


Figure 11 shows average examination recommended dollars by the top six sub-industries from FY 2001–FY 2004 and FY 2005–FY 2011. In all six sub-industries, examination recommended dollars significantly increased.

Figure 12 shows the top ten issues raised by examination (identified by Internal Revenue Code Section and measured by recommended amount) from FY 2007 through FY 2011 for the sample population. Two of the four issues with the most recommended dollars are Credit for Increasing Research Activities and transfer pricing. The Research credit issue continues to generate a lot of audit activity with unagreed recommended dollars of $19 billion and agreed dollars of $63 billion.

Figure 12 also reflects that the field responded positively to the tiering issue concept. Transfer pricing and research credit were both designated as Tier 1 issues. The graph above reflects material adjustments in transfer pricing and research credit issues.

Figure 13 shows how the different sub-industries contributed to the selected examination group's recommended dollars by year.

FIGURE 11. Average Audit Recommended Amount (in $ Billions), Top 6 Sub Industries of Selected CIC Taxpayers for Two Periods

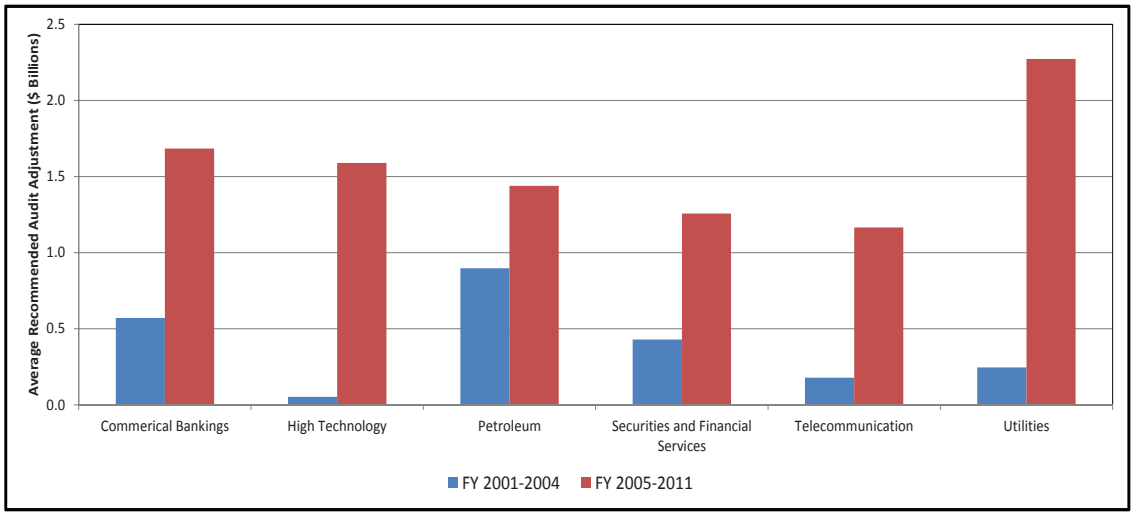
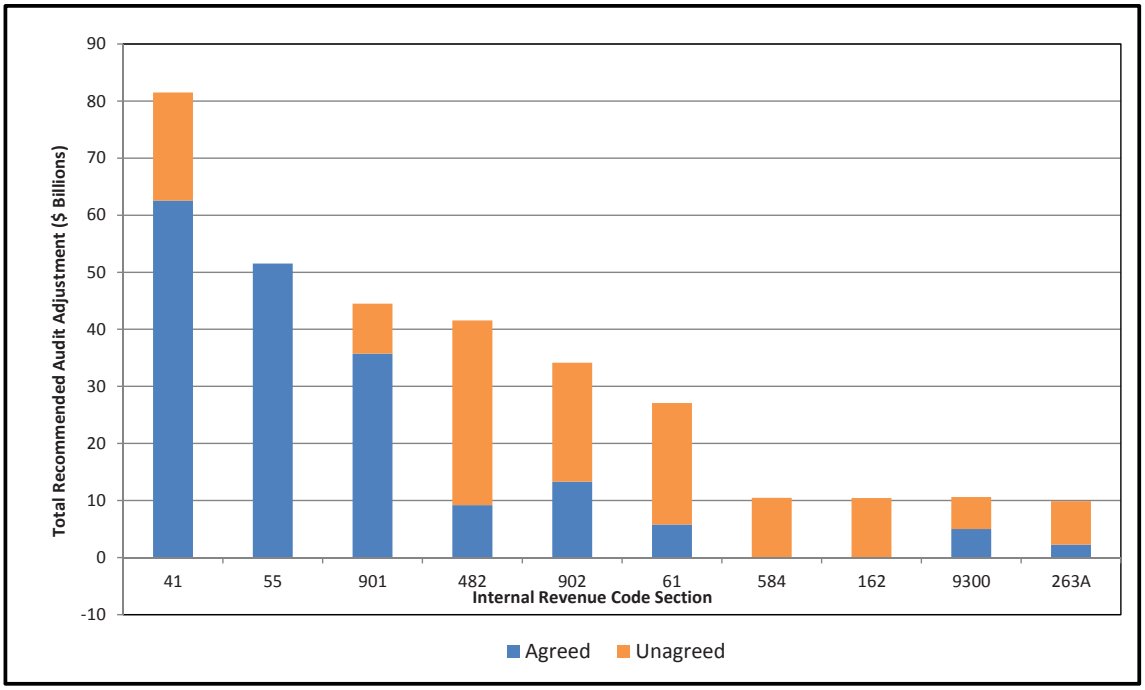


FIGURE 12. Total Amount Recommended (in $ Billions) Associated With the Eight IRC Sections Having the Largest Amounts, Selected CIC Taxpayers by Agreed vs. Unagreed Issues, FY 2007 through FY 2011

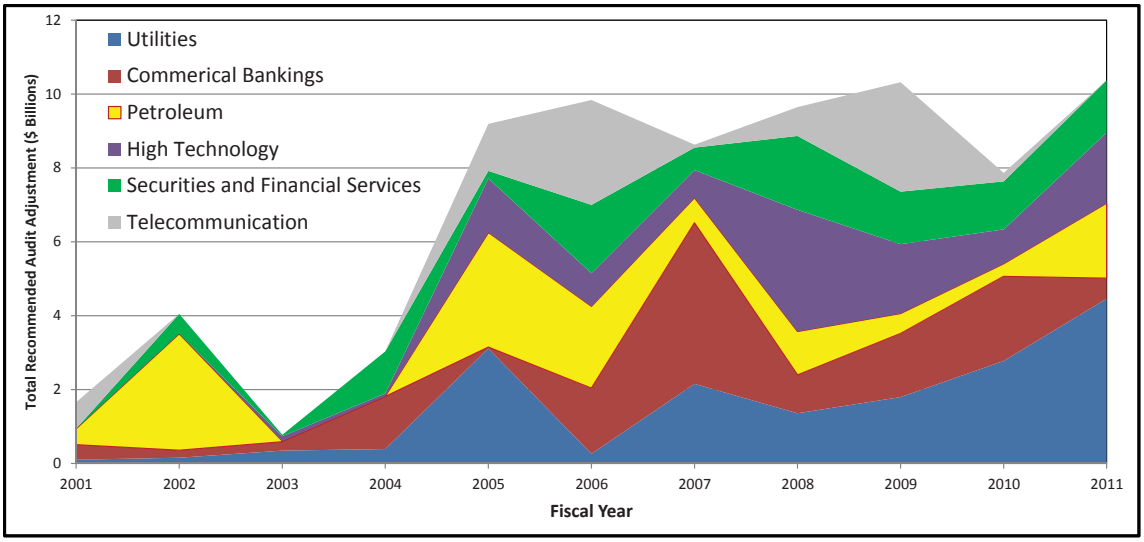


Note: see Table 4 for a description of these Code sections.

TABLE 4. Internal Revenue Code Section Descriptions

IRC Section	Brief Description
41	Credit for Increasing Research Activities
55	Alernative Minimum Tax Foreign Tax Credit
61	Gross Income Defined
162	Trade or Business Expense
263A	Capitalization
482	Transfer Pricing
584	Net Operating Loss Deduction
901	Foreign Tax Credit - Direct
902	Foreign Tax Credit - Deemed paid credit
9300	Sale In / Lease Out (SILO)

FIGURE 13. Total Recommended Audit Adjustment (in $ Billions) of the Top 6 Sub-Industries Among Selected CIC Taxpayers, FY2001-FY2011



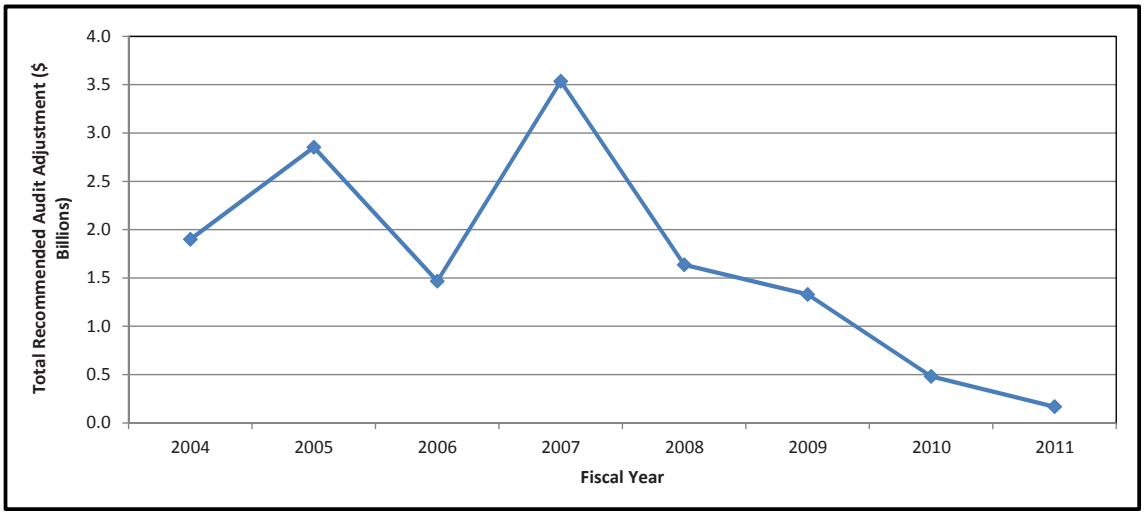
In FY 2002, the Petroleum industry contributed the most to examination recommended dollars, whereas in FY 2007, the Commercial Banking industry contributed the most recommended dollars.

A major LB&I initiative in the early 2000's was corporate tax shelters. Tax shelter cases took longer to close due to the complexity of the issue and the amount of documents involved coupled with extensive depositions. This is the reason why tax shelters were excluded from the currency initiative of 2004. As shown in Figure 14, related tax shelter recommended dollars contributed to the overall examination results in all of the years under study, but especially in FY 2005 and FY 2007.[10] LB&I pushed the examination of listed transactions and initiated promoter examinations in 2001. In a report to Congress, dated 11/20/2003, it was stated that, "The audits began in LMSB in 2001, with 22 entities under investigation. As is apparent, we have increased the number of audits significantly. This includes large accounting firms and major law firms, as well as banks and a number of boutique and mid-size promoters."[11]

[10] These cases have an identified listed tax shelter. Not all of the recommended dollars necessarily are generated from the shelter activity.

[11] Remarks given by IRS Commissioner to the United States Senate, hearing on Abusive Tax Shelters.

FIGURE 14. Total Recommended Audit Adjustment (in $ Billions) for Cases Closed with a Tax Shelter Project Code, by Fiscal Year of Case Closure



Conclusion

We find that the increase in LB&I examination results is attributable to the following five factors:

1. **The currency initiative:** As part of the LB&I currency and cycle time improvement initiative, the field closed more cases in FY 2004 (deadline for IC cases) and 2005 (deadline for CIC cases). The "old" CIC tax years were closed out as a result of this initiative during FY 2005, which resulted in an increase in examination recommended dollars.
2. **Tax shelter activity:** LB&I investigation of tax shelter promoters started in FY 2001 along with increased corporate tax shelter examinations. As stated above, the cycle time on these cases were longer than non-shelter cases. Shelter cases tended to close out of examination as unagreed cases. The results of these examinations did not start hitting the statistics until FY 2004, with peak years in FY 2005 and FY 2007.
3. **Issue tiering:** LB&I focus on specific issues stated in FY 2006 (issue tiering). Examination teams were mandated to work Tier 1 issues, which included transfer pricing and research credit issues. The result of this mandate generated many of notice of proposed adjustments for tax years FY 2007 through FY 2011.
4. **Sub-industry examination results:** An analysis of the examination results by sub-industry shows that the top six sub-industries increased their proposed dollars recommended from FY 2001–FY 2004 to FY 2005–FY 2011. The sub-industries with the biggest average increase were utilities and high-tech.
5. **The importance of 147 selected cases:** The data for the top six sub-industries include only a portion of the selected population of 147 taxpayers. The 147 taxpayers not in the top six industries were not included. The recommended dollars by sub-industry spiked in different time periods. The High Technology sub-industry (which incurred most of the transfer-pricing issue) and Telecommunications sub-industry generated most of the audit recommended dollars in the FY 2006–FY 2011 period; whereas the Petroleum sub-industry generated most of the audit recommended dollars in the FY 2001–FY2005 period. The Utilities sub-industry peaked in FY 2005 and FY 2011.

IRS Enforcement and State Corporation Income Tax Revenues

Margot Howard,[1]
University of North Carolina at Chapel Hill

1. Introduction

I examine the relation between IRS enforcement and state corporate income tax (SCIT) revenues. Many states are currently facing severe financial crises. For example, in early 2012 Moody's revised its outlook on Washington bonds from stable to negative.[2] In January 2013, Illinois postponed a bond auction amid concerns about the state's fiscal stability after S&P downgraded the state's debt (Nolan and Peters 2013). In 2011, Congress considered a measure that would have allowed states to declare bankruptcy in order to "get out from under crushing debts" (Walsh 2011). In July 2012, the State Budget Crisis Task Force, founded by Richard Ravitch and Paul A. Volcker, released its report on the threats to near- and long-term state fiscal sustainability.[3] One of the major threats discussed in the report is "narrow, eroding tax bases and volatile tax revenues." The report states, "The personal income, sales, and *corporate income taxes* are states' most economically sensitive and volatile revenues and they have grown in importance in recent years…. Since 1990, states' reliance on income taxes has continued to increase and the tax itself has become more volatile." [Emphasis added] Considering the current condition of state finances, it is imperative to understand the different factors that can affect state tax revenues.

I investigate one mechanism that may affect SCIT revenues: IRS enforcement.[4] The results suggest that there is a positive relation between tax enforcement at the federal level and SCIT revenues, even after controlling for SCIT rate, sales factor weighting, and federal corporate taxable income reported in the state. These results can be viewed in the context of Desai *et al.* (2007), with state governments as another set of outside stakeholders benefiting from higher levels of IRS enforcement. Following Guedhami and Pittman (2008), the proxies for IRS enforcement include IRS corporate income tax return audit rates, various IRS employment levels, and the number of fraud proceedings. In addition, based on prior evidence that the implementation of FIN 48 was associated with lower levels of state corporate tax avoidance (Gupta *et al.* 2013), I hypothesize and find that the positive relation between IRS monitoring and SCIT collections has been attenuated by the implementation of FIN 48.

In this paper, I attempt to examine both the mechanical and non-mechanical channels through which federal-level monitoring may be related to SCIT revenues. I control for the mechanical portion of the relation, driven by the fact that federal taxable income is typically the starting point for SCIT calculations, by including in my regression the federal corporate taxable income in a state in a given year. I then use the IRS enforcement variables mentioned above to examine the non-mechanical portion of the relation. However, even with the seemingly obvious mechanical relation it does not necessarily follow that a higher level of federal-level monitoring is associated with a higher level of SCIT revenues. IRS enforcement can focus on items, such as credits, that may not affect SCIT calculations. Even if additional federal taxable income flows through to the state-level return as the result of an IRS audit, that adjustment could lead to little, if any, additional tax at the state level as a result of the complexities of state-specific corporate income tax calculations. In addition, I examine *contemporaneous* IRS monitoring and SCIT collections. Since it typically takes years to complete a federal tax audit, this significantly lowers the likelihood that this relation is just the result of mechanical calculations.

[1] I appreciate helpful comments and suggestions from Josh Coyne, Brian Erard, Ed Maydew, Katie McDermott, Jenna Meints, Doug Shackelford, workshop participants at the University of North Carolina, and participants at the 2013 IRS-TPC Research Conference. All errors remain my own.

[2] For the revised opinion announcement see: http://www.moodys.com/research/MOODYS-REVISES-STATE-OF-WASHINGTON-RATING-OUTLOOK-TO-NEGATIVE-FROM—PR_236450.

[3] The State Budget Crisis Task Force focused their analysis on six states: California, Illinois, New Jersey, New York, Texas, and Virginia.

[4] Ideally, I would study IRS enforcement at the federal level in conjunction with enforcement by state tax authorities. Unfortunately, there is a large variance in what, if any, statistics states provide on the enforcement actions of their tax authorities. However, considering the size and resources of the IRS compared to state departments of revenue and the fact that federal corporate taxable income is incorporated into state corporate income tax calculations, the relation between IRS monitoring and state corporate income tax revenues is worthy of investigation.

I contribute to the literature in several ways. Prior studies have documented the relation between IRS monitoring and debt pricing, financial reporting quality, and cash effective tax rates. This paper provides evidence that IRS monitoring is also related to SCIT revenues. In the context of Desai *et al.* (2007), the state government appears to be another outside stakeholder that benefits from IRS enforcement, suggesting that states should consider federal-level enforcement when discussing state-level corporate income tax revenue and policy. This paper is also the first to document an attenuated relation between IRS enforcement and SCIT revenues after FIN 48 was implemented, providing further evidence that financial statement policies can affect cash tax outcomes. Comparing the results of this study to prior research also indicates that the SCIT landscape has shifted in recent years (as more states have adopted a more heavily weighted sales factor) and that this shift may be important in examining the determinants of SCIT revenues.

The paper proceeds as follows. Section 2 provides background information on state corporate income taxes. Section 3 discusses the related literature. Section 4 explains the hypothesis development. Section 5 discusses the data and research design. Section 6 provides the results. Section 7 concludes.

2. State Corporate Income Tax Background

Before examining SCIT collections, it is important to understand the basic components of the calculation of a firm's SCIT liability. First a corporate taxpayer must determine whether it has nexus with a state for corporate income tax purposes. Nexus exists when the taxpayer has a substantial enough connection to the state such that the state has the right to impose its income tax on the taxpayer. Then the taxpayer must look to a state's statutes and regulations to determine whether it will file a separate return or a combined return as part of a consolidated group.[5] Once the taxpayer has determined the correct filing group, it can begin the calculation of its SCIT liability. That calculation typically begins with federal taxable income. Each state then prescribes its own specific set of adjustments to that tax base, such as adding back state income taxes that were subtracted on the federal tax return.

If a firm operates in more than one state it must then separate its apportionable income from its allocable income. Apportionable income is typically considered to be "business income," income that is tied to the firm's core business, while allocable income is "nonbusiness income." The same item of income may be considered apportionable or allocable depending on the nature of the business of the firm. For example, a manufacturer would typically consider interest received on a bond to be allocable since it is not related to the general business of the firm. However, a financial services entity may consider the same interest to be apportionable income since holding bonds may be considered part of the central business of the firm. The taxpayer must then add back to apportionable income any specifically disallowed deductions. In many states, these disallowed deductions include royalty or interest payments to related passive investment companies or PICs.[6]

Once a firm has calculated its apportionable income it must multiply that figure by its apportionment factor. The formulae for calculating apportionment factors are state-specific, although generally they are based on some combination of property, payroll, and sales factors. The property factor is calculated as the fraction of total property in the United States that is located within that state. The payroll and sales factors are calculated similarly. However, states differ in their treatment of what is included in each of those factors and when an item is considered "sourced" to that state in calculating the numerator.

Once the firm has multiplied its apportionable income by its apportionment factor to arrive at apportioned state income, it must add any allocated income sourced to that state to arrive at state taxable income. State taxable income is multiplied by the applicable rate and then allowable credits are subtracted to arrive at the firm's liability.

3. Related Literature

This study is related to two streams of literature, research investigating SCIT policy and research investigating enforcement. Some SCIT policy studies focus on firm responses to those policies. For example, firms shift

[5] The terms "consolidated" and "combined" returns can have different implications for state tax purposes. However, that distinction is beyond the scope of this paper.

[6] See Dyreng *et al.* (2013) for a detailed discussion on the PICs.

their tax bases to favorable jurisdictions (as based on tax rates and sales factor weighting) and structure sales to reduce exposure to the throwback rule (Klassen and Shackelford, 1998).[7] Other SCIT policy studies examine the economic consequences of those policies. Goolsbee and Maydew (2000) and Lightner (1999) find that a state's payroll factor has a negative relation with manufacturing employment in that state. Gupta *et al*. (2009) examine the relation between various state tax policies and the level of SCIT revenue collected in the state for the years 1982 through 2002. They find that a higher sales factor weight is associated with lower SCIT revenue. Gupta *et al*. (2013) investigate the effect of FASB Interpretation No. 48 (FIN 48) on multistate income tax uncertainty. The model in Mills *et al*. (2010) illustrates that a mandatory disclosure environment (such as FIN 48) will deter certain taxpayers from engaging in tax avoidance transactions. Gupta *et al*. (2013) build on this finding and hypothesize that SCIT payments and revenues will increase with the implementation of FIN 48. They find that state effective tax rates and SCIT collections increased around the implementation of FIN 48, suggesting that the new rules for uncertain tax positions are related to lower levels of state tax avoidance by firms.

Desai *et al*. (2007) model the interaction between a corporate tax system and corporate governance. Their results suggest that higher levels of enforcement can benefit outside stakeholders by deterring managerial diversion. Guedhami and Pittman (2008) introduced into the accounting/finance literature the use of data from Transactional Records Access Clearinghouse (TRAC) on the level of IRS monitoring. They provide evidence that debt financing is less costly for private firms when there is a higher probability of a face-to-face IRS audit. They also hypothesize and find that private firms with high ownership concentration will have a stronger association between IRS enforcement and less costly debt financing, since these firms are particularly susceptible to agency issues between inside owners and outside shareholders. Higher levels of IRS monitoring have also been linked to higher quality financial reporting and increased firm cash effective tax rates (Hanlon *et al*., 2012; Hoopes *et al*., 2012).

Similar to this paper, Gupta and Lynch (2012) also examine the overlap of these two areas of research. Using data on corporate income tax enforcement expenditures collected from various state departments of revenue, they look at the association between those state-level expenditures and SCIT collections for 2000 through 2008. Their results suggest that state corporate income tax enforcement expenditures in year t are associated with increased SCIT collections in year t+2. They control for IRS enforcement in several of their specifications and find an insignificant or negative relation between federal-level enforcement and SCIT collections, although they note that those results may be because the variables they use "are likely an imperfect proxy for federal *corporate* enforcement." As mentioned earlier, ideally I would have run the analyses in this paper with a control for state-level enforcement as well. However, I do not have access to such data. Since the results in Gupta and Lynch (2012) show that both state- and federal-level enforcement may be related to SCIT collections, and this paper does not include a proxy for state-level enforcement, it is important to use caution when interpreting the results of this paper. Gupta and Lynch (2012) briefly address FIN 48, but only to note that their results are robust even after excluding 2007 and 2008 from their sample. They do not discuss the potential interaction between the enforcement and the FIN 48 financial reporting environments.

4. Hypothesis Development

4.1 IRS Monitoring and State Corporate Income Tax Collections

It may seem that the positive relation between contemporaneous federal tax monitoring and SCIT revenues is obvious. Hoopes *et al*. (2012) show that higher levels of federal tax monitoring are related to higher contemporaneous cash effective tax rates, which suggests higher levels of federal taxable income in that year. As discussed earlier, states generally begin their SCIT computations with federal taxable income. Therefore, it would seem that SCIT revenues should have a positive relation with IRS enforcement. However, there are several reasons why this may not be the case.

First, filing groups for federal and state tax returns often differ. While firms are generally required to file on a consolidated basis for federal purposes, states vary in their filing rules. Some states require each taxable

[7] In some states the sales factor has an additional restriction known as the "throwback rule." The throwback rule states that if the firm is not taxable in the destination state of a sale, that sale is thrown back into the numerator of the sales factor of the origination state.

entity to file a separate tax return. Even states that require or allow related taxpayers to file in a group may require that the group include only entities with activity in that state, causing the group to differ from the federal consolidated group. Consider a situation where, for federal tax purposes, one entity can use the losses of a related entity to offset its taxable income, resulting in no federal corporate income tax liability. Under state group filing rules, those entities could be required to file separate returns, prohibiting the profitable entity from offsetting its taxable income with the losses of the other entity, resulting in a SCIT liability for the profitable firm. It is not readily obvious how all federal audit adjustments will flow through to the state returns.

Second, it is not clear that the items investigated by an IRS audit will always affect state taxes. For example, enforcement efforts that are related to federal-level credits could have no effect on state returns for that year. In addition, as discussed earlier, there are many complexities involved in the SCIT computations that are unrelated to the federal tax return.

Third, in my analyses I examine *contemporaneous* IRS enforcement and SCIT revenues. IRS audits typically take multiple years to complete (Gleason and Mills, 2002). After a firm has determined its final federal audit adjustments it must recalculate its SCIT liabilities, report to the state(s), and often undergo a state audit process. This protracted timeframe means that it is unlikely that IRS enforcement activity is related to SCIT revenues in that same year through the mechanical SCIT calculation process.

Therefore, it is not immediately clear what connection, if any, exists and so I state my hypothesis in the null:

H1: IRS enforcement has no effect on state corporate income tax collections.

4.2 IRS Monitoring in a Post-FIN 48 World

FASB Interpretation No. 48, "Accounting for Uncertainty in Income Taxes" (FIN 48), was intended to clarify the treatment of uncertainty related to the accounting for taxes under Statement 109. Statement 109 did not provide a specific threshold or guideline to be used in situations where the tax outcome was uncertain. Under FIN 48, a firm can only "recognize the financial statement effects of an uncertain tax position when it is more likely than not, based on the technical merits, that the position will be sustained upon examination." The guidance sets forth a two-step process. The first step is to determine whether a particular tax position satisfies the "more likely than not" threshold. The second step is the measurement of the position that satisfies that threshold, which involves an analysis of different possible outcomes and their related probabilities. FIN 48 was effective for public entities for fiscal years beginning after December 15, 2006.[8]

Mills *et al*. (2010) model how the interaction between the government and public corporate taxpayers changed with the move to the mandatory disclosure environment of FIN 48. They show that FIN 48 makes the government weakly better off, but that taxpayers are not necessarily harmed. However, certain taxpayers will be worse off because they will be deterred from entering into tax avoidance transactions. Based on that model Gupta *et al*. (2013) hypothesize that SCIT payments by firms "will increase because taxpayers will claim fewer weak tax positions once the tax authority observes more information about the strength of tax positions." The authors find that firm state effective tax rates and SCIT collections do in fact increase around implementation of FIN 48. However, although the authors control for several state tax policies, they do not control for enforcement activity at the state or federal level. Without controlling for enforcement, it is impossible to know if the increased effective tax rates were related to FIN 48, enforcement activities, or both. If, as Gupta *et al*. (2013) posit, corporate taxpayers became more conservative on their tax returns around FIN 48, then it is possible that there was decreased opportunity for IRS enforcement to affect SCIT revenues. This leads to my second hypothesis:

H2: The relation between IRS enforcement and state corporate income tax revenue was reduced with the implementation of FIN 48.

[8] See Blouin *et al*. (2007) for a more detailed discussion of FIN 48.

5. Sample and Empirical Specification

5.1 Data and Sample

I hand-collected data on 43 states that impose a SCIT over the time period 1995 through 2010. Nevada, South Dakota, Washington, and Wyoming are excluded because they do not impose a SCIT. Michigan and Texas are excluded because the relevant corporate taxes in those states are not based solely on income. Ohio is included only for years 1995 through 2004 because starting July 1, 2005, the state transitioned to a Commercial Activity Tax that is based on gross receipts rather than income.[9] Alaska is excluded because its revenue data does not include pure income tax data. The state tax policy variables are adopted from Gupta *et al.* (2009). My enforcement variables are based on Guedhami and Pittman (2008). See Appendix A for variable definitions and data sources.

Table 1 provides descriptive statistics on all 682 observations. The statistics on the tax policy variables are generally consistent with Gupta *et al.* (2009). Any differences are the result of the later time period used in this paper. For example, Gupta *et al.* (2009) find a mean value of 44.9 for the sales factor weight (indicating that, on average, the sales factor accounted for 44.9% of the overall apportionment factor) in their sample (covering 1982 through 2002), compared to 53.8 in my sample (covering 1995 through 2010).[10] This difference is indicative of the trend during this time period of states moving from the traditional equally-weighted three factor formula to a more heavily weighted sales factor or even a single sales factor apportionment formula.[11]

TABLE 1. Descriptive Statistics for the State-Level Variables, 43 States, 1995–2010

Variable	Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum
SCIT	762,204	1,300,693	28,273	188,016	359,757	789,655	11,849,097
SCIT_GSP	0.0033	0.0015	0.0006	0.0023	0.003	0.004	0.0105
SALES	0.664	0.473	0	0	1	1	1
TXRATE	0.076	0.016	0.046	0.064	0.075	0.088	0.12
FLOWTHRU	0.696	0.079	0.436	0.64	0.7	0.757	0.885
FEDBASEGSP	0.062	0.052	-0.004	0.029	0.045	0.076	0.424
UNEMP	5.14	1.74	2.27	4	4.85	5.74	12.43
LN_POP	15.13	0.93	13.29	14.41	15.23	15.67	17.44
IRS_AUDIT	0.014	0.006	0.006	0.009	0.012	0.02	0.026
EMP	0.0149	0.0027	0.0114	0.0132	0.0146	0.0159	0.0228
REV_AGT	0.0023	0.0004	0.0019	0.002	0.0021	0.0025	0.0034
CI	0.0005	0.0001	0.0004	0.0004	0.0005	0.0006	0.0007
FRAUD	0.00005	0.00003	0.00002	0.00002	0.00004	0.00007	0.00012

682 observations; variables are calculated as detailed in Appendix A.

[9] Inferences are unchanged when I remove the Ohio observations from my sample.

[10] The mean of the *SALES* variable in Table 1 is 0.664 rather than 0.538 since *SALES* is an indicator variable rather than the actual sales factor weight. See footnote 12 and Appendix A.

[11] This trend has been widely documented, including Mazerov (2001) and Harrie (2008).

5.2 Empirical Specification

5.2.1 Base Regression Model

Following Gupta *et al.* (2009), my main specification is:[12]

$$SCIT/GSP_{it} = \alpha_i + \beta_1 TXRATE_{it} + \beta_2 SALES_{it} + \beta_3 FLOWTHRU_{it} + \beta_4 FEDBASEGSP_{it} + \beta_5 UNEMP_{it} + \beta_6 LN\_POP_{it} + \beta_7 IRS\_AUDIT_t + S_i + \varepsilon_{it} \quad (1)$$

TXRATE is the top statutory marginal state corporate income tax rate for the year. *SALES* is an indicator variable equal to one if the weight on the sales factor is 50% or greater in a state in a given year, zero otherwise.[13] Following the results of Gupta *et al.* (2009), I expect β_1 to be positive and β_2 to be negative. *FLOWTHRU* is defined as the proportion of total business federal income tax returns filed in a state that are from flow-through entities. Gupta *et al.* (2009) use this variable in an attempt to control for tax planning activities through the use of flow-through entities, and so in line with their results I expect β_3 to be negative. *FEDBASEGSP* is estimated federal corporate taxable income reported from each state, scaled by gross state product. Federal corporate taxable income is estimated by taking IRS corporate income tax collections in each state and grossing them up by the top marginal federal corporate income tax rate in that year. This variable should help control for any portion of the relation between IRS enforcement and SCIT collections that is driven by the mechanical process of using federal taxable income in calculating SCIT liabilities, so I expect β_4 to be positive. *UNEMP*, the state unemployment rate, is included as a control for general economic conditions in the state. Higher unemployment rates indicate poor economic conditions, which may be associated with lower corporate revenues and lower SCIT collections. Therefore, I expect β_5 to be negative. *LN_POP* is the natural log of the state's population, included as a control for the general size of the state. Based on the results of Gupta *et al.* (2009) I expect β_6 to be positive. Considering H1, I make no predictions on the sign or significance of β_7.

IRS_AUDIT, the enforcement variable in my main specification, is the percentage of federal corporate tax returns filed that are audited by the IRS. *EMP* is equal to the number of permanent IRS employees at the end of the year, scaled by the total number of corporate returns filed during the year. *REV_AGT* is equal to the number of IRS revenue agents at the end of the year, scaled by the total number of corporate returns filed during the year. *CI* is equal to the number of IRS criminal investigators at the end of the year, scaled by the total number of corporate returns filed during the year. *FRAUD* is equal to the number of corporate fraud assessments for the year, scaled by the total number of corporate returns filed during the year. As expected, Table 2 shows that the enforcement variables are highly correlated with each other, suggesting that they are all capturing a similar construct. All enforcement variables are adopted from Guedhami and Pittman (2008) and vary only by year, not by state.

[12] Gupta *et al*. (2009) find that endogeneity is an issue with the *SALES* (the percentage weight of the sales factor in the state's apportionment formula) and *TXRATE* (top marginal state corporate tax rate) variables in their data. Using the same version of the Hausman test used in their analysis I find that the endogeneity of *SALES* and *TXRATE* is not an issue with my sample. As a robustness check on my results, I also conducted my analysis using two stage least squares (2SLS). As part of that analysis I followed the guidance in Larcker and Rusticus (2010), including the use of the overidentification test and acceptable levels of F-statistics, to confirm that I used appropriate instruments. Inferences are unchanged whether I use the 2SLS or OLS specification. Therefore, I focus on the OLS specification in this paper.

[13] Gupta *et al*. (2009) include *SALES* in their model as the value of the sales factor weight. I use an indicator variable since it is not a truly continuous variable, but rather has observations mainly clustered around 33%, 50%, and 100%. Inferences are unchanged when *SALES* is included in the same form as in Gupta *et al*. (2009).

TABLE 2. Correlations* Between the Variables

	SCIT_GSP	TXRATE	SALES	FLOWTHRU	FEDBASEGSP	UNEMP	LN_POP	IRS_AUDIT	EMP	REV_AGT	CI	FRAUD
SCIT_GSP		0.34	0.24	-0.17	0.10	-0.12	-0.05	0.21	0.08	0.15	0.09	0.15
TXRATE	0.42		0.21	-0.15	0.18	-0.10	-0.10	0.04	0.05	0.05	0.06	0.05
SALES	0.25	0.21		0.04	0.01	0.15	0.45	-0.05	-0.07	-0.07	-0.07	-0.07
FLOWTHRU	-0.18	-0.15	0.02		-0.03	0.31	-0.02	-0.51	-0.68	-0.68	-0.75	-0.68
FEDBASEGSP	0.11	0.19	0.14	-0.13		-0.14	-0.04	0.07	0.00	0.04	0.02	0.05
UNEMP	-0.10	-0.06	0.16	0.24	-0.19		0.24	-0.07	-0.13	-0.16	-0.28	-0.20
LN_POP	0.00	-0.08	0.45	0.00	0.27	0.23		-0.02	-0.03	-0.03	-0.03	-0.03
IRS_AUDIT	0.29	0.03	-0.04	-0.39	0.15	-0.06	-0.03		0.63	0.80	0.69	0.90
EMP	0.07	0.07	-0.07	-0.76	0.09	-0.17	-0.05	0.48		0.95	0.95	0.85
REV_AGT	0.19	0.06	-0.06	-0.67	0.14	-0.22	-0.04	0.65	0.86		0.94	0.93
CI	0.10	0.07	-0.07	-0.78	0.12	-0.28	-0.05	0.43	0.97	0.84		0.91
FRAUD	0.15	0.07	-0.07	-0.75	0.15	-0.30	-0.05	0.55	0.91	0.87	0.92	

* Pearson (above diagonal) and Spearman (below diagonal) correlation coefficients for all variables used in primary analyses.

I also include state indicator variables. I do not include year indicator variables because the enforcement variables vary only by year, not by state, leading to multicolinearity in a model with year indicator variables.

In order to address concerns related to variable scaling I also tested a log specification, again following Gupta *et al*. (2009):

$$LN\_SCIT_{it} = \alpha_i + \beta_1 TXRATE_{it} + \beta_2 SALES_{it} + \beta_3 FLOWTHRU_{it} + \beta_4 FEDPERCAP_{it} + \beta_5 UNEMP_{it} + \beta_6 LN\_GSP_{it} + \beta_7 IRS\_AUDIT + \varepsilon_{it} \quad (2)$$

LN_SCIT is the natural log of corporate income tax revenue collections in a given state in a given year. *FEDPERCAP* is estimated federal corporate taxable income reported from each state, scaled by state population. *LN_GSP* is the natural log of the gross state product. All other variables are the same as in the main specification. The results (untabulated) under this log specification (using *IRS_AUDIT* and the other federal enforcement variables) are qualitatively similar to the results of the main specification. Therefore, for the remainder of the paper I focus my discussion on the main specification, equation (1).

5.2.2 FIN 48

To address whether the relation between IRS monitoring and SCIT decreased with the implementation of FIN 48 I create an indicator variable, *FIN48*, to denote the period when FIN 48 was in effect.[14] I also interact *IRS_AUDIT* with *FIN48* (*IRS*FIN48*) to capture the effect of federal enforcement in the post-FIN 48 period:[15]

$$SCIT/GSP_{it} = \alpha_i + \beta_1 SALES_{it} + \beta_2 TXRATE_{it} + \beta_3 FLOWTHRU_{it} + \beta_4 FEDBASEGSP_{it} + \beta_5 UNEMP_{it} + \beta_6 LN\_POP_{it} + \beta_7 IRS\_AUDIT_t + \beta_8 FIN48_t + \beta_9 IRS*FIN48_t + \varepsilon_{it} \quad (3)$$

6. Results

The results for the main specification of equation (1) are included in Panel A of Table 3. Panel B includes the results of equation (1) using the other proxies for IRS enforcement. Looking at Panels A and B, the results are

[14] FIN 48 was effective for all public entities for fiscal years beginning on or after December 31, 2006. Therefore, *FIN48* equals one for years 2007 through 2010 and equals zero for all other years.

[15] As with equation (1), in alternate specifications of the model I include the other IRS enforcement variables and interact them with *FIN48*.

generally consistent across specifications. *TXRATE* is significantly positive across specifications, as is *SALES*. In contrast, Gupta *et al.* (2009) find that *SALES* has a statistically significant negative coefficient.[16]

TABLE 3. OLS Regression Results for Alternative Specifications of Equation 1

Variable (Predicted Sign)	Panel A	Panel B			
	Orig. Specification	Alternate Proxies for the Enforcement Variable			
	IRS_AUDIT	EMP	REV_AGT	CI	FRAUD
TXRATE (+)	0.0259 ***	0.0240 **	0.0261 ***	0.0234 **	0.0246 ***
	(0.009)	(0.0095)	(0.0089)	(0.0096)	(0.0093)
SALES (-)	0.0006 ***	0.0006 ***	0.0007 ***	0.0006 ***	0.0006 ***
	(<0.001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
FLOWTHRU (-)	0.0071 ***	0.0058 ***	0.0130 ***	0.0065 ***	0.0108 ***
	(0.001)	(0.0013)	(0.0012)	(0.0019)	(0.0011)
FEDBASEGSP (+)	0.003 ***	0.007 ***	0.0044 ***	0.0067 ***	0.0041 ***
	(0.001)	(0.0012)	(0.0011)	(0.0012)	(0.0011)
UNEMP (-)	-0.0002 ***	-0.0002 ***	-0.0003 ***	-0.0002 ***	-0.0002 ***
	(<0.001)	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
LN_POP (+)	-0.0022 ***	-0.0022 ***	-0.0023 ***	-0.0023 ***	-0.0023 ***
	(<0.001)	(0.0002)	(0.0002)	(0.0003)	(0.0002)
ENFORCEMENT (?)	0.0906 ***	0.1135 ***	1.8240 ***	3.6398 ***	19.7332 ***
	(0.007)	(0.0284)	(0.1732)	(1.2313)	(2.0430)
INTERCEPT	0.0305 ***	0.0302 ***	0.025 ***	0.0303 ***	0.0297 ***
	(0.004)	(0.0037)	(0.0036)	(0.0037)	(0.0038)
R^2	0.7926	0.7438	0.7866	0.7406	0.7750
No. of Observations	682	682	682	682	682

Robust standard errors in parentheses. State indicator variables are included in the model, but the coefficients are not included here for sake of brevity. ***, **, and * represent 1%, 5%, and 10% significance levels respectively.

It seems possible that the differences in coefficients on *SALES* could be driven by the trends in apportionment formulae over the past 30 years. The mean sales factor weight in 2010 is 61.1, while the mean sales factor weight in 1995 is 49.3. The mean value for 1982 would undoubtedly be even lower. It is possible that the first wave of states that moved to a more heavily weighted sales factor paid for that in the form of lower SCIT revenues (compared to states that kept the equal weighted factor). As more states have increased the weight on the sales factor, somewhat leveling the apportionment playing field, it is possible that there is no longer a detrimental effect on SCIT revenues in comparison to other states.

The results of my model along with the results in Gupta *et al.* (2009) suggest that the landscape of SCIT policy has changed. Although the ultimate economic effects (both in terms of SCIT revenues and investment and employment within the state) of modifying a state's apportionment factors have long been debated (Mazerov (2001), Hamm and Verma (2002), Harrie (2008), and Swenson (2011), among others), moving to a more heavily weighted sales factor has long been cited as a tool to encourage economic development in a state.

[16] When I run the regression in equation (1) without the enforcement variable (similar to the specification used in Gupta *et al.*, 2009) I get similar results in terms of sign and significance as in Panel A, except that the coefficient on *FLOWTHRU* is insignificant.

In its Multistate Audit Technique Manual, the California Franchise Tax Board states, "[t]o promote investment within our state, California moved to a double weighted sales factor...." The difference in results between the time period examined by Gupta *et al*. (2009) and the time period examined in this paper suggests that states need to be careful to consider the current, not historical, state tax landscape (as well as any imminent changes) when determining how state tax policy changes will affect their revenues.

Also in contrast to Gupta *et al*. (2009), *FLOWTHRU* has a positive significant coefficient (significantly negative in Gupta *et al*., 2009) and *LN_POP* has a negative significant coefficient (significantly positive in Gupta *et al*., 2009). The change in coefficient on *FLOWTHRU* may be attributable to the expanding use of LLCs. The mean value of *FLOWTHRU* is .59 for 1995, but grew to .79 in 2010, indicating increased use of flow-through entities, including LLCs, during my sample period. Many states did not enact LLC legislation until the early to mid-1990's, towards the end of the sample period in Gupta *et al*. (2009). With the increase in LLCs in my sample period, used not only for tax planning purposes, but also for legal liability purposes, *FLOWTHRU* may now be capturing the level of business activity in the state rather than tax planning activity. Without access to data for the earlier part of the sample used in Gupta *et al*. (2009), it is difficult to draw conclusions about these differences in results.[17]

Consistent with the notion that higher federal taxable income is related to higher SCIT revenues due to the incorporation of federal taxable income into SCIT calculations, the coefficient on *FEDBASEGSP* is significantly positive in all specifications. However, even after controlling for the level of federal corporate income tax collections from each state, the coefficients on all of the enforcement variables are positive and statistically significant at the .01 level.[18,19]

The results of equation (3) are displayed in Table 4 and are generally consistent with the earlier results. Again, while the coefficient on *FEDBASEGSP* is significantly positive across all specifications, the coefficients on all of the various enforcement proxy variables also remain positive and statistically significant. The coefficient on *FIN48* is positive and significant in four of the five specifications, suggesting that SCIT revenues increased after the implementation of FIN 48. This is consistent with the results in Gupta *et al*. (2013). The coefficients on the interactions between the enforcement proxies and *FIN48* are negative across all specifications and they are statistically significant in the models with *IRS_AUDIT, REV_AGT*, and *FRAUD* as the enforcement variable. The results are consistent with H2, indicating that the positive relation between IRS enforcement and SCIT revenues was reduced by the implementation of FIN 48.

[17] I have access to tax rate and apportionment information going back to only the mid-1990s, prohibiting me from analyzing the full period examined in Gupta *et al*. (2009).

[18] Since the enforcement variables are only known ex post, the relation between SCIT revenues and enforcement in equation (1) assumes that taxpayers somehow anticipate the level of enforcement for the year. In a sensitivity check (results untabulated) I use one year lagged enforcement instead of current year enforcement and inferences remain unchanged.

[19] I also ran tests with equation (1), but with an additional interaction term between *FEDBASEGSP* and the given enforcement variable (e.g. *FEDBASEGSP*IRS_AUDIT*) to examine the relation between the mechanical (*FEDBASEGSP*) and non-mechanical (enforcement variables) aspects of the federal corporate income tax process. The results (untabulated) are inconclusive. Across all specifications the coefficient on the interaction term is positive, but the coefficient is significant only in the specifications where the enforcement variable is related to IRS employment (EMP, REV_AGT, and CI). In addition, in those employment-related specifications the coefficient on *FEDBASEGSP* becomes negative.

TABLE 4. Regression Results for Alternative Specifications of Equation 3

Variable (Predicted Sign)	Panel A	Panel B			
	Orig. Specification	Alternate Choices for the Enforcement Variable			
	IRS_AUDIT	EMP	REV_AGT	CI	FRAUD
TXRATE (+)	0.0270 ***	0.0276 ***	0.0278 ***	0.0274 ***	0.0271 ***
	(0.008)	(0.0084)	(0.0081)	(0.0085)	(0.0085)
SALES (-)	0.0005 ***	0.0005 ***	0.0006 ***	0.0005 ***	0.0005 ***
	(<0.001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
FLOWTHRU (-)	0.0034 ***	0.0002	0.0079 ***	0.0041 **	0.0046 ***
	(0.001)	(0.0014)	(0.0017)	(0.0019)	(0.0014)
FEDBASEGSP (+)	0.0033 ***	0.0052 ***	0.0033 ***	0.0048 ***	0.0038 ***
	(0.001)	(0.0011)	(0.0010)	(0.0011)	(0.0011)
UNEMP (-)	-0.0002 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***	-0.0003 ***
	(<0.001)	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
LN_POP (+)	-0.0021 ***	-0.002 ***	-0.0022 ***	-0.0021 ***	-0.0021 ***
	(<0.001)	(0.0002)	(0.0003)	(0.0002)	(0.0003)
ENFORCEMENT (?)	0.0682 ***	0.0853 ***	1.5113 ***	5.1274 ***	14.1394 ***
	(0.009)	(0.0283)	(0.2064)	(1.2282)	(2.0993)
FIN48 (+)	0.0053 ***	0.0019 *	0.0049 ***	0.0031	0.0031 ***
	(0.001)	(0.0011)	(0.0014)	(0.0019)	(0.0008)
*ENF*FIN48* (-)	-0.3819 ***	-0.0699	-2.1311 ***	-4.9441	-104.889 ***
	(0.120)	(0.0878)	(0.7196)	(4.9600)	(38.5925)
INTERCEPT	0.0314 ***	0.0312 ***	0.027 ***	0.0289 ***	0.0309 ***
	(0.004)	(0.0037)	(0.0037)	(0.0036)	(0.0038)
R^2	0.8031	0.7864	0.8096	0.7906	0.8004
No. of Observations	682	682	682	682	682

Robust standard errors in parentheses. State indicator variables are included in the model, but the coefficients are not included here for sake of brevity. ***, **, and * represent 1 percent, 5 percent, and 10 percent significance levels respectively.

It appears that as state tax avoidance behavior by firms decreased, the relationship between federal-level enforcement and SCIT collections was weakened. These results suggest that enforcement and financial regulation should be examined together, not just separately.

7. Conclusion

In this paper I examine the relationship between tax enforcement at the federal level and SCIT revenues. My results suggest that higher levels of federal tax enforcement are associated with higher levels of SCIT collections, even after controlling for applicable tax rates and sales factor weighting. These results can be viewed in the context of Desai *et al.* (2007), with state governments as an additional set of outside stakeholders benefiting from IRS enforcement. This relationship seems to be the result of both mechanical and non-mechanical aspects of the SCIT environment. These results hold for a variety of proxies for IRS enforcement, including IRS corporate income tax audit rates, various IRS employment levels, and the number of fraud proceedings.

In examining how enforcement matters in a post-FIN 48 environment, I find that the positive relation between IRS monitoring and SCIT revenues is reduced during the time period that FIN 48 has been in effect for public companies. This study can help inform policymakers as they discuss solutions to the current financial crises faced by states. The results in this paper suggest that such discussions should extend beyond just state tax policy choices to consider federal-level enforcement, the financial reporting environment, and the interaction of these various factors.

APPENDIX
Variable Definitions (sources in parentheses)

Dependent Variables

SCIT	State corporate income tax revenue collections (Census Bureau)
SCIT/GSP	SCIT divided by GSP

Tax Policy Variables

SALES	Indicator variable equals 1 if weight on sales factor is 50% or greater, zero otherwise (Commerce Clearing House)
TXRATE	Top statutory marginal state corporate income tax rate (Commerce Clearing House)

Enforcement Variables

IRS_AUDIT	Percentage of corporate returns filed that were audited by an IRS revenue agent (TRAC)
EMP	Number of permanent IRS employees at the end of the year, scaled by the total number of corporate returns filed (IRS)
REV_AGT	Number of IRS revenue agents at the end of the year, scaled by the total number of corporate returns filed (IRS)
CI	Number of IRS criminal investigators at the end of the year, scaled by the total number of corporate returns filed (IRS)
FRAUD	Number of corporate fraud assessments for the year, scaled by the total number of corporate returns filed (IRS)

Other Variables

FLOWTHRU	Percentage of business returns filed by flow-through entities in a particular state, measured as the number of partnership and S corporation returns filed divided by the total of partnership, S corporation, and C corporation returns filed (IRS)
FEDBASE	Federal corporate income tax collections by state, grossed up by the top marginal tax rate in effect for the year (IRS)
FEDBASEGSP	FEDBASE divided by GSP
GSP	Gross state product (Bureau of Economic Analysis)
UNEMP	State unemployment rate (Bureau of Labor Statistics)
POP	State population (Bureau of Economic Analysis)
LN_POP	Natural log of POP
FIN48	Indicator variable that equals 1 for years 2007 through 2010, 0 otherwise
IRSFIN48	Interaction of IRS_AUDIT and FIN48

References

Blouin, J., C. Gleason, L. Mills, and S. Sikes. (2007). What can we learn about uncertain tax benefits from FIN 48? National Tax Journal 60, 521–535.

California Franchise Tax Board Multistate Audit Technique Manual. December 2002.

Desai, M, A. Dyck, and L. Zingales. (2007). Theft and taxes. Journal of Financial Economics 84, 591–623.

Dyreng, S., B. Lindsey, and J. Thornock. (2013). Exploring the Role Delaware Plays as a Domestic Tax Haven. Journal of Financial Economics 108, 751–772.

Financial Accounting Standards Board (FASB). (2006). Accounting for Uncertainty in Income Taxes, an Interpretation of FASB Statement No. 109. FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.

Gleason, C. and L. Mills. (2002). Reporting and Recording of Contingent Tax Liabilities. The Accounting Review 77, 317–342.

Goolsbee, A. and E. Maydew. (2000). Coveting Thy Neighbor's Manufacturing: The Dilemma of State Income Apportionment. Journal of Public Economics 75, 125–143.

Guedhami, O. and J. A. Pittman. (2008). The importance of IRS monitoring to debt pricing in private firms. Journal of Financial Economics 90, 38–58.

Gupta, S., J. Moore, J. Gramlich, and M.A. Hofmann. (2009). Empirical Evidence on the Revenue Effects of State Corporate Income Tax Policies. National Tax Journal 62, 237–267.

Gupta, S. and D. Lynch. (2012). The Effects of Changes in State Tax Enforcement on Corporate Income Tax Collections. Working paper, Michigan State University.

Gupta, S., L. Mills, E. Towery. (2013). The Effect of Mandatory Financial Statement Disclosures on Tax Reporting and Collections: The Case of FIN 48 and Multistate Tax Avoidance. Working paper, The University of Texas at Austin.

Hamm, W., and A. Verma. (2002). "If California Adopted the Single-Factor (Sales) Apportionment Formula, What Would be the Impact?" State Tax Notes 10 June.

Hanlon, M., J. Hoopes, and N. Shroff. (2012). The Effect of Tax Authority Monitoring and Enforcement on Financial Reporting Quality. Working paper, Massachusetts Institute of Technology.

Harrie, D. (2008). "Utah Enacts Corporate Income Apportionment Change to Attract Services." State Tax Notes 24 March.

Hoopes, J., D. Mescall, and J. A. Pittman. (2012). Do IRS Audits Deter Corporate Tax Avoidance? The Accounting Review 87, 1603–1639.

Klassen, K., and D. Shackelford. (1998). State and Provincial Corporate Tax Planning: Income Shifting and Sales Apportionment Factor Management. Journal of Accounting and Economics 25, 385–406.

Larker, D. F., and T.O. Rusticus. (2010). On the use of instrumental variables in accounting research. Journal of Accounting and Economics 49, 186–205.

Lightner, T. (1999). The Effect of the Formulary Apportionment System on State-Level Economic Development and Multijurisdictional Tax Planning. Journal of the American Taxation Association 21, 42–57.

Mazerov, M. (2001). "The Single-Sales-Factor Formula: A Boon to Economic Development or a Costly Giveaway?" State Tax Notes 21 May.

Mills, L., L. Robinson, and R. Sansing. (2010). FIN 48 and Tax Compliance. The Accounting Review 85, 1721–1742.

Nolan, K. and M. Peters. (2013). "Illinois Yanks Bond Amid Pension Woes." Wall Street Journal 31 January.

State Budget Crisis Task Force. (2012). Report of the State Budget Crisis Task Force.

Swenson, C. (2011). On the Effectiveness of Single Sales Factors for State Taxation. Working paper, The University of Southern California.

Walsh, M. W. (2011). "A Path Is Sought for States to Escape Their Debt Burdens." New York Times 20 January.

4

Lessons From Other Tax Adminstrations

Christian

Koh ◆ Merriman ◆ Vielma

Cleland

Why Evasion Under a National Sales Tax Would Explode the Tax Gap: Lessons Learned from the States

P. Cary Christian,
Georgia Southern University

Conventional wisdom has it that evasion is nearly nonexistent with respect to consumption taxes, perhaps partially explaining why evasion research in the United States has focused almost entirely on income tax evasion. More recent research on both consumption and income tax enforcement indicates that the conventional wisdom suffers from a number of poor assumptions that, upon closer inspection, have no grounding in evidence. If a national sales tax were instituted to replace the income and payroll tax systems currently in place, the acceptance of these assumptions in designing the enforcement mechanisms for the national system would result in a dramatic increase in the tax gap, substantially reduced revenues, and ultimately a requirement to increase rates to account for revenue losses from evasion. Ultimately, it would require a large expansion of Internal Revenue Service (IRS) compliance enforcement efforts rather than contraction or, as has been proposed in some cases, outright elimination of the IRS. I will address the following **four assumptions** in this paper:

- Evasion under a sales tax system is nonexistent,
- The amount of sales taxes collected by entities with a propensity toward theft of sales tax is immaterial,
- State estimates of sales tax compliance rates can be used to estimate sales tax lost to evasion or theft, and
- Under a national sales tax there would be no need for the Internal Revenue Service.

Methodology

The discussion and conclusions in this paper were developed based on qualitative, quantitative, and quasi-experimental research carried out over an 18-month period within the Department of Revenue (DOR) of Florida—a major state that relies heavily on the sales tax for revenue (59 percent of total taxes). This research project was composed of four major components:

1. Interviews of DOR field personnel were conducted, accompanied by review of current and historical management data, reports, audit and investigative results, and other relevant documents covering a period of 10 years. This portion of the project was designed to construct a baseline for compliance activities currently carried out by DOR personnel and to develop an understanding of the existing enforcement paradigm. Additionally, interviews of subjects in criminal investigations over a 10-year period were reviewed. Finally, interviews of subjects and potential subjects were conducted during the quasi-experimental portion of this study to provide a profile or typology of those who evade or steal sales taxes collected.
2. Since there is no third-party verification of sales taxes collected, an observational survey was developed to test the efficacy of the use of tax evasion predictors in a sales tax system. The survey was designed as a proxy for DOR personnel's evaluation of potential fraud indicators through the review of available reporting and compliance data supplemented by other readily-available third-party and observational data. The purpose of this survey was to test the ability to accurately predict fraudulent activities with a minimal amount of data gathering from outside the DOR.

3. An analysis of the use of task force operations was conducted to gauge the impact on enforcement outcomes of magnifying available enforcement resources through cooperative operations with other federal and state agencies.
4. In the quasi-experimental portion of this project, detailed third-party data were accumulated in several industries to identify potential sales tax evasion and theft more accurately. This quasi-experimental activity was designed as a pilot test of a Targeted Industry Enforcement Program. Companies identified as potentially involved in fraudulent activities were contacted by letter and were provided with detailed information providing evidence of a compliance problem. Those companies were required to respond and provide information to verify that their reporting was accurate or enter into agreements to rectify their compliance issues. These issues were resolved through a variety of actions including referral to collections, full audit, or criminal investigation and prosecution. Interactions with these subjects were documented and contributed to the development of the evader typology mentioned above.

A discussion of the specific results of each part of this project goes well beyond the scope of this paper. However, much of what was learned in carrying out the above research informs the conclusions reached in this paper. Complete details related to the above study are available (Christian, 2010).

Assumption One: Evasion Under a Sales Tax System Is Nonexistent

Sales tax enforcement efforts in the states are structurally similar to the income tax enforcement model used by the IRS, but adapted to sales tax enforcement. The same voluntary reporting structure is used and enforced through punishing deterrence policies linked to delinquency identification and audit capabilities. This adaptation of the income tax compliance enforcement system to sales tax systems is the beginning of the problem: sales tax compliance enforcement issues are very different from income tax enforcement issues. Income taxes are relatively easier to enforce due to the level of third-party reporting required in such systems. Based on data from the IRS 2007 Statistics of Income, 82.59 percent of all income reported on tax returns is subject to some type of third-party verification (Internal Revenue Service, 2007). Based on IRS tax-gap estimates, the majority of the tax gap is related to income that is not subject to such third-party verification. With respect to the tax gap related to individual income tax underreporting, the net misreporting percentage with respect to amounts subject to little or no information reporting is 56 percent and the misreporting percentage with respect to amounts subject to some information reporting is 11 percent. By contrast, the misreporting percentages for amounts subject to substantial information reporting and amounts subject to substantial information reporting *and* withholding are 8 percent and 1 percent, respectively (Internal Revenue Service, 2012). Table 1 summarizes the impact of third-party verification with respect to individual income tax returns based on both reported and estimated unreported revenues. Unreported income is based on the individual income tax underreporting gap of $190 billion divided by an assumed marginal tax rate of 28 percent.

TABLE 1. Impact of Third-Party Verification of Income in Income Tax Enforcement

Amount Reported by Taxpayers	Estimated Unreported	Amount Verifiable by Third-Party Information	Amount Verifiable as a Percent of Amount Reported	Amount Verifiable as a Percent of Amount Reported + Unreported
$8.8 trillion	$.7 trillion	$7.3 trillion	82.59%	76.84%

Research indicates third-party verification plays an important role in compliance enforcement on several levels. The 1982 Taxpayer Compliance Measurement Program (TCMP) found a positive correlation between underwithholding of income tax and a subsequent underreporting of the tax liability (Chang & Schultz, 1990). Martinez-Vazquez, Harwood, and Larkins (1992) observed that people with liquidity problems were less likely to pay commercial debts and theorized that liquidity problems may have the same effect on the behavior of taxpayers. Using experimental methods, they found that if the possibility of evading taxes in a safe manner existed, a near-majority of people would take that chance, and the proportion of individuals choosing to evade who were in an illiquid position was significantly larger. Blanthorne (2000) found that taxpayers who have the

opportunity to underreport income actually underreported more, in both frequency of underreporting and in the amount underreported, and had lower tax reporting ethics than taxpayers who did not have the opportunity to underreport. Carnes and Englebrecht (1995) found that tax compliance increases as the visibility of income to the taxing authority increases. Antonides and Robben (1995) found that the probability of tax evasion was related to the opportunity available to the taxpayer to conceal income.

There is a generally implied assumption that there is virtually no opportunity for evasion with respect to sales tax because it is collected, and thus verified, by third-party business entities. This assumption misses the point: in a retail sales tax system there is no mechanism to verify that all of the sales tax collected is actually remitted to the government. Therefore, the real issue is not sales tax evasion; it is sales tax theft by the parties who collect the tax as an agent of the state.[1]

In a sales tax system, retail businesses become collection agents for the government and agency theory yields important insights in the analysis of sales tax compliance enforcement efforts. Two primary concerns addressed by agency theory are the problems of adverse selection and moral hazard (Droege & Spiller, 2009). Adverse selection occurs when a principal selects an inappropriate agent based on false or inaccurate information. Moral hazard refers to the situation where the agent does not provide appropriate effort to achieve the goals of the principal. Agency theory assumes adverse selection can be controlled if the principal has access to all available information, and that the required information can be obtained for a price. The principal must balance the cost of acquiring the information needed with the potential gain from selecting an appropriate agent. Moral hazard can be controlled through either behavioral contracts designed to control the activities of the agent, or through outcome-based contracts that are designed to align the goals of the principal and agent and allow the principal to monitor specific outcomes produced by the agent rather than the agent's activities (Droege & Spiller, 2009).

The state will encounter problems with both adverse selection and moral hazard in its dealings with its collection agents and must address several unique problems:

First, the state cannot choose only those collection agents it wishes to work with. Businesses meeting the minimal requirements to obtain a license to collect sales tax become agents of the state. In current sales tax systems there are few reasons for disqualification and even new businesses owned by known tax cheats or their family members cannot generally be denied a license. As a result, the state is guaranteed an adverse selection problem regardless of the availability of information about potential agents.

Second, agency theory assumes a direct relationship between the principal and the agent that makes it possible to either monitor the agent's activities or require specified results based on contract specifications. Since all retail businesses engaged in the sale of taxable property become agents, direct monitoring is impossible: there are simply too many of them. Outcomes cannot be predicted since sales tax collected will be proportional to the sales the business is able to make and the mix of taxable and exempt items sold. Setting quotas for tax collected is not possible nor would it be good public policy. The agent's duty is simple: collect the tax on all sales of taxable items, properly account for the taxes collected, and remit the total amount of sales taxes collected to the government. To ensure the agent carries out this duty, the state must either allocate the resources required for direct monitoring, or develop methodologies for predicting outcomes at the single-business level for use in direct monitoring by exception.

Third, agency theory assumes that the agent will perform well based on incentives provided by the principal. In most state sales tax systems a minor collection allowance is given to the sales tax collection agent for its efforts. This incentive is normally a small percentage of the tax collected and is generally capped at a certain dollar amount per month. In terms of agency theory, these provisions provide no incentive at all. For example, in the State of Florida the collection allowance is capped at $30 regardless of whether the business collected hundreds of thousands of dollars in sales tax or just a few thousand. Miller and Whitford (2006) point out that

[1] Sales tax evasion does exist in the form of a business refusing to collect the tax, or by treating taxable items as nontaxable, usually as a form of achieving a competitive pricing advantage over competing retailers, and losses through this type of evasion can be substantial depending on the structure of the sales tax system.

incentives large enough to induce an appropriate level of effort in the agent are prohibitively expensive for public agencies, who rely more on coercive monitoring and sanctions.

Fourth, agency theory presumes that the principal is risk neutral and the agent is risk averse, a dynamic that does not exist in the sales tax principal-agent relationship. The risk presumptions of agency theory assume the incentives to the agent under the contract take into account the fact that the risk of failure is borne by the principal, whose portfolio is sufficiently diversified to absorb the risk of failure. In the sales tax collection contract, the agent has no incentive at all, other than possibly the incentive to steal as much tax money as possible. The agent bears no risk from subpar sales tax collections because of the lack of incentive payments. The agent is essentially working for no compensation. Moreover, the principal is not risk neutral because sales tax collections are a major component of state revenue.

Finally, agency theory assumes that information is a commodity, and that all information can be known and purchased. This is simply unrealistic in the sales tax collection principal-agent contract since the state must attempt to manage hundreds of thousands, or even millions, of separate retail establishments. The cost of acquiring the level of information required to effectively monitor every agent would be cost prohibitive as is indicated by the reliance of tax agencies on audit regimens rather than contract-management activities to enforce compliance.

These problems might be interpreted to mean that agency theory is only contingently valid in the case of the sales tax collection principal-agent relationship, and it is apparent that major presumptions of the theory are, indeed, invalid in this relationship. But agency theory is instructive nonetheless in that agency theory will predict very poor results for the state as principal for the reasons discussed above.

Indeed, an agency theory analysis of the state's position in the sales tax collection principal-agent relationship provides ample reasons why compliance is significantly lower with respect to retail consumption taxes versus income taxes, where third-party verification removes the opportunity to evade with respect to a large portion of taxable income. Agency theory informs the principal in the design of contracts that control adverse selection and moral hazard, but in the instant case, it is impossible for the principal to follow that guidance. Agency theory provides the warning that in the sales tax collection process, the state must find alternative means to monitor and enforce compliance because standard methods of controlling adverse selection and moral hazard will be of little use.

Agency theory makes an implicit assumption that the agent is dishonest, and this somewhat "politically incorrect" assumption contributes to the theory's power to predict poor results for the principal when control over agent dishonesty is not perfected (Bohren, 1998). It is difficult to acknowledge that so many people will choose to evade or steal tax monies when faced with the opportunity to do so, but the research on evasion continually supports this conclusion regardless of the theory or determinant of evasion under study. Corporations take a more realistic view of opportunistic crime and commit vast sums of money each year to internal controls, corporate security, employee screening, and outside consultants in an attempt to control employee theft or embezzlement. In spite of these efforts United States organizations still lose almost 5 percent of their revenues to fraud, an estimated $652 billion in 2006 (Ramamoorti, 2007). In the private sector it is deemed reasonable to assume that many will choose to steal, and it is expected that management will be forthright in accepting that premise and take steps to minimize the damage from theft. For unknown reasons we expect our citizens, who are the same individuals corporations pay to protect against, to be much more ethical and honest in their dealings with the government than they are with private-sector enterprises. The government response to this threat is to audit less than 1 percent of accounts each year for compliance, and at the state level, audit activity is so low that it is almost nonexistent. In an environment where there are no third-party controls to help enforce remittance of sales taxes collected, government cannot reasonably assume that sales tax theft will not occur regularly.

Given that the remittance of sales tax collected by business agents is not subject to any third-party verification, the aforementioned findings predict dire consequences for compliance under a sales tax regime. Yet Watrin and Ullman (2008) note that their work is the first to explicitly focus on the behavioral differences between compliance in the realm of income tax versus compliance related to sales taxes. The lack of specific research related to tax evasion in a consumption tax environment is surprising given the core differences between how

income taxes and consumption taxes are administered and given the lack of third-party controls in a sales tax system. Watrin and Ullman (2008) also found that none of the previous models developed for the analysis of consumption tax utility and optimal mixes of taxation regimes have even allowed for the possibility of tax evasion in a consumption tax setting, and there has never been a model that allows for evasion in income tax and consumption tax regimes at the same time. To provide context for this glaring omission, it must also be understood that research related to evasion in value-added consumption taxes (VAT) are inapplicable to studies of the American retail sales tax. VAT regimes are more easily enforced because the tax is collected at multiple stages during the production process. Taxes not collected at one stage can still be collected at a subsequent stage of production. More importantly, the VAT calculations at each stage of production leave a paper trail that makes it easier to find and prove evasion, and provides an incentive for proper reporting because of the built-in credit structure (Garner, 2005). This verification and incentive structure does not exist with a retail sales tax.

Transitioning to a national sales tax from the income tax would result in moving from a tax where approximately 76 to 83 percent of the tax base is verified through third-party reporting to a tax where zero percent of the tax base is verified through third-party reporting. The results of current state compliance enforcement efforts clearly indicate that sales tax evasion and theft occur, and the installation of a national sales tax would raise the stakes through greatly increased rates and less ability to verify the tax actually collected.

Assumption Two: The Amount of Sales Taxes Collected by Entities With a Propensity Toward Theft of Sales Tax Is Immaterial

Sales tax theft is primarily a small business problem. Larger companies have internal controls in place that make it more difficult to retain sales tax monies collected from customers and, at a minimum, would require a high degree of collusion among multiple positions within the business to accomplish the theft and conceal it. Larger businesses tend to rely instead on exploitation of the "gray areas" of the law to reduce liabilities, which may in some cases rise to the level of evasion. The effect is the same: the government loses revenue. Losses from larger businesses are more likely to stem from activities related to refusal or failure to collect tax than in theft, and will often represent avoidance rather than evasion.

For example, the State of Florida had lawsuits pending against several online travel companies who were collecting Florida sales tax based on the discounted price they paid to Florida hotels for blocks of rooms rather than on the proceeds from the sale of those rooms to individual customers.[2] While expensive in terms of lost sales tax revenues, this type of activity is arguably avoidance rather than evasion because the positions being taken by the companies are generally transparent and based on an interpretation of the law. In smaller businesses, whether incorporated, operating as partnerships, or sole proprietorships, the owner or owners exercise more control over all aspects of operations, and internal controls are generally lacking. This enables the theft of state funds without collusion, which is important since employees would have less incentive to participate in theft.

It should be noted that IRS tax gap research also identifies small businesses and individuals as responsible for the majority of the income tax gap at the federal level, with underreporting noncompliance accounting for about 83.5 percent of the gross tax gap (Internal Revenue Service, 2012). Some have argued that the sales by these small businesses are immaterial and evasion would result in a lower tax gap than currently exists under the income tax. It can be difficult to develop a proxy for sales by small businesses that would be subject to the national sales tax, but consider the following:

1. Businesses with fewer than 100 employees accounted for $8.8 trillion—or about 28.5 percent—of all receipts during 2007.[3] Using the higher threshold of fewer than 500 employees for classification as a small business (per the Small Business Administration), we find that these businesses accounted for

[2] Alachua County v. Expedia Inc., 1D12-2421, Florida First District Court of Appeal (Tallahassee) finally resolved this issue in favor of the online companies.

[3] Receipts were used for this analysis rather than retail sales since a national sales tax would be applied to receipts and to a much broader range of transactions than retail sales. The actual tax base estimated for a national sales tax is $11.244 trillion gross (Bachman, Haughton, Kotlikoff, Sanchez-Penalver, & Tuerck, 2006).

$12.4 trillion in receipts during 2007, which is about 40.3 percent of the total. Therefore, small business receipts are not a trivial amount (U. S. Census Bureau, 2007).

2. If we assume a national sales tax is designed to be revenue neutral with respect to the taxes it replaces, then based on the above percentages from $712 billion to $1 trillion in sales taxes will be collected by small businesses. Again, not a trivial amount.
3. It is also assumed that the "prebate" will be funded through new revenues raised by the sales tax in order to achieve revenue neutrality (Tuerck, Haughton, Bachman, & Sanchez-Penalver, 2007). The prebate amounts to $500 billion to $660 billion at risk, based on the assumed base reduction of $2.1 trillion. These monies are being transferred by the government, in advance, under the assumption that all sales tax collections will be remitted, but 28.5 percent to 40.3 percent of these collections will be at risk as well (another $142.5 billion to $265.7 billion).[4]
4. The total amount of sales tax and prebates collected by small businesses will therefore be in the range of $854.5 billion and $1.266 trillion.
5. The Internal Revenue Service estimates the net misreporting percentage (NMP) for amounts subject to little or no information reporting to be 56 percent (Internal Revenue Service, 2012).

In my studies I found that it was not uncommon to find individual businesses stealing 90 percent or more of all sales tax collected in a variety of industries, and evasion rates of up to 92 percent with respect to certain types of income industry wide (Christian, 2010). In many cases relatively small businesses selling high-priced items, such as automobiles, or high-volume services, such as security guard services, were able to steal $2 million or more in 2 years or less at sales tax rates of only 6 to 7 percent. Many small businesses would have previously paid little or no income tax, either because they operate at breakeven or less, because they avail themselves of many of the preferences found in the Internal Revenue Code, or because they were among those businesses that fail to report all of their income (contributing to the 56-percent misreporting percentage). They will now be entrusted with up to $1.266 trillion collected as an agent of the government and will recognize that the government has no means of knowing just how much in sales taxes they have collected. Many of these businesses would not steal sales tax collections under any circumstances. Others may not view this as an opportunity to steal, but may be struggling to meet payrolls, pay vendors, and pay rent or mortgages due. When faced with a choice of remitting funds to the government that the government does not know the business has collected or keeping the doors of the business open, many of these businesses will resort to "borrowing" funds for the short term with every intention of repaying these funds at a later date. Once the realization sets in that the government has no idea the funds were "borrowed," the "borrowing" becomes a common occurrence (Christian, 2010).

This type of "borrowing" occurs now at the federal level with respect to the payroll trust-fund taxes even though employers understand that their "borrowing" may be uncovered once employees file their income tax returns and claim their withholding credits. During 2012 the IRS issued 1.6 million delinquency penalties related to employment taxes, 3.9 million penalties for failure to pay employment taxes, and 1.5 million penalties for violations of federal tax-deposit rules with respect to employment taxes (Internal Revenue Service, 2012). Some businesses, such as criminal enterprises, those interested in terrorist financing, or businesses in need of money laundering vehicles will seize on the opportunity to steal sales taxes.

It should be noted that using a cutoff of 100 employees, or even 500 employees, to signify those businesses best able to steal sales tax monies is obviously not a hard line for this determination. Many businesses with far more employees exhibit a control structure conducive to theft and many businesses smaller than 100 or 500 employees may in fact have rather sophisticated internal controls.

It should also be recognized that failure on the part of companies of any size to collect sales taxes would add to the sales tax gap, as well. Moreover, a high-rate sales tax may increase the likelihood of expanding black markets and cash-only transactions, thus reducing sales tax collections.

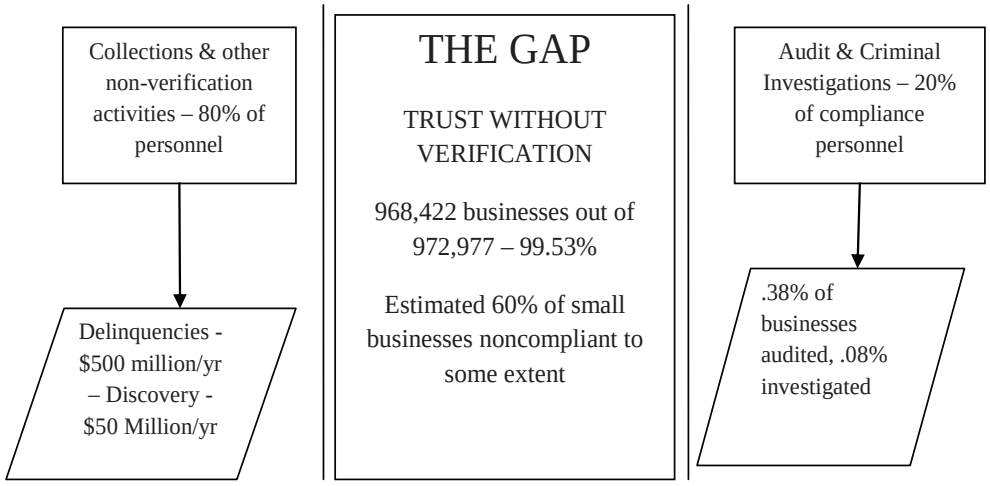
[4] To maintain revenue neutrality, the prebate will be collected on top of the revenue required to replace current tax receipts.

Assumption Three: State Estimates of Sales Tax Compliance Rates Can Be Used To Estimate Sales Tax Lost to Evasion or Theft

Many states boast of voluntary sales tax compliance rates of 95 to 99 percent, but it is important to understand that these quoted rates are generally based on delinquent returns and/or payments only ("delinquencies"). Delinquencies arise when a taxpayer has either failed to file a sales tax return, or has filed a return and failed to remit the sales tax due with the return. If no return is filed, an estimated delinquency amount is computed based on that specific taxpayer's filing history. Nothing in this process addresses those taxpayers who collect taxes and report smaller amounts of tax to the state, retaining the remainder. Further, it is unproductive to look for stolen sales tax among those taxpayers who are delinquent, because those who steal sales tax nearly always file returns on time and pay whatever tax they decide to report. In this manner, they "fly below the radar" and are rarely subject to other enforcement efforts (Christian, 2010).

Sales tax compliance efforts are generally geared toward delinquent taxpayers. For example, in Florida more than 80 percent of compliance enforcement personnel are tasked with collection of delinquencies or other nonverification activities. Collectors do not verify that amounts reported on returns are correct. Auditors and investigators review and verify the activities of only .47 percent (.0047) of taxpayers annually; meaning more than 968,000 businesses (99.53 percent) are not subject to any type of verification whatsoever. It is further estimated that as many as 60 percent of these businesses are noncompliant to some extent.

FIGURE 1. The Sales Tax Enforcement Gap



I evaluated sales tax theft in a number of industries, finding that the amount of sales tax theft far exceeds the amounts of delinquencies reflected on the books of the Florida DOR. In one such exercise, I began with several hundred dealers in a single county and used external data to develop fraud indicators. Using these indicators, I identified 192 dealers for further review. Each of these dealers was contacted and confronted with the evidence indicating they had underreported and underpaid their taxes. Given the possibility that each of these 192 cases had the potential to become a criminal tax investigation, full criminal investigation procedures were followed. The owners of the companies were called to come in for voluntary interviews and asked to bring records that might explain why they had failed to report all of their sales and remit the taxes collected. Most dealers cooperated immediately, appearing either in person or by sending an internal or outside accountant to the meeting. All were informed that the inquiry could become a criminal investigation and were advised they had the right to contact legal counsel before cooperating with the investigation, and several did, in fact, hire attorneys and refused voluntary interviews. For each of these dealers, the violations were for periods extending back three to four years (the statute of limitation for Theft of State Funds is 5 years in Florida). The latest period under investigation was at least 6 months prior to the date of contact with the dealers to prevent investigation

of dealers who were simply delinquent with their payments. Thus, all 192 dealers exhibited lengthy periods of consistent noncompliance. The following table summarizes the results of that review.

TABLE 2. Comparison of Sales Tax Theft to Delinquency Amounts for a Subset of Dealers

Number of dealers examined	192
Revenues reported	$1 billion
Sales tax reported	$36.7 million
Theft of sales tax identified	$21.4 million
Collection balances (delinquencies)	$.3 million
Percent of nonremitted tax identified through delinquencies	1.4%

There were significant revenues ($1 billion) and sales tax reported ($36.7 million) on the returns filed by these entities. However, my investigation showed that these companies failed to remit, on average, 36.8 percent of all the sales tax they collected. The vast majority of these companies filed their returns on time, but some were delinquent with respect to some periods resulting in $302,000 of delinquent taxes on the books that were subject to Department of Revenue collection activity. The delinquent taxes reflected were based on returns filed by the delinquent entities for which no payments had been received or based on estimates made by collectors based on prior filings of the entities. Collectors are not responsible for verifying that amounts reflected as due on those delinquent returns were correct and reflected all of the taxes collected by these taxpayers, or that prior returns that form the basis of current estimates of delinquent taxes were correct. Such verification is not in their job descriptions. The $302,000 delinquent out of the $36.7 million these companies had reported yields a delinquency rate for this group of companies of only .8 percent. The Department would have reported 99.2-percent voluntary compliance with respect to these taxpayers—an obviously incorrect conclusion. The end result is that by focusing only on delinquencies, the department might have eventually realized enforced collections of $302,000 from these companies but would have never been aware of and would have been unable to recover the additional $21.4 million stolen. Further, none of these businesses had been selected for audit for any of the periods under review.

There is no generally acceptable basis for estimating the total sales tax gap, and it cannot be derived from general audit and investigative results because only a tiny percentage of businesses collecting the sales tax are audited or investigated. Moreover, since audits are strategic in nature rather than random, there is no basis to generalize audit results to the entire population. One thing that is clear, however, is that delinquency rates are not a valid proxy for evasion and theft potential.

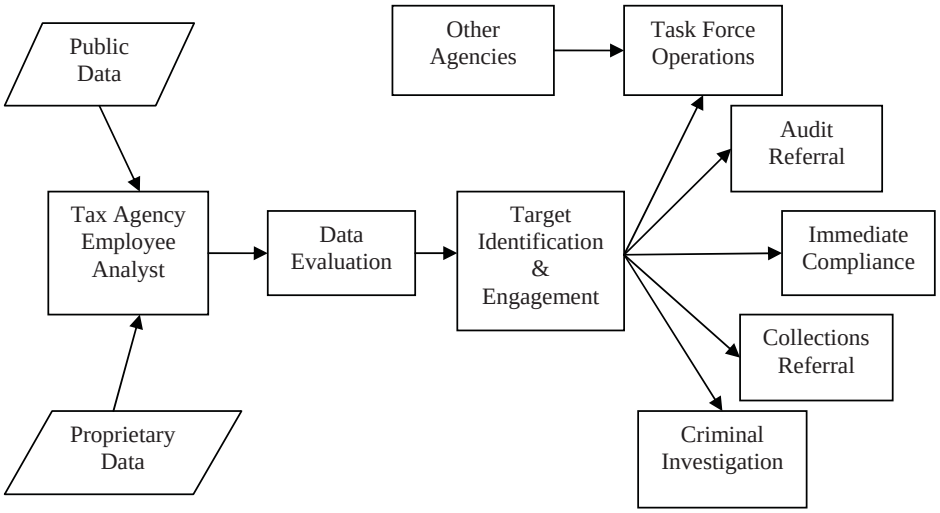
Assumption Four: Under a National Sales Tax There Would Be No Need for the Internal Revenue Service

There are those who advocate a national retail sales tax to replace all other taxes at the federal level. Some of those advocates attempt to "sell" the idea by playing to populist anger and frustration with the income tax system generally and the IRS specifically. They argue that the IRS would no longer be necessary in a sales tax regime. The author believes such arguments are based on faulty assumptions regarding evasion, or the lack of any assumptions regarding evasion in some cases. The assumption is made that the federal government would be just as efficient in collecting sales tax revenue (through the states) as the IRS is in administering and collecting those taxes currently in effect (Tuerck, Bachman, & Sanchez-Penalver, 2007). This assumption ignores the fact that the enforcement environments for these varying taxes have little in common. It also demonstrates a lack of understanding of the sales tax enforcement environment specifically, which is critically important to understand if states are to bear the burden of collecting a high-rate federal consumption tax.

While states do not actively attempt to estimate the true sales tax gap for publication (since they have no empirical basis for estimating what cannot be observed), they do indeed recognize the much higher risk of sales tax theft and most have active data analysis programs in place or under expansion to aid in the develop-

ment of compliance leads for audit and investigative activities. The task is complicated by the requirement to sift through and match tremendous amounts of internal and external data in an effort to frame and understand return data and flag those taxpayers who might be engaged in theft. The required analysis process can be somewhat represented by Figure 2.

FIGURE 2. Simple Lead Development Process



A detailed analysis and evaluation of both internal and external data are required to identify sales tax theft. While third-party reporting of credit-card sales and, in some cases, of wholesale alcohol and tobacco sales can be helpful, such reporting is not comprehensive enough to stand alone as compliance-enforcement tools. For example, many have suggested to me that comparing gross sales as reported on sales tax returns over the course of a year to gross sales reported on the business' income tax return should be a good way to estimate the volume of sales taxes due. At the current time, it is not possible for states to prepare an estimate in this manner. The fact pattern may vary from state to state depending on differing state reporting requirements, but I will continue to use Florida as an example.

In Florida, businesses that elect Subchapter S status and those organized as Limited Liability Companies do not have to file a state corporate return. There is no individual income tax in Florida, so sole proprietorships do not file an income tax return. As a result, gross sales data do not exist for comparison at the state level for most small businesses. The state does acquire federal income tax return data to match to business sales tax reporting to make sure that federal sales reported matches state sales tax reporting.[5] However, companies that steal sales tax will generally make sure the sales reported on sales tax returns are consistent with sales reported on federal income tax returns. In other words, these businesses are understating gross revenues on their federal returns as well, and contributing to the 56-percent net misreporting percentage referred to in the federal tax gap estimates.

A larger problem is determining the amount of taxable sales. In most businesses gross sales do not equal taxable sales since about 60 percent of all sales are exempt from sales tax under one provision or another. Many companies who steal do so by claiming fraudulent exempt sales. There is no way to determine what portion of sales is exempt when looking at gross sales on a federal income tax return. But perhaps more importantly from the standpoint of the argument set forth in this paper, these points become moot when we consider a national sales tax, because there will no longer be an income tax return at the federal level to use for measuring whether or not an entity reported all their sales.

[5] Matching gross sales reported on sales tax returns to sales reported on federal income tax returns is not definitive with respect to underreporting sales because sales taxes in Florida must be reported on the accrual basis while sales on the income tax return could be reported on cash basis, and frequently are by small businesses.

To better understand the dimensions of the analysis problem, consider the following:

- States currently focus most enforcement efforts on managing delinquencies rather than identifying theft. The result is an allocation of maximum resources to those efforts that produce the lowest yield in terms of revenue recovered. Optimal levels of enforcement resources and the allocation of those resources to enforcement strategies should be based on an analysis of marginal revenue compared to marginal costs. It is difficult in the current political environment to convince legislatures to adopt a marginal-cost approach to tax enforcement because their primary focus is generally on reducing taxes and regulation, and cutting costs rather than optimization.
- Budget constraints dictate that agencies must seek to improve compliance results in an ever more cost-effective manner by employing innovative compliance-enforcement solutions that do not require large increases in the overall level of resource dedication. Such solutions may involve reallocating and refocusing certain existing resources and greater investments in technology.
- Sales tax enforcement is more difficult than income tax enforcement because of the lack of third-party verification that is so integral to the income tax system. Sales tax enforcement requires more manpower and the use of more complex technologies because internal systems should ideally be interfaced with a wide variety of third-party systems to approach the level of verification that exists in an income tax system. If programmed interfaces are not possible, then more manual analysis is required and the cost of enforcement increases.
- To ensure adequate enforcement of a national sales tax, IRS human and technological capabilities would need to be expanded dramatically. Currently, IRS enforcement personnel are tasked with controlling evasion related to only about 17 to 26 percent of the total tax base. Under a national sales tax the tax base subject to evasion would arguably expand to 100 percent of the base when all avoidance and evasion possibilities are considered. Auditing at current rates would be insufficient and all systems would need to be retooled for the collection and analysis of greatly different types of data than what are currently used for enforcement purposes.
- IRS systems would need to be integrated with systems from other federal agencies and from state and local Governments to approach maximum effectiveness, and this would prove troublesome given the confidentiality requirements attached to tax information.
- The increase in the sales tax rate to a combined federal and state rate that exceeds 30 percent would increase the benefits related to evasion or theft to the point where theft could become both more commonplace generally and, because of the higher yield, a major source of funding for criminal and terrorist activities.

Moving all tax compliance responsibilities down to the state level would not work, either, without large increases in the enforcement budgets at the state level. States are only just beginning to upgrade their enforcement capabilities to a more sophisticated structure designed to deal with sales tax theft rather than primarily delinquencies, so many of the costs of retooling the IRS to enforce a national sales tax would be required at the state level as well. These costs cannot be avoided merely by moving the enforcement responsibility to the states. Moreover, states would have to be able to enforce the law consistently from state to state, a function better served by a federal-level entity. To expand state capabilities to enforce a sales tax with rates that are five to six times larger than the rates they currently administer, making theft so much more profitable, is no mean task. The enforcement issues at the state level are similar to the issues faced by the IRS with respect to income items where no third-party reporting exists, except that the states are not currently addressing in adequate fashion any noncompliance beyond delinquencies and the results of the strategic audits of a tiny proportion of the businesses collecting sales tax.

Studies have shown that there is also a significant indirect effect of targeted enforcement activities (Christian, 2010), particularly within an industry that is targeted for enforcement action. To achieve this indirect effect, which will result in lower enforcement costs overall, an adequately funded, well-conceived, and active enforcement capability is required. Talk of no longer requiring the enforcement capabilities of the IRS may be popular rhetorical red meat for the masses, but such a course of action would have disastrous consequences for effective and efficient revenue collection.

Conclusion

The tasks performed by collections personnel, auditors, and criminal investigators in modern revenue agencies are critical to the mission of the agencies, but there remains a gap in enforcement coverage that allows unacceptably large amounts of sales tax theft to escape detection. Most states are quietly aware of this issue and many have embarked on sophisticated lead development programs to address this shortcoming. The implementation of a national sales tax system without accounting for the loss of third-party verification and without retasking the income tax enforcement structure to address the substantially different needs of a sales tax enforcement regime would result in a substantial increase in both the ability of taxpayers to engage in noncompliant behavior and the amount of tax revenue lost to fraud and abuse. Additionally, a high-rate sales tax with inadequate or no administrative enforcement capabilities is a perfect tool for criminal enterprise: fraud, money laundering, and terrorist financing activities.

What is surprising is that with 45 sales tax systems in existence in this country, few attempts to study sales tax theft and evasion have been made at the level required to justify some of the critical assumptions inherent in what would be a massive tax policy shift at the federal level. It is true that the ability to study sales tax enforcement in the field can be extremely limited. Access to data have been a perennial problem for tax evasion researchers because of confidentiality issues, and the same holds true with respect to the retail sales tax. It is difficult to acquire data that has not been summarized to the point of being useless, as was noted by Alm (1991), but limitations should not preclude cooperative studies within state revenue agencies with appropriate privacy protections in place. Much more research would be required to determine what an appropriate compliance enforcement regime would look like with respect to a national sales tax.

References

Alm, J. (1991). A perspective on the experimental analysis of taxpayer reporting. *The Accounting Review*, 577–593.

Antonides, G., & Robben, H. (1995). True positives and false alarms in the detection of tax evasion. *Journal of Economic Psychology*, 617–640.

Bachman, P., Haughton, J., Kotlikoff, L., Sanchez-Penalver, A., & Tuerck, D. G. (2006). *Taxing sales under the fair tax: What rate works?* Retrieved June 12, 2013, from www.fairtax.org: http://www.fairtax.org/PDF/Tax%20Notes%20article%20on%20FT%20rate.pdf.

Blanthorne, C. M. (2000). *The role of opportunity and beliefs on tax evasion: A structural equation analysis.* Arizona, United States: Arizona State University.

Bohren, O. (1998). The agent's ethics in the principal-agent model. *Journal of Business Ethics*, 745–755.

Carnes, G. A., & Englebrecht, T. D. (1995). An investigation of the effect of detection risk perceptions, penalty sanctions and income visibility on tax compliance. *Journal of the American Taxation Association*, 26–41.

Chang, O., & Schultz, J. (1990). The income tax withholding phenomenon: Evidence from TCMP data. *Journal of the American Taxation Association*, 88–93.

Christian, P. C. (2010). *Sales tax enforcement: An empirical analysis of compliance enforcement methodologies and pathologies.* Retrieved from FIU Electronic Theses and Dissertations. Paper 335: http://digitalcommons.fiu.edu/etd/335.

Droege, S., & Spiller, S. (2009). Critique of a premise: Illuminating the cracks in the agency theory framework. *Journal of Legal, Ethical and Regulatory Issues*, 41–54.

Garner, C. A. (2005). Consumption taxes: Macroeconomic effects and policy issues. *Economic Review—Federal Reserve Bank of Kansas City, 90(2)*, 5–29.

Internal Revenue Service. (2007). *SOI Tax Stats—Individual Income Tax Returns.* Retrieved December 5, 2009, from Internal Revenue Service: http://www.irs.gov/taxstats/indtaxstats/article/0,,id=133414,00.html.

Internal Revenue Service. (2012). *SOI Tax Stats—IRS Data Book.* Retrieved May 23, 2013, from Internal Revenue Service: http://www.irs.gov/uac/SOI-Tax-Stats-IRS-Data-Book.

Internal Revenue Service. (2012, January 6). *Tax Gap for the Year 2006.* Retrieved March 23, 2013, from Internal Revenue Service: http://www.irs.gov/pub/irs-soi/06rastg12overvw.pdf.

Martinez-Vazquez, J., Harwood, G., & Larkins, E. (1992). Withholding position and income tax compliance: Some experimental evidence. *Public Finance Quarterly*, 152-206.

Miller, G., & Whitford, A. (2006). The principal's moral hazard: Constraints on the use of incentives in hierarchy. *Journal of Public Administration Research and Theory*, 213–233.

Ramamoorti, S. (2007). Fraud, the human factor. *Financial Executive*, 53–55.

Tuerck, D. G., Bachman, P., & Sanchez-Penalver, A. (2007, September). *Tax administration and collection costs: The fair tax vs. the existing federal tax system.* Retrieved June 12, 2013, from www.fairtax.org: http://www.fairtax.org/PDF/TaxAdminCollectionCosts071025.pdf.

Tuerck, D. G., Haughton, J., Bachman, P., & Sanchez-Penalver, A. (2007, February). *A comparison of the FairTax base and rate with other national tax reform proposals.* Retrieved 06 12, 2013, from www.fairtax.org: http://www.fairtax.org/site/DocServer/A_Comparison_of_the_FairTax_Base_and_Rate.pdf?docID=761.

U. S. Census Bureau. (2007, December). *2007 Statistics of U.S. Businesses Historical Data—Business & Industry.* Retrieved April 28, 2013, from http://www.census.gov: http://www.census.gov/econ/susb/data/susb2007.html.

Watrin, C., & Ullmann, R. (2008). Comparing direct and indirect taxation: The influence of framing on tax compliance. *The European Journal of Comparative Economics*, 33–56.

Factors Influencing Use Tax Payment in Illinois

Joanna Koh, Illinois Department of Revenue; David Merriman, University of Illinois Chicago; and Hector M. Vielma, Illinois Department of Revenue[1]

I. Introduction

Like most U.S. states, Illinois' primary sources of state tax revenue are the personal income tax and the general sales tax. Illinois' general sales tax of 6.25 percent on most items is actually an "occupation" tax imposed on sellers. The legal occupation-tax liability of sellers is based on the amount of tangible personal property purchased from them for use in Illinois.[2] Sellers reimburse themselves for this liability by charging buyers a sales tax.

Well-known constitutional restrictions prevent Illinois from requiring out-of-state sellers with no legal nexus (roughly physical presence) in Illinois to remit occupation taxes to the state even when it is known that buyers will use the purchased goods in Illinois. Any person or business that uses goods in Illinois that were purchased outside of Illinois at a lower tax rate or tax free is liable for the difference as a "use" tax. The use tax applies to items purchased through the mail, by phone, online from other states, or in-person via a cross-border sale for use in Illinois. The use tax is intended to create a level playing field between out-of-state sellers who do not collect sales tax on purchases by Illinois residents and Illinois brick-and-mortar retailers who are required to collect the Sales tax when a transaction occurs.[3]

Illinois' use tax has been in place since 1955. Purchases made over the Internet, through toll-free numbers, from mail-order catalogs and from out-of-state locations are examples of purchases subject to use tax. Illinois use tax rates are 6.25 percent of the purchase price of general merchandise and 1 percent of the purchase price of qualifying food, drugs, and medical appliances.[4] If the use tax amount is $600 or less, the tax is due by April 15 of the following year. If the use tax due is more than $600, it must be paid by the last day of the month following the month when the purchase was made. Illinois collected over $1.2 billion of state use tax in Calendar Year 2012. This was about 13 percent of the total revenue from the statewide sales and use taxes.

Retailers that have a physical presence in Illinois and make retail sales from out-of-state locations to customers in Illinois remit about half of the total use tax payments, which they collect from their customers in Illinois. The remaining share of state use tax that is collected consists mostly of either payments from individuals and businesses related to motor-vehicle purchases or payments from businesses that make a taxable purchase from a supplier who does not collect Illinois tax.[5]

[1] The authors thank Natalie Davila for helpful discussions at an early stage of this work, Andy Chupick for helpful comments on parts of this manuscript, participants in the 2013 IRS-Tax Policy Center Research Conference for helpful comments, and Brian A. Hamer for making the authors' collaboration possible. Views expressed in this paper are those of the authors and do not necessarily reflect the positions of their institutions.

[2] Tangible personal property excludes real estate, stocks, bonds, and other "paper" assets representing an interest in some asset.

[3] Remote sales have been federally regulated by the 1998 Internet Tax Freedom Act and two Supreme Court rulings. Under current law, online and catalog sellers (viz., remote sellers) are required to collect sales tax only in states where they have nexus. Otherwise, consumers who shop online and do not pay a sales tax at the time of purchase are required to pay the tax to their resident state. On May 6, 2013, the U.S. Senate passed the Marketplace Fairness Act bill, and an identical bill is referred to the House Subcommittee on Regulatory Reform, Commercial and Antitrust Law. This legislation would allow states that currently charge sales taxes to require large remote retailers to collect sales and use taxes on purchases made by their residents regardless of their physical presence in that state. Were this legislation to take effect, we expect that use tax compliance by remote sellers would increase significantly.

[4] Illinois Department of Revenue, Use tax Questions and Answers; available at http://tax.illinois.gov/Individuals/FAQs-Use-tax.htm.

[5] Use tax is reported on Form ST-1 for registered retailers and on Form ST-44 for nonretailers. For motor-vehicle purchases, RUT-25, RUT-50, and RUT-75 are used.

Illinois can easily monitor use tax payments due on automobiles by tracking automobile registrations. Similarly, business purchases that incur a use tax obligation are relatively easy to monitor compared with purchases by individuals because the business purchases tend to be large, recurring, and may be documented in public records. In this paper we are concerned with the use tax obligations of individuals whose purchase activities may not be easily monitored.[6]

In 2010 and 2011, Illinois offered a use tax amnesty, which allowed individual taxpayers to pay use tax without penalty and interest for purchases during the period from July 1, 2004, through December 31, 2010.[7] Taxpayers were encouraged to file and pay use tax from January 1, 2011, through October 15, 2011. This initiative did not receive much attention and the Illinois Department of Revenue (IDOR) attributes less than 4 million dollars of additional collections in 2011 to it.

In an effort to increase awareness of and compliance with the use tax, the Illinois General Assembly followed the lead of more than 20 other states and passed a law putting a use tax line on the personal income-tax form (IL-1040) beginning in Tax Year 2010. This change gives individuals a simpler alternative compared to the previously available Form ST-44 to pay their Illinois use tax liability.[8] Since 2010, the instructions for the IL-1040 have included a "lookup" table that guides tax filers to estimate their use tax liability based solely on their federal adjusted gross income (FAGI) in the event the filer does not have records of out-of-state purchases.

A small minority of Illinois tax filers make any use tax payments under either the ST-44 or the IL-1040. While it is difficult to attribute use tax liability to any particular tax filer, aggregate studies make clear that many tax filers have a use tax liability with which they are not complying. Prior to 2010, it was plausible to think of the noncompliance as inadvertent and resulting from ignorance of the legal mandate since most Illinois tax filers were probably unaware of the use tax and of procedures required to file Form ST-44. Since 2010, it is probably more plausible to think of widespread use tax evasion by Illinois tax filers who justifiably believe that there is a very small possibility that failure to comply with use tax requirements will result in any penalty.

Because we have no information about the use tax liability of individual taxpayers, we cannot directly measure tax compliance or avoidance. However, various studies discussed in Section 2 have estimated aggregate Illinois use tax liability. By comparing these estimates with our data on use tax payments, we infer that there is widespread noncompliance.

To better understand use tax payments in Illinois, we obtained IDOR's "warehouse" data on ST-44 tax returns covering the years 2005 through 2012. Also, we obtained access to the universe of almost 6 million of Illinois' individual income tax returns in Calendar Years 2010 and 2011 for this research. Illinois has a very simple tax form (see http://www.revenue.state.il.us/taxforms/IncmCurrentYear/Individual/IL-1040.pdf) that requires little information beyond the filer's FAGI, number of deductions, withholding amount, and tax-exempt retirement income. We use these data to relate tax-filer characteristics to the probability that they will make Illinois use tax payments on the IL-1040 in 2011.

II. Previous Literature

The general topic of compliance with tax laws has been a major area of study for tax scholars—particularly economists. Slemrod (2007) and Slemrod and Bakija (2008) discuss recent evidence on this topic. Much of the research in this area has focused on federal taxes, particularly the federal income tax. Compliance with state and local taxes, with the exception of tobacco taxes, has gotten less attention.[9] Recently, there have been a number of serious attempts to measure and understand sales and use tax compliance, particularly as it relates to Internet commerce.

---

[6] Some of these obligations could theoretically be disposed of by filing Form ST-44 but very few individuals file this form. Illinois Department of Revenue, Illinois Use tax; available at http://tax.illinois.gov/individuals/illinois-use tax.htm.

[7] Illinois Department of Revenue, Use Tax Amnesty Questions and Answers; available at http://www.revenue.state.il.us/Amnesty/Amnesty-FAQs-Use-tax.htm.

[8] Tax filers who owe a use tax of more than $600 per person are required to use Form ST-44.

[9] See Chernick and Merriman (2013) and Merriman (2010) and sources cited therein for a small sample of the literature on tobacco-tax compliance.

Einav, et al. (2012) investigate the sensitivity of Internet retail purchases to sales taxes using data from the eBay marketplace.[10] They use several approaches that all indicate online purchases are sensitive to sales-tax rates. In their most novel investigation, Einav, et al. note that because seller locations are revealed only after buyers have expressed interest in an eBay item by clicking on its listing, the required sales-tax payment can be treated as an exogenous "surprise" that affects the price of the item. The authors exploit this insight by comparing buyers who arrive at the same item page, some of whom are located in the same state as the seller (and therefore would be compelled to pay sales tax if they purchase the item) and some of whom are located in different states and therefore would accrue only a (potential) use tax liability that they might, or might not, later discharge by making a payment to their state of residence. Einav, et al. (2012) find a tax-price elasticity of about -2 for buyers who click on a listing; each 1 percent increase in the sales-tax rate causes the probability of a sale to fall by about 2 percent.

Alm and Melnik (2012) also used eBay data to study use taxes. They collected information on 21,000 eBay listings generated by roughly 7,000 sellers and over 9,000 buyers on a typical day to measure the extent of cross-border shopping and estimate its potential for generating state use tax revenue. They find that 94 percent of eBay purchases are made by out-of-state buyers. Alm and Melnik present state-specific estimates of use tax liability attributable to eBay sales under a number of assumptions about how their "typical day" data on a subset of sales categories can be generalized to all eBay sales during 2010. In Table 8 (p.28), Alm and Melnik estimate that Illinois use tax liabilities as a result of eBay purchases in 2010 were nearly $41 million. They conservatively estimate that 2007 eBay U.S. e-commerce sales of about $23 billion represented roughly 17 percent of the total U.S. e-commerce sales of $137 billion. If Illinois were typical of the nation and if eBay sales were representative of all e-commerce sales, this would suggest a 2007 Illinois use tax liability of about $246 million (= 137/23 x 41 million).[11]

Using a different methodology, Chupick and Davila (2009) estimate unpaid use tax from online sales to Illinois households and businesses for Calendar Years 2005 through 2010. They start with actual[12] U.S. online sales, estimate Illinois' share, subtract out sales that either are not taxable or where enforcement of the use tax is strong, and estimate the tax due. Chupick and Davila estimate use tax liabilities ranging from $103 million in 2005 to $169 million in 2010. Their 2009 estimate of $169 million appears to suggest less potential additional revenue than Alm and Melnik.

While there have been a number of attempts to measure theoretical use tax liabilities, only recently have there been significant attempts to study and encourage compliance. One major compliance initiative has been the effort by a large number of states to incorporate payment of use tax liabilities into the personal income-tax filing process. Manzi (updated 2012) provides a very useful description of these efforts.

Manzi reports that 25 states have a line on the income-tax return that allows payment of use taxes. An additional seven states provide information about how to discharge use tax liabilities in the booklet that explains procedures for complying with the state income tax. Manzi also provides data showing that, in states that allow payment of use tax on the personal income-tax form, the percent of returns with a nonzero use tax payment varies from 0.3 in California to 9.8 in Maine. Some states employ a "lookup table" that provides an estimate of use tax liability as a function of taxpayer income. Manzi reports that the participation rate—i.e., rate of tax returns with a nonzero use tax payment is 3.1 percent on average in states with a lookup table but only 0.6 percent in states without a lookup table. She notes that Indiana and Rhode Island have tried to increase collections by sending information about the use tax to a random sample of taxpayers. Kentucky, Massachusetts, and Maine have also introduced individual compliance programs.

Gunter (2011) has done an in-depth analysis of factors influencing use tax payments on Maine's personal income-tax returns over the years 2003 to 2009. According to Manzi, Maine has the highest rate of income-

[10] Einav et al. (2012) also provide a review of related literature beyond that discussed here.

[11] Alm and Melnik note that their calculations assume that "all of the observed transactions are subject to use tax, when in fact some transactions are not legally taxable. As a result, these calculations are an upper bound on potential use tax under-collection." (p.27) On the other hand, Alm and Melnik's work covers only online transactions. Catalog purchases and purchases through physical cross-borders are not studied. As a result, in our view, the estimate that we derive using Alm and Melnik's figures should not be viewed as an upper bound for the potential use tax liability in 2007.

[12] Projected sales were used in 2008, 2009, and 2010 since data were not yet available for those years at the time they did their research.

tax returns with use tax payments of any state in the country. Gunter is the first paper in the literature to our knowledge to use information on use tax payments from income-tax returns to study factors that influence compliance. He reports that between 11 and 13 percent of income-tax returns in Maine have use tax payments during his sample period. His paper provides an important model, inspiration, and basis of comparison for the analyses of Illinois' personal income-tax returns that we discuss in sections 3 and 4.

Gunter uses a balanced panel of Maine taxpayers and estimates a linear probability model with use tax payment (=1 if some use tax payment and zero otherwise) as the dependent variable. He finds that taxpayers who used a paid preparer are roughly 8 percent less like to pay use tax than similar taxpayers who did not use a paid preparer. In order to quell concern that taxpayers who use paid preparers are self-selected individuals with an especially strong desire to minimize their tax burden, Gunter runs a specification that controls for the use of a tax preparer in the previous year. He finds that a taxpayer who switches to using a preparer is less likely to pay use tax compared with a taxpayer who does not use a preparer in the current or previous year.

Conventional wisdom holds that small business owners are especially likely to evade and avoid taxation. Gunter tests and rejects this hypothesis. He finds that small business owners are slightly *more* likely to pay use taxes on their income taxes in Maine.

Gunter also finds that taxpayers who make charitable contributions are more likely to pay use taxes. In particular, the 1 percent of taxpayers who take an option to check a box and make a voluntary charitable or political contribution are 23 percent more likely to pay use taxes than taxpayers who do not.

Gunter notes that previous research on tax avoidance has found that noncompliance increases with the amount of tax due. He finds that the probability that a Maine income-tax return will have a positive use tax payment increases with the refund due.

Recently Anderson (2013) reported the results of a field experiment conducted in cooperation with the Nebraska Department of Revenue in which a thousand randomly choosen taxpayers received a postcard encouraging them to report their use tax liability when they filed their state income taxes. Anderson finds that this nudge more than doubled the likelihood of use tax reporting and increased the amount of revenue collected.

III. Descriptive Statistics About Illinois Use Tax Payments

Table 1 shows the number of ST-44 and IL-1040 returns, the number of IL-1040 returns with use tax payments in 2010 and 2011 and the use tax paid on each type of return. Both businesses and households may incur use tax liability in Illinois and both may discharge their liability using the ST-44.[13] Of course, households can also pay their use tax using the IL-1040. In many cases, nonincorporated businesses can also pay their use tax using the IL-1040. With our data, there is no unambiguous method to determine whether a particular use tax payment was made for a household or a business activity. In most years, roughly one-third of ST-44 returns list a Federal Employer Identification Number (FEIN) and two-thirds list a Social Security (SS) number. ST-44 use tax payments associated with a FEIN almost certainly resulted from business activity. ST-44 use tax payments associated with an SS number may come from either a household or a business activity. Roughly half of ST-44 use tax payments are associated with returns using an FEIN, while the other half of payments are coming from returns listing an SS number.[14]

[13] Registered retailers discharge their use tax liability using form ST-1.

[14] Our discussion of ST-44 filers and payments is based on our aggregation of data on individual filings. We also obtained data about ST-44 filers compiled within IDOR from accounting reports. While the accounting report data should, in principle, be based on the disaggregated data, our totals do not match accounting totals. In general, we find slightly less revenue and slightly fewer returns than accounting reports. We suspect that the discrepancies in these figures are due to the fact that we attribute an ST-44 return to the year in which it was filed while the accounting reports may attribute a return to the year in which the liability occurred. We have not been able to verify this explanation.

TABLE 1. Illinois Use Tax Payments Before and After a Use Tax Payment Option Was Added to Personal Income Tax Return Form

Calendar year	Number of ST-44 returns	Total use tax on ST-44 returns (million dollars)	Number of IL1040 returns	Number of IL1040 returns with non-zero use tax payment	Total use tax payment on IL1040 returns (million dollars)
2005	1,857	2.58	na	np	np
2006	4,520	4.00	na	np	np
2007	6,366	5.26	na	np	np
2008	9,801	4.86	na	np	np
2009	8,055	5.34	na	np	np
2010	6,415	6.09	4,747,133	242,412	10.22
2011	27,618	8.12	5,124,947	239,900	10.92
2012	4,256	5.94	na	na	na

NOTE: na=not available, np=not possible. Analysis of IL-1040 is restricted to 5,124,947 matched returns in 2010 and 2011. Dependents, returns with over $1mm FAGI, returns with zero or negative FAGI, and returns of nonresidents are dropped from the analysis. Prior to 2010 it was not possible to make a use tax payment on the IL1040 form. The large number of ST-44 returns in 2011 is the result of a use tax amnesty in that year. Details are discussed in http://tax.illinois.gov/Amnesty/Amnesty-FAQs-Use-Tax.htm.

As shown in Table 1, Illinois' nearly 13 million residents filed fewer than 10,000 ST-44 returns each year except in 2011, when there was a tax amnesty that resulted in a temporary surge in returns. The introduction of an alternative use tax filing mechanism—a line on the IL-1040—after 2010 appears to have had little impact on ST-44 tax payments. In 2011, there were more than eight times as many IL-1040 returns with a positive use tax payment than ST-44 returns. While the average IL-1040 use tax payment was small, total revenue collected on this form exceeded that collected using the ST-44 by a large margin.

Despite the addition of a use tax payment option on the IL-1040, total Illinois use tax payments in 2010 of $16.3 million are less than half of the $41 million legal liability Alm and Melnik (2012) estimate is due on eBay purchases alone and less than one-fourth the amount that Chupick and Davila (2009) estimate is due on all online transactions. Neither Alm and Melnik nor Chupick and Davila provide an estimate of the use tax liability due on offline cross-border transactions, so these estimates should be viewed as quite conservative estimates of the use tax that is legally due to Illinois.

We obtained an alternative estimate of legal Illinois resident use tax liabilities by calculating the liability that would be incurred if all tax filers used the Illinois' use tax lookup table which is shown in Figure 1.

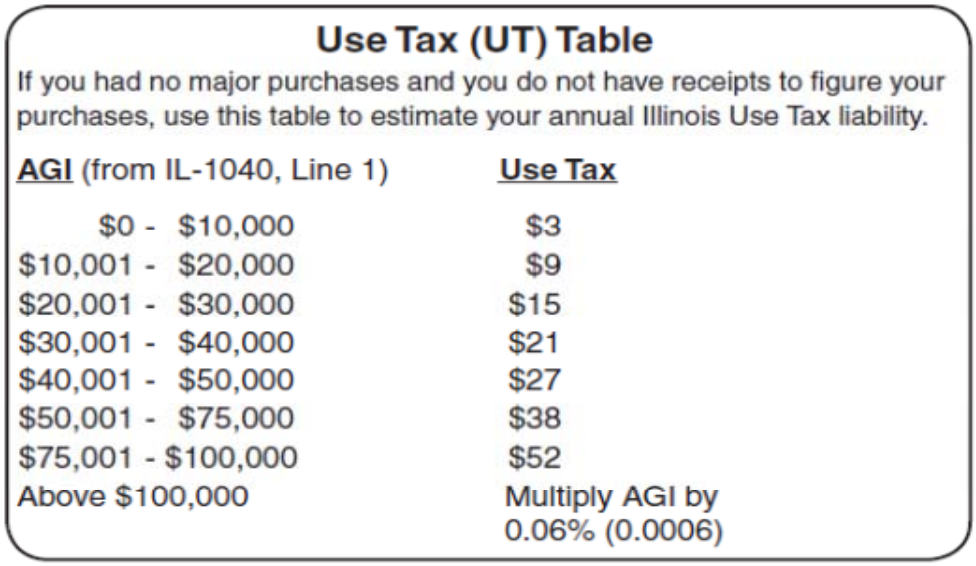
Illinois' use tax lookup table assesses a use tax of 0.06 percent of a filer's FAGI in the center of each lookup category[15] and the 2010, 2011, and 2012 use tax lookup tables charge identical rates. If these rates are applied to 2011 IL-1040 returns, tax filers would hypothetically generate use tax payments of $214 million from Illinois residents who were not claimed as dependents on someone else's tax return. This is more than 20 times as much as the amount of use tax actually paid on the 2011 IL-1040.

While we cannot determine the precise amount of use tax noncompliance, various estimates suggest that it is large. However, a significant minority (about 4 percent) of IL-1040 tax filers pay use taxes despite the fact that, like other U.S. states, Illinois makes almost no effort to audit personal income tax filers' use tax payments.

We create a "panel" of tax returns by attempting to match the primary SS number on all 2010 returns with a 2011 return with the same primary SS number. We create Table 2 using this full panel of returns. Persistence is one of the most dependable regularities in the study of human behavior. Table 2 shows that use tax payment behavior is, perhaps unsurprisingly, quite persistent among Illinois taxpayers.

[15] The rate used in Illinois use tax lookup table was apparently justified on the basis of unpublished research by IDOR employees who estimated that uncollected use taxes on nonauto e-commerce sales in 2008 were about 0.06 percent of Illinois' FAGI. Apparently, IDOR did not estimate use tax liabilities stemming from other (non-e-commerce) transactions, nor did IDOR estimate whether average use tax liability varied with a tax filer's FAGI.

FIGURE 1. Illinois Use Tax Liability Schedule



Use Tax (UT) Table

If you had no major purchases and you do not have receipts to figure your purchases, use this table to estimate your annual Illinois Use Tax liability.

AGI (from IL-1040, Line 1)	Use Tax
$0 - $10,000	$3
$10,001 - $20,000	$9
$20,001 - $30,000	$15
$30,001 - $40,000	$21
$40,001 - $50,000	$27
$50,001 - $75,000	$38
$75,001 - $100,000	$52
Above $100,000	Multiply AGI by 0.06% (0.0006)

(Source: http://www.revenue.state.il.us/taxforms/IncmCurrentYear/Individual/IL-1040-Instr.pdf.)

As shown in Table 2, about 4.9 million filers submitted tax returns in both 2010 and 2011 and had no use tax payment in either year. Also, about 685 thousand additional returns with no use tax payment were filed in 2010 but could not be matched to a 2011 return, and about 658 thousand returns with no use tax payment in 2011 could not be matched to a 2010 return.[16] About 4.6 percent of all 2010 IL-1040 returns and about 4.4 percent of all 2011 IL-1040 returns included a use tax payment.[17]

Payment behavior is quite persistent. Roughly half (82,489) of the 165,198 filers who made a use tax payment of $1 to $50 in 2010 also made a use tax payment of $1 to $50 in 2011. Most of the other half (72,200) made no use tax payment in 2011, but a small minority of filers (around 10,000) made a larger use tax payment in 2011. The same general pattern holds for other use tax payment categories so that the numbers in the main diagonal of Table 2 (where the use tax payment category is the same in 2010 and 2011) are larger than off-diagonal elements. Put another way, a filer who paid use tax in 2010 had a more than 50-percent probability of paying use tax in 2011, while all IL-1040 tax filers had only a 4.4-percent probability of paying use tax in 2011.[18]

Illinois collected $11.5 million of use tax revenue in 2010 and $12.1 million in 2011 from approximately 260,000 to 270,000 tax filers, which represents a 4.5-percent payment rate. This compares with an average 2009 payment rate of 3.1 percent among states that have both a use tax reporting line on their income-tax return and a lookup table.[19] While Illinois' IL-1040 payment rate is above the national average, it is only about half of the payment rate in Vermont and Maine.

[16] We matched 2010 and 2011 returns based on the "primary" SS number on the 2010 return. If the holder of this SS number did not file a 2011 IL-1040 or if she or he filed a 2011 IL-1040 but was not listed as the "primary" SS number, our computer algorithm will fail to find a match.

[17] We do not know what share of tax filers actually incurred a use tax liability. However, Madden and Rainie (2003) report that in December of 2002, on any given day, 5 percent of Internet users (3 to 6 million individuals) made an online purchase and 61 percent of Internet users (about 67 million people) made an online purchase at some point. By 2010, these percentages surely must have been higher.

[18] Some readers might erroneously expect that persistence of payment among those who paid in a previous year could eventually lead to widespread payment of the use tax. The Appendix demonstrates that this expectation is misplaced.

[19] Manzi, Nina, updated 2012.

TABLE 2. Cross-Tabulation of 2010 and 2011 Use Tax Payments by Illinois Tax Filers

		Amount of Use Tax Payment on 2010 Income Tax Return							
	Category	**No matching 2010 return**	**No use tax payment**	**$1 to $50**	**$51 to $100**	**$101 to $300**	**$301 to $600**	**Over $600**	**TOTAL 2010 returns**
	No matching 2011 return		685,031	13,757	1,355	464	56	14	700,677
Amount of Use Tax Payment on 2011 Income Tax Return	No use tax payment	657,910	4,896,250	85,483	19,107	5,472	687	151	5,007,150
	$1 to $50	11,559	72,200	82,489	8,525	1,824	152	8	165,198
	$51 to $100	1,014	20,471	10,523	24,753	2,538	164	22	58,471
	$101 to $300	349	6,531	2,114	3,359	8,179	477	28	20,688
	$301 to $600	33	804	161	165	672	725	33	2,560
	over $600	12	219	35	28	76	111	79	548
	TOTAL 2010 returns								**5,955,292**
	TOTAL 2011 returns	**670,877**	**4,996,475**	**180,805**	**55,937**	**18,761**	**2,316**	**321**	**5,925,492**

Notes: Each cell shows the number of Illinois tax returns with characteristics in column and row headings. The greyed cells inside the black box represent returns of taxpayers present in both 2010 and 2011. The first row of numbers and the left-most column of numbers show the number of tax returns that were present in only one of the years. For example there were 685,031 returns with no use tax payment in 2010 and no matching return in 2011. Similarly there were 657,910 tax returns with no use tax payment in 2011 and no matching return in 2010.

IV. Cross-Tabular Evidence About Taxpayer Characteristics That Influence the Probability of Illinois Use Tax Payment

In this section, we examine potential factors that may influence tax filers' propensity to pay use taxes when sellers without Illinois nexus cannot be required to remit the tax. We begin by examining some simple cross-tabulations and similar statistics to examine individual factors that influence the probability of payment. In the next section, we report the results of regression analyses that control for a variety of factors simultaneously.

We study data about Illinois use tax payments to learn about use tax policy specifically and also to gain broader lessons about factors that may affect tax compliance more generally. Tax filers may be less prone to fulfill their obligation to pay use tax than to pay other taxes because: (a) it is difficult for tax authorities to audit use tax liability and therefore to compel compliance; (b) there could be a high compliance burden since use tax liability may result from multiple small transactions (in the absence of the use tax, a typical tax filer would have little reason to track transactions that create a use tax liability); and (c) the use tax is generally small and obscure and, at least until 2010, many Illinois tax filers may not have been aware of their potential use tax liability at the time they filed their personal income-tax form.

Tax authorities could seek to increase use tax compliance by supporting federal legislation or an amendment of the U.S. Constitution that would place responsibility for remitting the use tax on out-of-state sellers.[20] Doing this would make it much easier for tax authorities to audit and compel use tax compliance.

The option to pay use taxes as part of the filing of a personal income-tax return combined with the use tax lookup table is a key administrative procedure designed to minimize the compliance burden. The huge increase in the number of use tax payers—from about 8,000 (who filed the ST-44) in 2009 to more than 240,000 (ST-44 filers plus those who paid some use tax when filing their IL-1040 return) in 2010 and 2011—suggests the procedure is effective.

As shown in Table 3 below, almost three out of five Illinois tax filers who paid any use taxes for 2010 and 2011 paid the exact amount indicated by the use tax lookup table.[21] This suggests that many tax filers do not

[20] As discussed above, legislation that would place responsibility for remitting the use tax on out-of-state sellers is under review in the U.S. Congress in 2013.

[21] In Tables 3, 4, 6, and 7, we restrict the sample of returns as noted. These tables are not strictly comparable to Tables 2 and 5 which use a broader sample.

track individual purchases and may find the lookup table a convenient way of reducing compliance costs.[22] Because it decreases perceived compliance costs, the addition of a use tax payment line to the IL-1040 and the addition of a use tax lookup table in the IL-1040 instruction booklet can be responsible for the dramatic increase in the number of use tax filers since 2010.

TABLE 3. Actual Use Tax Payment Compared to Amount Suggested by Use Tax Lookup Table

Tax Year	Use tax taxpayers total	Paid same as Lookup Table	Paid more than Lookup	Paid less than Lookup
2010	242,412	57%	6%	37%
2011	239,900	58%	6%	36%

NOTE: Dependents, returns with over $1mm FAGI, returns with zero or negative FAGI, and returns of nonresidents are dropped from the analysis.

Another factor that may be responsible for the increase in the number of use tax filers is the increase in knowledge about the use tax after 2010. In the absence of survey data about the use tax, the surest way to know that a tax filer is aware of the use tax is that the filer either paid the use tax in a previous year or employed a paid preparer. We know from Table 2 that taxpayers who paid the use tax in 2010 were much more likely to pay the use tax in 2011 than those who did not. This fact might be explained by the hypotheses that people who paid the use tax in 2010 are: (a) more likely to have a use tax liability in 2011; (b) more likely to voluntarily comply with a liability when they have it and hence more likely to pay in 2011; or (c) more knowledgeable about the use tax and thus more likely to comply.

Table 4 provides some additional insight about the relationship between knowledge of the use tax and the probability of payment. We know that paid preparers have been exposed to training about the use tax and are knowledgeable about Illinois tax law. Consistent with Gunter's findings in Maine, we find, as shown in Table 4, that Illinois returns prepared by paid preparers are less likely to have use tax payments than returns that are self-prepared. In 2010, 6.2 percent of self-prepared returns had use tax payments compared with only 4.5 percent of paid preparer returns. Similarly, in 2011, 5.8 percent of self-prepared returns had a use tax payment but only 3.6 percent of paid preparer returns had such a payment.

TABLE 4. Paid Preparers and Use Tax Payments

Return prepared by	Tax Year 2010			Tax Year 2011	
	All tax filers	Use tax filers	Percent that paid use tax	All tax filers	Use tax filers
Self	37.7%	45.6%	6.2%	38.2%	5.8%
Paid preparer	62.3%	54.4%	4.5%	61.8%	3.6%

NOTE: Dependents, returns with over $1mm FAGI, returns with zero or negative FAGI, and returns of nonresidents are dropped from the analysis.

Of course, tax filers who prepare their own returns are likely to be systematically different from tax filers who use paid preparers in a number of ways, so the higher rate of use tax payment among this group is not definitive evidence that use of paid preparers causes reduced use tax payment. To better understand the relationship between the use of a paid preparer and payment of the use tax, we confined our analysis to tax filers who filed in both 2010 and 2011 and who switched between using a self-prepared and paid-preparer return (or vice-versa) in these years.[23] Our results are displayed in Table 5.

[22] Technically, using the use tax lookup table does not relieve tax filers of the obligation to track their use tax obligations. Instructions for Form IL-1040 say: To determine the Illinois Use tax you owe, check your records to see if you were charged tax on internet, mail order, or other out-of-state purchases and use the Use tax (UT) Worksheet to calculate your tax … use the UT Table [i.e., the use tax "lookup table"] to help you estimate the use tax you owe. Enter the Illinois Use tax from the UT Worksheet or UT Table on Form IL-1040, Line 23 …. If we find that you owe additional tax, we may assess the additional tax plus applicable penalties and interest. We conduct routine audits based on information received from third parties, including the U.S. Customs Service and other states. (words in [ ] added).

We speculate that many tax filers believe that using the use tax lookup table is in practice a quasi-guarantee that they will not be charged penalties and interest.

[23] This idea was suggested by Gunter, who found that "a taxpayer who switches to using a preparer is less likely to pay use tax compared to a taxpayer who does not use a preparer in the current or previous year." (p. 9)

TABLE 5. Probability of Use Tax Payment in 2011 Conditional on Whether Use Tax Was Paid in 2010

(Among those who switched from self prepared to paid preparer returns and vice-versa)

2010 use tax payment	2010 return prepared by	2011 return prepared by	Probability use tax was paid in 2011
No	Self	Paid preparer	2.3%
	Paid preparer	Self	1.7%
Yes	Self	Paid preparer	18.0%
	Paid preparer	Self	25.0%

NOTE: Data include all returns matched in 2010 and 2011.

Those tax filers who did not pay a use tax in 2010 and switched from a self-prepared 2010 return to a paid-preparer return in 2011 were *more* likely to pay use tax than those who switched from a paid-preparer return to a self-prepared return (2.3 percent versus 1.7 percent). Contrary to Gunter (2011), this suggests that paid preparers might encourage use tax payments.

However, the story may not be so simple. As we show in the bottom panel of Table 5, among those who both paid use tax in 2010 and switched from self-prepared to paid-preparer returns, the 2011 payment rate of 18 percent is low compared with the 25 percent payment rate of those who paid use tax in 2010 and switched from paid-preparer returns to a self-prepared return. Thus the bottom panel of the Table 5 is consistent with Gunter's finding that paid preparers *reduce* the use tax payment rate but apparently inconsistent with the top panel of Table 5.

The analysis of Table 5 is limited because it does not control for other factors that may affect use tax payment. In the next section, we control for other influences on use tax payment using regression analysis.

V. Regression Evidence About Taxpayer Characteristics That Influence the Probability of Illinois Use Tax Payment

The discussions in the previous two sections were quite informal and designed primarily to provide descriptive information about which Illinois tax filers pay the use tax. In this section, we attempt to be more rigorous by controlling for a number of variables simultaneously that might influence use tax payment. We caution that we do not observe use tax liability and have both a short time series (only 2 years) and a very limited set of information about tax filers—we have only the data reported on Illinois' very simple personal income-tax form—and can make limited causal inferences about use tax payments.

Table 6 shows descriptive statistics about key variables included in our analyses.

Our dependent variable (UT_11) is a dichotomous (1/0) variable that equals one if a tax filer made a use tax payment and zero otherwise. About 4.7 percent of tax filers in our data set made a use tax payment in 2011. Various factors such as a tax filer's attitude toward honesty, risk, government, and their financial situation probably play a role in the decision about whether or not to make a use tax payment in 2010. In all of our regressions, we control for whether the filer made a payment in 2010 (UT_10). About 5.1 percent of filers in our sample made such a payment. Since we use matched 2010 and 2011 tax returns and control for whether the tax filer made a use tax payment in 2010, our regressions should be interpreted as explaining changes in taxpayers' behavior between 2010 and 2011.

TABLE 6. Descriptive Statistics of Key Variables Used in the Analysis

(Based on 5,124,947 IL1040 matched returns in 2010 and 2011; dependents, returns with $1 million or more AGI, returns with zero or negative AGI, and returns of nonresidents are dropped from the analysis.)

Type of variable	Minimum	Maximum	Mean	Std. Deviation	Description
Dependent Variable					
UT_11	0	1	0.05	0.21	1=paid use tax in 2011, 0=not paid
Independent Variables					
UT_10	0	1	0.05	0.22	1=paid use tax in 2010, 0=not paid
agi_2011	1	999,823	62,574	76,695	Federal adjusted gross income
TAXPREP2011	0	1	0.62	0.49	1=used paid tax preparer; 0=self prepared
ESTIMATED PAYMENT	0	1	0.03	0.16	1=made estimated tax payment of more that $500 and under 65 yrs old, 0=not paid
DONATION_DUM_2011	0	1	.004	.067	1=paid check-off donation, 0=not paid
TAX_PMT_DUE_2011	-173,372	621,421	0 67	1,384	positive=refund due; negative=tax pmt due to the state
COUNTY2011	0	1	0.66	0.47	1=border county, 0=nonborder county

Valid Filing Statuses, 2011		
	Frequency	Percent
JOINT	2,103,322	41.0
DEATH	16,104	0.3
MARRIED FILING SEPARATE	63,263	1.2
SINGLE	2,932,581	57.2
WIDOW	9,677	0.2
Total	5,124,947	100.0

We start with a very simple regression specification that includes only a constant and a dummy variable indicating whether a use tax payment was made in 2010 (UT_10) and gradually enter additional variables to better understand the determinants of use tax payments in 2011 (UT_11). It is clear from Gunter's (2011) analyses and from cross-tabulations of our data (not shown) that the probability of paying use tax varies with tax-filer income as measured by FAGI. This correlation is probably due to several factors, including the fact that as tax filers' income rises they become less cash constrained and therefore have less incentive to avoid the use tax. The higher payment rate among higher income tax filers also may be because these tax filers have more complex returns and perceive themselves to have a higher probability of facing an audit.

We have dropped tax filers with negative or zero FAGI and those with FAGI of $1 million or more from our sample: these tax returns may use sophisticated accounting techniques, and the returns may get extra scrutiny and the payment of use taxes may be determined by special considerations. The mean income of tax filers in our sample is more than $62,574 (median is $40,647). In our regression analysis, we enter the natural log of FAGI as an independent variable under the (maintained) hypothesis that the probability of use tax payment increases roughly proportionately with FAGI. We also allow for the possibility that the slope of this relationship changes at high income levels by including a zero/one dummy variable that interacts with FAGI over $250,000.

We also include controls for use of a paid tax preparer for reasons discussed above. As discussed in Gunter (2011), conventional wisdom suggests that small business owners may be particularly savvy about ways to reduce their tax burden and therefore may be less likely to pay the use tax. Unfortunately, the data available from the IL-1040 do not allow us to directly observe whether the filer is a business owner. In lieu of direct evidence, we reasoned that small business owners are more likely to pay quarterly estimated income taxes since they may not have withholding of business income. We create a dummy variable equal to one if the tax filer had estimated tax payments of more than $500 (indicating taxable income of at least $10,000) in 2011 and did not claim an exemption for age 65 or older (which might indicate that filer is a retiree). We suspect that many, but perhaps not all, of the returns that met these two conditions were filed by small business owners.[24] Thus, the estimated coefficient on this variable gives some indication about whether small business owners were more likely to pay use taxes.

We also include a dummy variable equal to one if the filer elected to make a voluntary donation to a number of designated Illinois charities in conjunction with the filing of their tax return. Only about 0.4 percent of tax filers make such a donation. This select group is probably either extremely generous or may feel unusually economically secure, so it would not be surprising if this group also was likely to pay the use tax.

We include a set of variables measuring filing type (single, married filing jointly, and so forth) because we speculate that tax filers may take the views of their partner into account when making ambiguous tax choices. We speculate that holding FAGI constant, couples filing jointly may be more likely to pay the use tax because they want to avoid putting a partner at risk of a negative outcome. Of course, couples with different filing types may be very different in a variety of ways for which we cannot fully control, so that our analyses should be taken as suggestive rather than definitive.

We include a dummy variable equal to one in Illinois' border counties (where 66 percent of filers live) and zero elsewhere. The idea here is that those who live in border counties may have less incentive to shop via the Internet because the retail stores in their home jurisdiction face more intense interstate competition.[25] We hypothesize that because those in border counties might be less likely to shop via the Internet, they might be less likely to pay the use tax.

Since tax filers may be loss averse, they may be less likely to pay a use tax if they owe taxes than if they have a refund due. Because of this, we include a variable measuring the refund or tax payment due on the return and allow it to enter the regression equation asymmetrically.[26]

Table 7 shows our regression results. Because we have a very large sample size, almost all of the t-statistics on the regression coefficients are quite large. The column I regression has a simple interpretation. It shows that the probability of a filer paying the use tax in 2011 is just 2 percent if they did not pay use tax in 2010. A filer who did pay use tax has a nearly 58-percent probability (= 0.021 + 0.556) of paying the use tax. All of the other coefficients in the regression table are quite small relative to the coefficient of UT_10.

[24] We acknowledge that this variable does not identify small business owners since the number of returns that met these conditions is far lower than the number of federal returns in Illinois that report schedule C (business) income.

[25] Agrawal 2011 provides evidence that proximity to state borders results in increased sales-tax competition.

[26] In 2011 there was an unusually small number of Illinois tax filers with a refund due. The reason for this is that when Illinois increased its (flat) personal income-tax rate in January 2011, many Illinois employers did not increase withholding in a timely fashion. As a result, many tax filers found that they had to make a payment or got a smaller refund when they filed their 2011 taxes in the first quarter of 2012. We see no reason why this decrease in refunds due should affect the interpretation of coefficients that measure the relationship between the refund due and the probability of a use tax payment.

TABLE 7. Illinois Use Tax Regression Results

Dependent Variable 2011 Use Tax Dummy:

Data: 2010 and 2011 Illinois individual income tax returns excluding those of dependents and nonresidents as well as those with zero or negative AGI or with AGI over $1 million.

Number of Records 5,124,947

VARIABLE	VARIABLE	MODEL							
		I	II	III	IV	V	VI	VII	VIII
	Intercept	0.021	-0.088	-0.081	-0.076	-0.076	-0.059	-0.059	-0.059
2010 USE TAX PMT	2010 use tax dummy; 0=no; 1=use tax paid	0.556	0.550	0.549	0.548	0.548	0.547	0.547	0.547
2011 ASYM_LOGAGI	LOG_AGI		0.010	0.010	0.010	0.010	0.008	0.008	0.008
	ASYM_LOG=0 if FAGI<$250K; LN(FAGI) if FAGI>$250K		0.001	0.001	0.001	0.001	0.001	0.001	0.001
2011 TAX PREPARER	Paid preparer dummy; 0=self, 1=paid preparer			-0.011	-0.011	-0.011	-0.011	-0.011	-0.011
2011 ESTIMATED PAYMENT	0=no pmt,1=Pmt over $500 under age 65				0.031	0.031	0.030	0.030	0.030
2011 DONATION-CHECK-OFF	Donation dummy; 0=no 1=yes					0.037	0.037	0.037	0.037
2011 FILING TYPE	Single						Omitted	Omitted	Omitted
	Head of household						-0.010	-0.010	-0.010
	Joint filing						0.004	0.004	0.004
	Separate filing						-0.002	-0.002	-0.002
	Widow						0.007	0.006	0.006
	Deceased						-0.005	-0.005	-0.005
Border County	Border county dummy; 0=no; 1=border county							-0.001	-0.001
FINAL_REFUND_2011	Refund_due: positive=refund due, neg=tax pmt								-0.00000007
PMT_ASYM_2011	FINAL_REFUND_2011 * PMT_ASYM_DUM_2011; 1 = Refund_Due > 0, 0 = Refund Due <0								0.00000033

*T statistics for all coefficients are greater than or equal to three except for Final_refund_2011 which has a t stat of 0.7 in column 8 and PMT_ASYM_2011 which has a t stat of 2.5 in column 8. Adjusted R-squareds are about .32 for all regressions.

The coefficients in column II show that the probability of paying the use tax increases with FAGI but not at a very fast pace. According to our regression results, a tax filer with the mean FAGI of $62,574 who did not pay the use tax in 2010 would have a probability of paying the use tax of about 2.5 percent. A tax filer with twice as much income would have a probability of paying the use tax of only about 3.2 percent.[27]

The coefficient of ASYM_LOG allows the slope of the relationship between the probability of a use tax payment and FAGI to change for incomes above $250,000. The coefficient on ASYM_LOG suggests that the probability of use tax payment rises quite rapidly with FAGI above $250,000. According to the regression coefficients in column II, a tax filer with FAGI of $125,000 would have a probability of paying use tax of about 3.2 percent. Doubling FAGI to $250,000 increases the probability to only about 3.9 percent; however, doubling FAGI again (to $500,000) increases the probability of a use tax payment to about 6.1 percent.

[27] Simulated probabilities are calculated by multiplying coefficients by appropriate values of independent variables. For example 2.5% = -0.888 + 0.010 x ln(62574). Other simulated values are calculated analogously.

After we control for income and payment of the use tax in the previous year, filing a self-prepared return (rather than having a return submitted by a paid preparer) has essentially the same impact on the probability of making a use tax payment as doubling the tax filer's FAGI. This coefficient is essentially unaffected by the addition of more independent variables in columns IV to VIII.

In column IV, we add a dummy variable equal to one if an estimated payment was made and zero otherwise. The coefficient of this variable is quite large, indicating that making an estimated payment is associated with a 3-percent increase in the probability of paying the use tax. This is perhaps counter-intuitive since it suggests those who make estimated payments, who we speculate are likely to be business owners, are much more likely to pay the use tax than a randomly selected filer.

As we noted above, about 0.4 percent of tax filers make a voluntary donation to one of the Illinois charities indicated on the IL-1040. As shown by the estimated coefficients in columns V to VIII, this group of donors is much more likely to pay use tax than others. A tax filer who did not pay use tax in 2010 but made a check-off donation has a 6.2-percent probability of making a use tax payment compared with a 2.5-percent probability for a tax filer who did not make a check-off donation. This is consistent with the hypothesis that use tax payment is, in effect, similar to a charitable donation. This is consistent with Gunter's (2011) finding using Maine data.

In column VI, we added various dummy variables measuring filing type to the regression. The estimated coefficients measure the impact of filing type relative to the omitted category of a single filer. A head of household (nonmarried person with one or more dependents) is about 1 percent less likely to pay use tax while widows or widowers and those filing jointly are more likely to pay the use tax. Since the baseline probabilities of paying the use tax are low (about 2 percent for a filer who did not pay use tax in 2010), the impact of filing type is relatively large. The baseline probability of a use tax payment by a formerly single person who became a single parent in 2011 (and therefore switched filing status from single to head of household) would fall from about 2 percent to only about 1 percent. If instead she married in 2011, her baseline probability of paying the use tax would be almost 2.5 percent.

Living in a border county reduces the probability of paying the use tax as hypothesized and shown in columns VII and VIII. The locational impact is quite small, however, living in one of these counties reducing the probability of payment by only about 0.1 percent.

Finally, in column VII we added a variable measuring the refund that is due[28] and allow for asymmetry around zero so that having a refund due is allowed to have a different effect than having to make a tax payment. We included these variables to investigate the role of "loss aversion," which might make tax filers less willing to pay the use tax if paying it requires them to write a larger end-of-tax-year check.

The coefficient for FINAL_REFUND_2011 is not significantly different from zero despite the very large sample size in our analysis. The coefficient on PMT_ASYM_2011 (FINAL_REFUND_2011*PMT_ASYM_2011) is statistically significant but very small. The coefficients are easiest to interpret using numerical examples. The coefficient on FINAL_REFUND_2011 is negative, implying that *the larger the tax payment due, the higher the probability a tax filer will pay a use tax.* In practice, the estimated coefficient suggests that there is essentially no relationship between the amount of tax payment due and the probability a tax filer pays the use tax.

Similarly, according to the regression results, the estimated coefficient suggests essentially no relationship between the amount of refund due and the probability of a tax filer paying the use tax.

VI. Discussion

Authors of economics textbooks recommend that we evaluate systems of government finance based on equity, efficiency, and administrative ease (Hyman 2002). By these criteria, the use tax has both virtues and flaws. The use tax is a modification intended to make the sales tax more equitable and efficient by reducing or eliminating

[28] This value is negative if a tax payment is due.

behavioral distortions that would arise if out-of-state purchases for in-state use were taxed at a different rate than in-state purchases. A universally imposed use tax would remove the incentive for buyers to purchase equally priced identical goods over the Internet or from cross-border shopping. This could reduce wasteful travel and shipping and increase the efficiency of the tax system. A universally imposed use tax also would increase horizontal equity as those who live near a low-tax border and those with easy access to Internet purchases (e.g., those with credit cards) would no longer pay a lower after-tax price than other citizens.

Unfortunately, imposition of the use tax is far from universal. Use tax compliance by individual consumers is essentially voluntary for many out-of-state purchases. Compliance has been limited both because narrowly self-interested consumers benefit from noncompliance and because even consumers with a broader or more risk-averse conception of self-interest previously had no low-cost method of complying with the use tax. The addition of a use tax payment line on the IL-1040 and the associated lookup table was designed to reduce compliance costs of tax filers.

The reduction in compliance costs associated with the addition of a use tax payment line on the IL-1040 and the associated lookup table appears to have increased revenue from the use tax but may have done little to increase the efficiency of the sales tax and may have reduced horizontal equity. Because a majority of Illinois use tax-paying filers use the lookup table, it functions as essentially a 0.06-percent personal income-tax surtax for many compliers. The addition of a use tax line to the IL-1040 will not diminish behavioral distortions if compliers simply use the lookup table to assess their tax liability because tax payments will not vary with the level of low-tax purchases. Thus, taxpayers will have no incentive to curtail low-tax purchases. Furthermore, the addition of a use tax payment line on the IL-1040 could reduce horizontal equity because risk-averse, naïve, or charitable tax payers may comply while others do not. Because of this, taxpayers in similar circumstances with similar incomes may pay different taxes.

VII. Conclusion

Illinois' use tax is legally mandated but difficult to monitor and collect from individual buyers. Since Tax Year 2010, the Illinois personal income-tax form has contained a line so that tax filers can remit their use taxes with minimal administrative complexity. We examined data from a panel of 2010 and 2011 Illinois personal income-tax returns. With this short time series and limited number of independent variables, we were not able to make definitive causal inferences about factors that determine filers' use tax payments.

However, our analysis does show that only a small fraction of tax filers pay any use tax. By far the most important predictor of a 2011 use tax payment is a 2010 use tax payment, suggesting that persistence is an important factor. Consistent with Gunter—so far as we are aware, the only other empirical research directly on our topic—we find that the probability that a tax filer pays the use tax increases with income and is associated with charitable contributions. Like Gunter, we find evidence consistent with the hypothesis that business owners are more likely to pay use tax than randomly selected tax filers. We also find that filing type matters, although the impact is relatively small. In contrast to Gunter, we find mixed evidence about the impact of paid preparers on use tax payment. We find only small impacts from other variables that we examined.

We find that about 60 percent of tax filers that pay the use tax pay exactly the amount suggested by the use tax lookup table. It may therefore be advisable for Illinois and other states to undertake additional research to refine the payments suggested by use tax lookup tables.

APPENDIX
How Many Use Tax Payers Will There Be in the Long Run?

Policymakers might wish to know the share of filers that will pay the use tax in the long run. Illinois has very limited experience in this area so it is difficult to make an empirical estimate. Based on our analysis of the first two years' experience, however, two facts are salient. First, only a small percentage of taxpayers pay any use tax. Second, use tax payment is persistent. Those who paid in Year 1 are much more likely to pay in Year 2 than the average. The high level of persistence of those who do pay the tax might lead to the hope that eventually such behavior will become widespread. On the other hand, our analysis suggests that even in the best-case scenario, persistent payment behavior would lead to only modest growth in the number of use taxpayers. To understand this claim, suppose that

$$(1) \quad P_t = \lambda F_t$$

Where P_t = the number of filers paying the use tax in time period *t,* λ = a constant, F_t = total number of IL-1040 returns in time period *t,* and *t* = the first year in which the use tax line is on the IL-1040 form. Further assume that the number of tax filers is approximately constant from year to year so that $F_{t+1} = F_t = F$.

In Illinois, tax filers who paid the use tax in 2010 had an almost 50-percent chance of paying the tax again in 2011. To capture this fact, suppose

$$(2) \quad P_{t+n} = .5P_{t+n-1} + \lambda\left(F - P_{t+n-1}\right) = \left(.5 - \lambda\right)P_{t+n-1} + \lambda F$$

Equation (2) says that half of the tax filers that pay the use tax in a given year also pay the use tax in the next year, which captures the persistent behavior that we observe thus far in Illinois. Equation (2) also says that other taxpayers who did not pay the use tax in the previous year are just as likely to pay the use tax as a randomly selected taxpayer was in the first year, and therefore the probability that they will pay is the same λ. This is probably an overly optimistic assumption since those who did not pay in Year 1 are less likely to be disposed to pay than the average in Year 2. Nonetheless, we adopt this assumption since it provides a "best case" scenario when projecting the number of future use tax payers. Equation (2) implies that

$$(3) \quad P_{t+n} = \left(.5 - \lambda\right)\left(\left(.5 - \lambda\right)P_{t+n-2} + \lambda F\right) + \lambda F = \left(.5 - \lambda\right)^2 P_{t+n-2} + \left(.5 - \lambda\right)\lambda F + \lambda F$$

Generalizing this

$$(4) \quad P_{t+n} = \left(.5 - \lambda\right)^n \lambda F + \left(.5 - \lambda\right)^{n-1} \lambda F + \left(.5 - \lambda\right)^{n-2} \lambda F + \ldots + \left(.5 - \lambda\right)\lambda F + \lambda F$$

rewrite as

$$(5) \quad P_{t+n} = \left(.5 - \lambda\right)^n \lambda F + \sum_{1}^{n} \left(.5 - \lambda\right)^{n-i} \lambda F$$

we can rewrite equation (5) as

$$(6) \quad \frac{P_{t+n}}{F} = \lambda \sum_{0}^{n} \left(.5 - \lambda\right)^{n-i}$$

If $0 < \left(.5 - \lambda\right) < 1$ then $\sum_{0}^{n} \left(.5 - \lambda\right)^{n-i}$ converges, and this equation can be used to predict the share of taxpayers that will pay the use tax at any point in the future.

Consider the simple example with λ =4% for N≤20 displayed in the table below. With these parameters, Illinois starts out with about 4 percent of filers paying the use tax. This rises gradually to about 7 percent of filers by the third year (Year 3) since the use tax line is on the tax form but remains at 7 percent in subsequent years. Of course, the assumptions used in creating the table are very simple, but they make the general point that, even with a rate of payment-persistence of 50 percent, Illinois should expect to experience only a short-term increase in the number of use tax payers. In the long-term, the share of use tax payers will remain quite small, given the parameters we see in the first 2 years of the program.

Years since use tax payment line was put on 1040	Share of general filers paying use tax	Predicted share of filers that pay use tax if one-half of those paying in the preceding year pay in next year
t	lambda	Share paying
1	4%	4%
2	4%	6%
3	4%	7%
4	4%	7%
5	4%	7%
6	4%	7%
7	4%	7%
8	4%	7%
9	4%	7%
10	4%	7%
11	4%	7%
12	4%	7%
13	4%	7%
14	4%	7%
15	4%	7%
16	4%	7%
17	4%	7%
18	4%	7%
19	4%	7%
20	4%	7%

References

Agrawal, David R. 2011. The Tax Gradient: Do Local Sales Taxes Reduce Tax Differentials at State Borders? (November 11, 2011). Available at SSRN: http://ssrn.com/abstract=1909035 or http://dx.doi.org/10.2139/ssrn.1909035.

Alm, James, and Mikhail I. Melnik. 2012. "Cross-Border Shopping and State Use Tax Liabilities: Evidence from eBay Transactions," Public Budgeting and Finance 32(1):5–35.

Anderson, John E. 2013 "Paying the State Use Tax: Is a 'Nudge' Enough?" Department of Economics, University of Nebraska-Lincoln, mimeo.

Chernick, Howard, and David Merriman 2013. "Taxes and Cigarettes: Using Littered Pack Data to Estimate Tax Avoidance in NYC" National Tax Journal 66(3):635–668 (September 2013).

Chupick, Andy, and Natalie Davila. 2009. "A New Method for Estimating Illinois' E-Commerce Losses," Illinois Department of Revenue (Available at http://www.tax.illinois.gov/AboutIdor/TaxResearch/InternetSalesTaxLosses.pdf).

Einav, Liran, Dan Knoepfle, Jonathan D. Levin, and Neel Sundaresan. 2012. "Sales Taxes and Internet Commerce," NBER Working Paper Series (Working paper 18018) http://www.nber.org/papers/w18018.

Gunter, David. 2011. "Who Pays Use Tax?" Maine Revenue Service (unpublished manuscript cited with permission from the author). See http://www.taxadmin.org/fta/meet/12rev_est/pres/gunter.pdf for a related PowerPoint presentation.

Hyman, David. 2002. Public Finance: A Contemporary Application of Theory to Policy. 7th edition. Fort Worth, TX: Harcourt College Publishers.

Madden, Mary, and Lee Rainie. December 2003. "America's Online Pursuits" Pew Internet & American Life Project available at http://www.pewinternet.org/~/media//Files/Reports/2003/PIP_Online_Pursuits_Final.PDF.PDF.

Manzi, Nina, updated April 2012. "Use Tax Collection on Income Tax Returns in Other States," Research Department Minnesota House of Representatives. (Available at http://www.house.leg.state.mn.us/hrd/pubs/usetax.pdf.)

Merriman, David. 2010 "The Micro-Geography of Tax Avoidance: Evidence from Littered Cigarette Packs in Chicago," American Economic Journal: Economic Policy 2(2):61–84 May 2010.

Slemrod, Joel. 2007. "Cheating Ourselves: The Economics of Tax Evasion" The Journal of Economic Perspectives, Vol. 21, No. 1 (Winter, 2007), pp. 25–48.

Slemrod, Joel, and Jon Bakija. 2008. Taxing Ourselves, 4th edition: "A Citizen's Guide to the Debate Over Taxes." 4th ed. Cambridge, MA: MIT Press.

Shedding Light Into the Black Hole: Research and Evaluation Into the Hidden Economy

Alice Cleland,
Inland Revenue, New Zealand

1.0 Introduction

The Hidden or Cash Economy[1] undermines the integrity of the tax system. Moreover, evidence from the literature suggests that in times of economic recession, Hidden Economy activity is likely to increase with taxpayers opting out of the system.

To address this risk, funding was allocated to Inland Revenue New Zealand by the New Zealand Government in 2010 to undertake work on the Hidden Economy Initiative. An approach was developed to influence Hidden Economy compliance and social norms by making it easy for people to comply, and by creating a challenging environment for those operating within the Hidden Economy.

A longitudinal and multi-tiered approach was proposed that was intelligence-led and evidence-based to move people and businesses from noncompliance or partial compliance to a social commitment to meet their tax and social obligations.

To this end, a comprehensive programme of work was developed by the National Research & Evaluation Unit, to:

- Better understand the Hidden Economy risk;
- Identify the enablers and barriers to compliance;
- Inform the design of interventions to mitigate/treat those risks; and
- Assess the effectiveness of the implemented interventions/"treatments", and specifically, whether they have resulted in long-term improvement in compliance/behaviour change.

The purpose of this paper is to provide an overview of Inland Revenue New Zealand's Hidden Economy Research & Evaluation Programme. The conceptual basis of the Programme will first be discussed (Section 2.1), followed by an outline of the components of the Programme (Section 2.2); highlights of some of the key research and evaluation findings (Section 2.3); the lessons learned and implications discussed (Section 2.4).

2.0 Hidden Economy Research & Evaluation Programme

The overarching objective of the Hidden Economy Research & Evaluation Programme is to provide an ecological, longitudinal, and evidence-based approach that positively influences taxpayer compliance (refer to Appendix A). The specific objectives of the Programme are to:

- Provide insight into customer compliance attitudes, behaviours, and norms.
- Improve understanding of the drivers of noncompliance for different segments of the customer base.

[1] The Hidden Economy is defined as comprising of monetary and non-monetary transactions intentionally not declared or accurately reported. The key characteristics/behaviours that identify people operating within the Hidden Economy include: being outside the tax system; being inside the tax system but not filing; underreporting their income/inflating their expenditure; earning income from illegal activities; deliberately using fraud.

- Inform the development and implementation of Inland Revenue's Hidden Economy strategy.
- Inform the design and implementation of Hidden Economy interventions/ treatments to ensure that they remain responsive to the identified risks.
- Keep Inland Revenue informed about international best practice, and to ensure that Inland Revenue's practices remain relevant and innovative.
- Increase research and evaluation collaboration, and information sharing with other tax jurisdictions.

To achieve the Programme objectives, research and evaluation activity employing a mixed method approach was used.[2]

2.1 Conceptual Basis of the Programme

The Hidden Economy Research & Evaluation Programme was developed after having conducted a literature review (which included an overview of the Hidden Economy research being conducted by other tax jurisdictions), key informant interviews (with subject matter experts within Inland Revenue), and consulting internal strategic documents. The findings from this review highlighted the knowledge gaps both nationally and internationally and how the National Research & Evaluation Unit could fill some of those gaps. A conceptual framework for Hidden Economy research & evaluation and an evaluation strategy were developed.

One of the conclusions from the literature review was the acknowledgement that tax compliance behaviour was complex and that there are myriad factors influencing whether customers choose to comply with their tax obligations. There was also the recognition that there was a "psychology" to tax compliance and that for behaviour change to occur, a holistic view of the customer was needed.

An ecological framework for conceptualising the factors influencing taxpayer compliance was therefore proposed and provides the theoretical basis of the Hidden Economy Research & Evaluation Programme. While the approach itself is not new (refer to Australian Public Service Commission, 2007; Central Office of Information, 2009; Halpern et al., 2004), it does provide a taxonomy that simplifies the otherwise complex myriad of factors that are likely to be influencing taxpayer attitudes and behaviours.

On applying this framework, it is postulated that taxpayer compliance is influenced by factors at the following levels:

- Individual level: factors that are intrinsic to the person (such as their attitudes and beliefs), as well as their demographic characteristics (such as their age, gender etc.);
- Social level: the influence of their peers and social networks etc.; and
- Environmental level: the impact of the wider environment, (such as natural disasters, the changing nature of technology etc.).

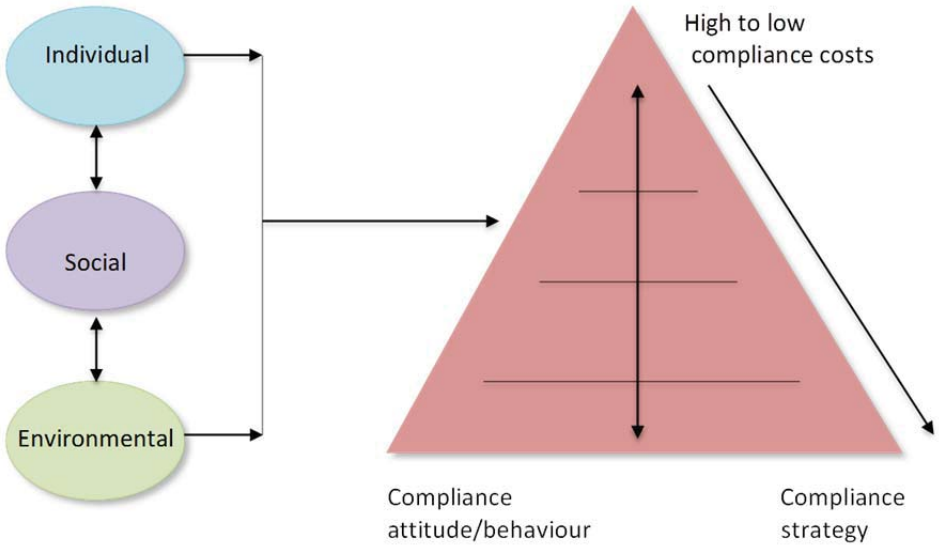
Figure 1 provides a diagrammatic illustration outlining how an ecological framework could be applied to Inland Revenue's compliance model.[3] A more comprehensive overview of the factors that are likely to influence taxpayer compliance is provided in Appendix B.[4]

[2] This approach includes not only the analyses of primary sources of data (including survey work with taxpayers), but also analyses of administrative data and secondary data sources/information where appropriate. It is hoped that such a triangulated approach to data collection would address inherent limitations associated with survey methodology.

[3] Please note that the purpose of the ecological framework is to provide a taxonomy for research and evaluation activity; it is not intended to replace Inland Revenue New Zealand's BISEP compliance model. For more information about the BISEP model, refer to Braithwaite (2003).

[4] While the list of factors provided in Appendix B is not exhaustive, the model was derived after a review of the national and international literature on tax compliance/Hidden Economy.

FIGURE 1. An Ecological Approach to Understanding Taxpayer Attitudes and Behaviours



2.2 Components of the Programme

After applying an ecological approach to the design of the Hidden Economy Research & Evaluation Programme, seven key components emerged. Refer to Figure 2. Each of the identified areas is now discussed.

FIGURE 2. Overview of the Components of Inland Revenue New Zealand's Hidden Economy Research & Evaluation Programme

2.2.1 Understanding Individuals' HE Attitudes, Behaviours, and Drivers

Developing insights into taxpayers' attitudes and behaviours, and understanding the drivers of tax compliance, remain at the core of the Hidden Economy Research & Evaluation Programme.

As outlined earlier, human behaviour is complex and is influenced by a myriad of factors that may either be internal (e.g., cognitions, attributions) or external (e.g., the global financial crisis) to the individual. The need to better understand the psychological and behavioural drivers of behaviour supports international trends on tax compliance research, which have seen shifts away from the use of "economic models" to account for taxpayer behaviour.

Thus, one of the first pieces of research to be completed from the Programme included a national survey of Hidden Economy attitudes and behaviours. The purpose of this survey was to establish a baseline measure of Hidden Economy attitudes and behaviour, and to monitor these trends in the long term. The survey has been able to provide Inland Revenue with a measure of:

- The demand and supply side of the Hidden Economy (and in particular the level of social acceptance for cash job activity);
- The prevalence of cash job activity in the general populace;
- The types of industries where cash job activities may be occurring;
- And general public perception of the effectiveness of Inland Revenue in being able to detect such activity.

As Inland Revenue's programme of work on the Hidden Economy further progresses, it is anticipated that findings from research such as this would inform the development of large-scale activities to shape tax compliance behaviour (such as national and regional social marketing campaigns).

2.2.2 Understanding Businesses' HE Attitudes, Behaviours, and Drivers

Ensuring that businesses receive the right level of support, that the tax system remains responsive to customer needs, and maintaining long-term customer compliance are high-level objectives that most tax administrations are striving to achieve.

For businesses that are operating within the tax system, the type of noncompliant behaviour that may be associated with Hidden Economy activity could range from poor recordkeeping, underreporting of income, not filing or paying (or late filing and payment), aggressive tax planning, fraud and evasion activities, to name but a few. Owing to the "hidden" nature of Hidden Economy activity, developing a better understanding of how and why some businesses are choosing to operate outside the tax system would also be an equally important knowledge gap to fill.

To develop a better understanding of businesses and their tax compliance behaviour, with particular focus on Hidden Economy activity, one project that was completed under the Hidden Economy Research & Evaluation Programme was research on small and medium-sized enterprises (SMEs) from the construction and trade industry. The reason the construction and trade industry was researched was because of Inland Revenue's current compliance activities on this industry. This research helped to:

- Develop an understanding of the characteristics of SMEs from the industry (such as how they sourced work; how income was generated; how they managed their bookkeeping etc.);
- Identify their attitudes towards tax and the Hidden Economy, and the types of Hidden Economy activity they were engaging in;
- The drivers of compliance/noncompliance and the characteristics of compliers/noncompliers;
- Identify strategies/interventions that could be adopted by Inland Revenue to facilitate industry compliance behaviour.

Consideration is now being given to how the findings of this research could be used to design, develop, and implement interventions that could be applied to promote industry compliance.

2.2.3 Understanding HE Risk Posed by Special Groups of Interest

In addition to developing insight into general customer behaviour, there is also a need to better understand the tax compliance behaviour of different segments of the customer base. For example, one-quarter of New Zealand's population are now overseas-born, with Asia being the most common region of birth for those born overseas (Statistics New Zealand, 2013).

In response to the changing nature of New Zealand's population, the needs of "ethnic SMEs" have been investigated (Chand & Cleland, 2012), and a literature review on immigrant entrepreneurship and tax compliance completed (Yuan, Cain, & Spoonley, 2013).

Research exploring how being a migrant shapes tax compliant behaviour has also been recently completed. Some of the objectives of this research include developing a baseline measure of:

- The characteristics of migrant SMEs (such as how they structure their business; manage their bookkeeping; where their sources of income are from etc.);
- Their compliance attitudes and behaviours, and the drivers of compliance;
- The influence of "cultural distance" on compliance; and
- Factors that would foster migrant compliance.

It is anticipated that findings from this research would contribute to the knowledge vacuum that currently exists on migrant businesses in New Zealand and inform Inland Revenue's compliance activities on this segment of the customer base.

2.2.4 Investigating the Impact of Changes in the Tax Environment

The application of an ecological approach emphasises the need to apply a holistic framework to researching the Hidden Economy, including the need to consider not just factors at the micro level but also macro level influences on taxpayer compliance. These may include, for example, taking into consideration developments in the national and international tax environment, and the impact of these on the evolving business needs of tax administrations.

Two such macro-level influences that were researched as part of the Hidden Economy Research & Evaluation Programme were to:

- Explore the impact of adverse events on SME tax compliance behaviour (with particular focus on the Christchurch 2010 and 2011 earthquakes in New Zealand); and
- Investigate how technology may be acting as an enabler of noncompliance.

2.2.4.1 The Impact of Adverse Events

The Christchurch earthquakes of 2010 and 2011 are continuing to have an effect on New Zealand's economy, and will probably continue to do so for many years. Where challenging circumstances are created as a result of challenging events, there is a need to understand the "challenges" that may arise so that any potential risks can be mitigated, the right level of support is provided (to minimise opportunities for noncompliance), and that the lessons learned could be applied in the future should similar circumstances arise.

Thus, a longitudinal project on the impact of adverse events was initiated (refer to Poppelwell et al., 2012; and Poppelwell et al., 2013). The research has provided Inland Revenue with:

- A baseline measure of Christchurch SME's attitudes and behaviours towards tax compliance;
- Identified what impact Inland Revenue's presence has had on SME tax compliance behaviour; and
- Established how the different phases of disaster recovery have affected SME compliance.

A three-to-five-year longitudinal project has been scoped and is currently informing Inland Revenue's compliance initiatives in Christchurch.

2.2.4.2 Use and Application of Technology

Technological advances are occurring at a phenomenal rate, and as are the increased opportunities not to comply and engage in noncompliant activities via electronic means. In particular, the advent of the internet has implications for all tax administrations and poses a potential risk to revenue collection.

To better understand some of the compliance risks associated with electronic commerce (e-commerce)/ information communication technology (ICT), research has been undertaken to develop a baseline measure of:

- Online, and in particular, off-shore consumer behaviour (such as types of goods/services that are being bought and the value of those goods);
- The motivators/barriers behind online purchasing behaviours;
- Tax compliance attitudes and behaviours (with particular focus on whether technology may be abetting individuals to operate outside the tax system).

The aforementioned research provides: high-level background information on the potential size of the e-commerce industry within New Zealand; the prevalence of consumers who may be importing goods below the de minimis threshold to avoid paying GST on those goods; and the implications of this for tax revenue collection for Inland Revenue.

2.2.5 Inform HE Strategy and Intervention Design and Implementation

According to Inland Revenue's Hidden Economy Approach, Inland Revenue is committed to applying an "evidence-based approach" to address the Hidden Economy risk. Such an approach, therefore, requires the need for research to aid in the understanding of the customer base, provide the contextual background for compliance initiatives, and inform the design and development of such interventions.

As the Hidden Economy Programme develops, it is envisaged that the National Research & Evaluation Unit would work closely with its colleagues to continually inform Inland Revenue's strategy for the Hidden Economy and help design initiatives that are based on an understanding of the customer and is responsive to their needs.

2.2.6 Evaluating the Effectiveness of HE Interventions/Programs of Work

In their paper on how other tax authorities tackled the Hidden Economy, the UK's National Audit Office (2008) identified that one gap in the research was that there was "little published information available on the cost-effectiveness of the measures used by tax authorities to tackle the Hidden Economy" (National Audit Office, p.9, 2008). What this highlights is the need to have robust evaluation processes in place to assess whether the interventions/programs of work applied have been effective.

Moreover, even prior to the implementation of any interventions, baseline measures are needed to allow for pre- and post-intervention comparisons. Long-term monitoring is also suggested to help determine whether the interventions have resulted in longstanding changes in compliance behaviour.

The need for a more systematic and "experimental" approach to intervention/policy design and implementation is supported by the work of the Behavioural Insights Team in the UK, who promote the use of randomised control trials when developing public policy (refer to Haynes et al., 2012). How this approach could be applied to Hidden Economy research and evaluation is worth considering further.

Conducting evaluations is an integral part of the Hidden Economy Research & Evaluation Programme, and an evaluation strategy was developed to assess the effectiveness of Inland Revenue's compliance activities on the Hidden Economy. The development of the strategy was guided by the intervention logic for the Hidden Economy Programme – refer to Appendix C.

One example of an evaluation project completed under the Programme was an assessment into the effectiveness of Inland Revenue's campaign on the hospitality industry, and in particular, the restaurant sector. The evaluation was able to provide:

- Insights into restaurant owners' motivators and barriers to compliance;
- Suggestions on how Inland Revenue could improve on the delivery of its compliance initiatives within the sector;
- Ideas on how awareness of Inland Revenue's compliance activities could be generated.

2.3 Highlights From the Programme

Research and evaluation activity on the Hidden Economy has been ongoing in the past few years, and a number of outputs have been produced from the Hidden Economy Research & Evaluation Programme. Below is a high-level synthesis of some of the key findings. The findings are now informing Inland Revenue's strategy towards the Hidden Economy and the development and implementation of compliance initiatives.

2.3.1 Measuring the Size of the Hidden Economy

The size of the Hidden Economy in New Zealand is unknown. The merits of deriving a single figure to represent the size of potential revenue loss to the New Zealand Government through Hidden Economy activity was explored, in a review of the different methodological approaches to measuring the tax gap. The review found that while it may be conceptually appealing to present a quantifiable figure of the potential size of the Hidden Economy, deriving this figure was problematic for the following reasons:

- There are currently no standard approaches for assessing and interpreting the tax gap;
- Tax authorities apply different definitions of the tax gap which makes it difficult to directly compare tax gap figures across different countries;
- Tax gap calculations are subject to assumptions that cannot always be verified in practice; and
- The statistical robustness and accuracy of current tax gap measurements was queried.

Owing to the methodological issues inherent in all tax gap measurements, it was concluded that any measurements pertaining to the tax gap needed to be treated with caution. The conclusions from the paper supported Inland Revenue's current stance not to estimate the size of the Hidden Economy and/or tax gap within New Zealand.

2.3.2 Scale of the Hidden Economy

As mentioned earlier, when the Hidden Economy Research & Evaluation Programme was developed, research to better understand the scale and nature of Hidden Economy activity at the wider societal level was identified as a priority.

In response to this gap in knowledge, research was conducted at national, regional, and business levels to investigate the:

- Level of Hidden Economy activity in the wider community;
- Perceived prevalence of Hidden Economy activity;
- Level of social acceptability/tolerance for such activity;
- Perceived deterrence effect of Inland Revenue.

Supporting the international literature on cash job activity, (e.g., European Commission, 2007), the research found that construction and repair services and home assistance services were the two most common industries where cash jobs were bought and sold from (refer to Figures 3 and 4).

FIGURE 3. Industries Cash Jobs Were Bought From

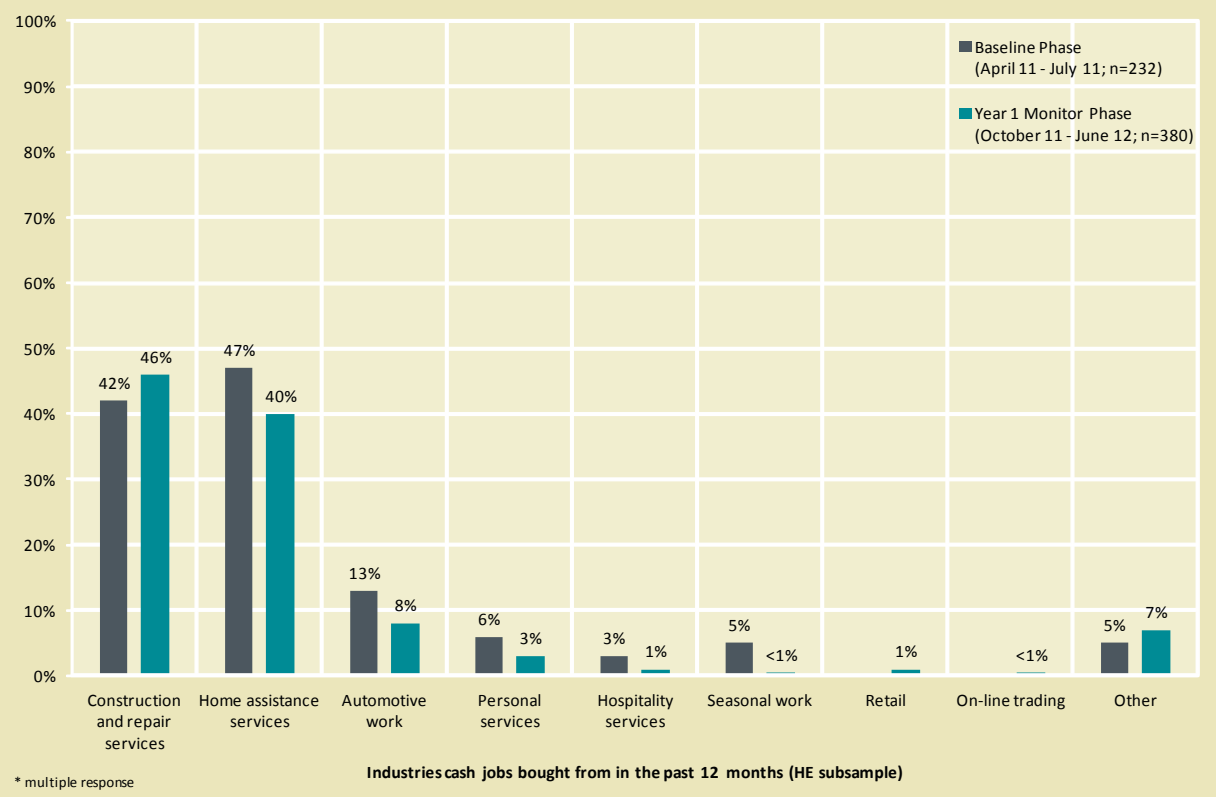
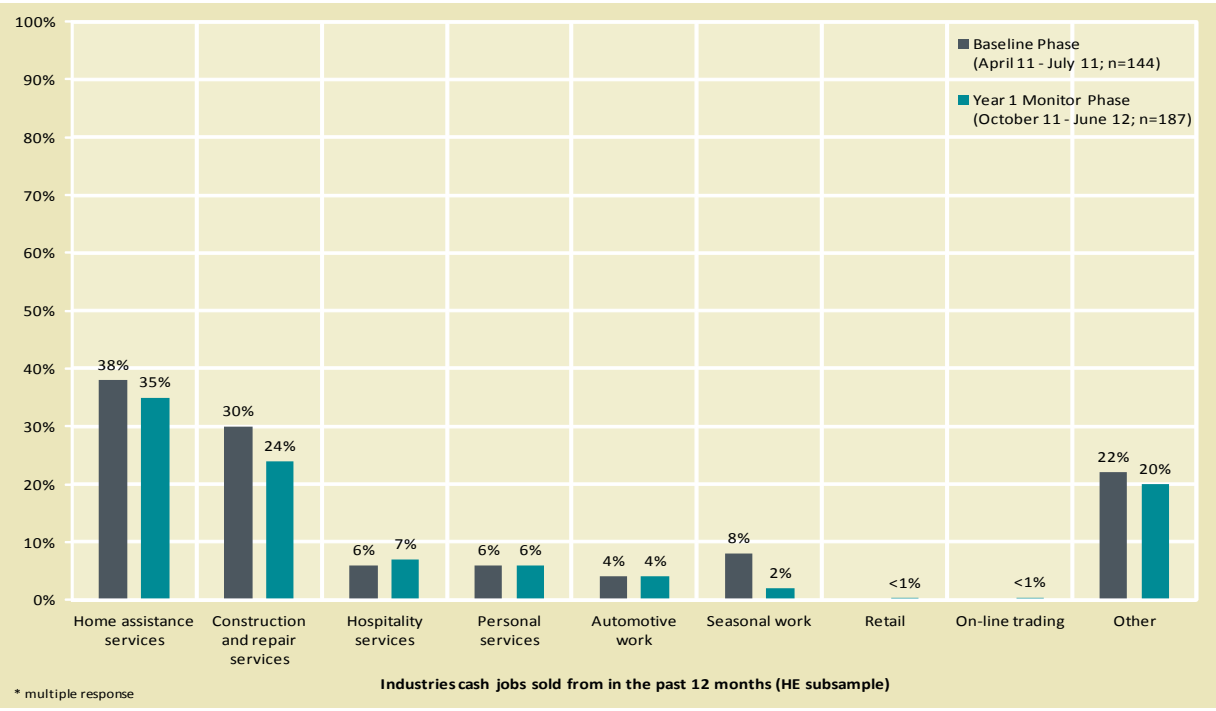
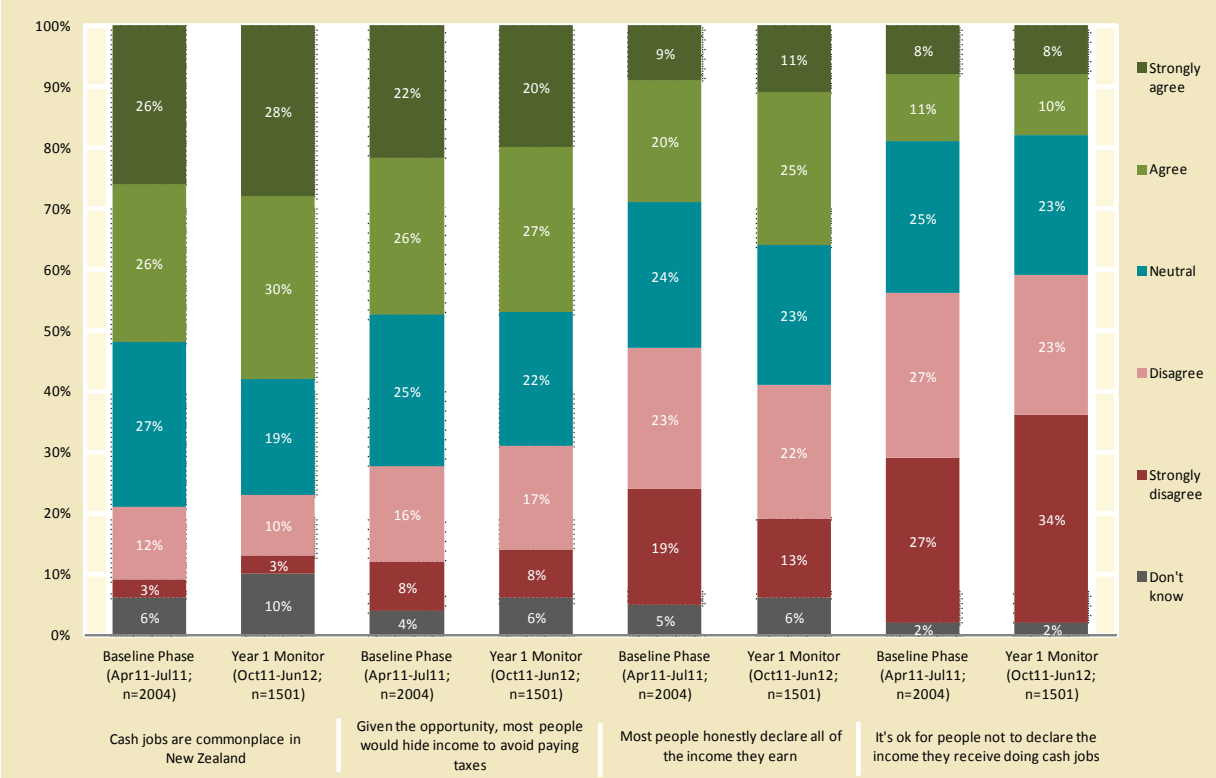


FIGURE 4. Industries Cash Jobs Were Sold From



The general perception by just over half the respondents surveyed was that cash jobs were "commonplace" in New Zealand (refer to Figure 5). Moreover, under half believed that "given the opportunity, most people would hide income to avoid paying taxes." In terms of social tolerance for such activity, however, 57% believed that it was not "ok" for people not to declare all their income. Regarding whether respondents perceived that cash job activity would be detected by Inland Revenue, the responses were evenly distributed across the response categories.

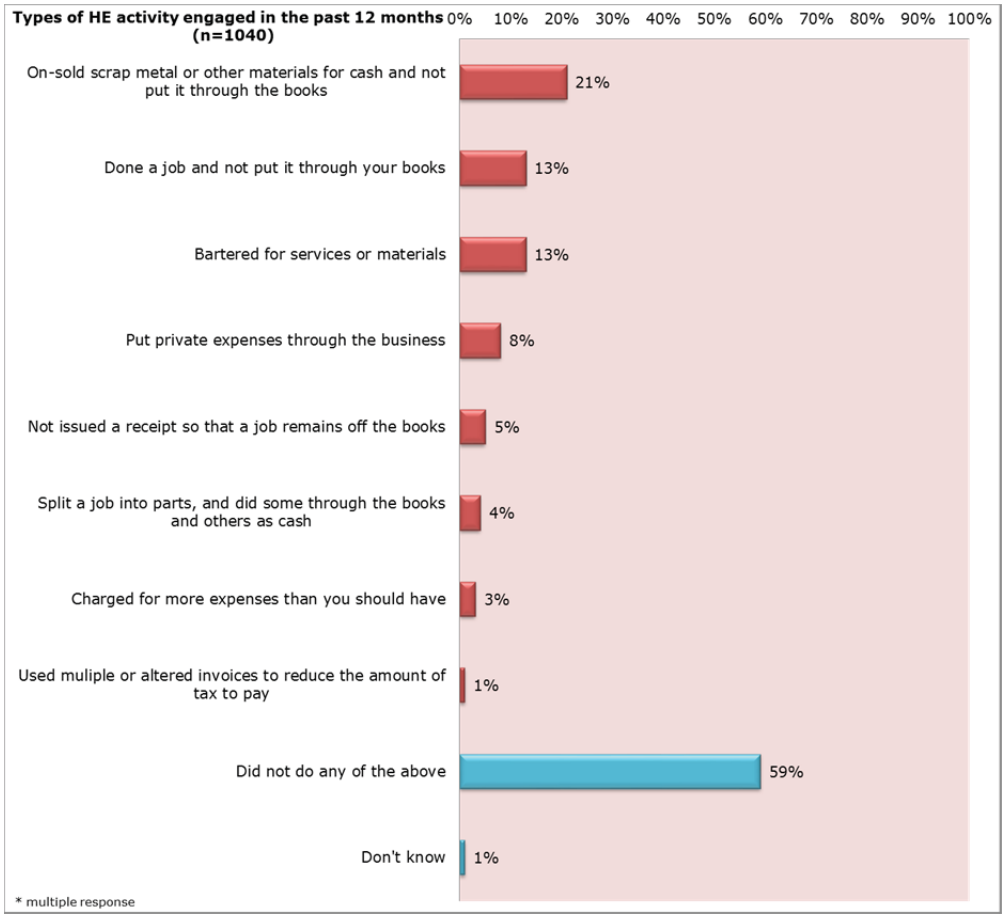
FIGURE 5. Personal and Social Norms Towards the Hidden Economy



In addition to research that focused primarily on engagement in cash job activity, whether the New Zealand public also engaged in other types of Hidden Economy activity was also investigated. Specifically, Small and Medium-sized Enterprises (SMEs) from the construction and trade industry were surveyed to better understand their compliance characteristics.

The research found that for construction and trade SMEs, the on-selling of scrap metal or other materials for cash and not declaring the income was the most common type of noncompliant activity reported (21%), followed by doing cash jobs (13%), and bartering for services or materials (13%) (refer to Figure 6). For SMEs with employees, a small proportion also reported treating employees as if they were "self-employed" (12%) or having employed someone under-the-table (9%) (refer to Figure 7).

FIGURE 6. Types of Noncompliant Activity Engaged by Respondents From the Construction and Trade Industry



2.3.3 Motivators and Barriers to Hidden Economy Participation

The literature suggests that in times of economic hardship, Hidden Economy activity is likely to increase. This was supported by the research, which found, for example, that economic and financial hardship was one of the predictors of Hidden Economy participation for SMEs from the construction and trade industry. These findings also corroborated results from a national survey on Hidden Economy attitudes and behaviours, which found that there was strong social acceptance for using the state of the economy as justification for cash job participation (refer to Figure 8).

FIGURE 7. Noncompliant Behaviour With Construction and Trade SMEs Who Had Employees

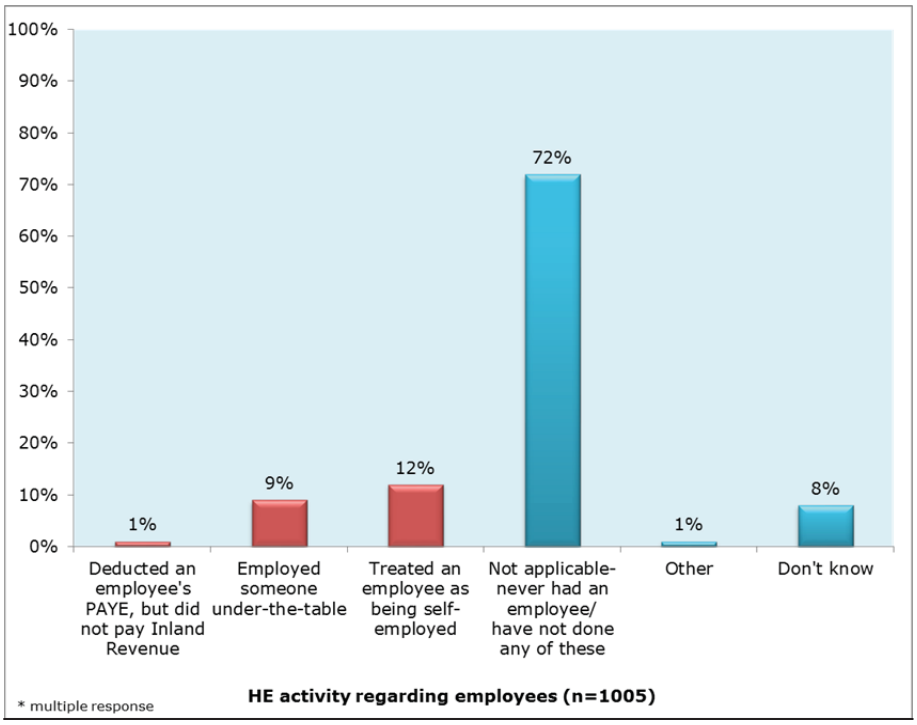
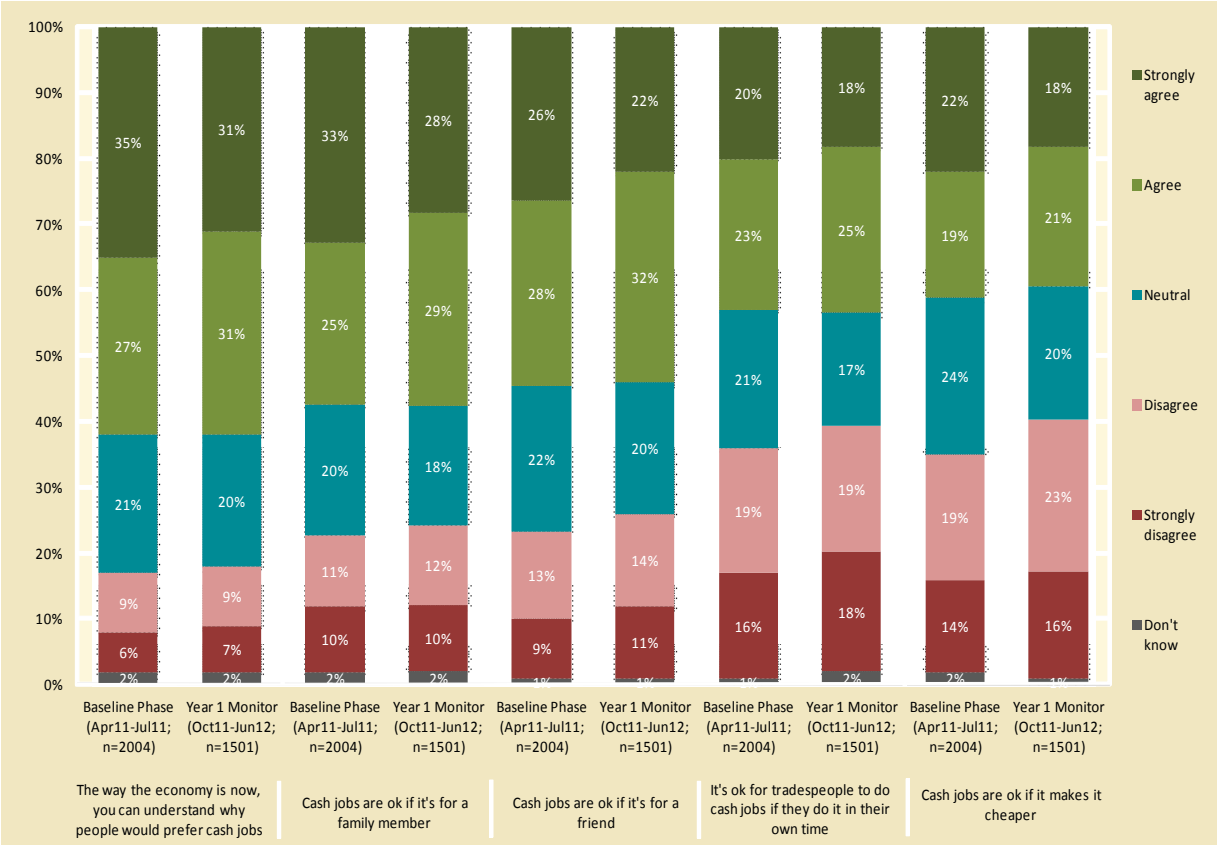


FIGURE 8. The Economy, Social Acceptability, and Price of Cash Jobs as Motivators for Cash Job Participation



What the findings from Figure 8 also illustrate is the importance and influence of interfamilial and interpersonal relationships in maintaining cash job activity. The role of these relationships was further explored in a qualitative case study on the drivers of Hidden Economy participation within a region of New Zealand (refer to Figure 9). As with the findings from the quantitative research, financial and interpersonal factors were found to be strong drivers of Hidden Economy participation within this qualitative research.

The following were found to be predictors of construction and trade SMEs not participating in the Hidden Economy:

- Respondents' moral conscience (tax morale);
- Their perceived risk of detection (deterrence);
- Belief that cash jobs were not commonplace; and
- Belief that others were honest with their tax obligations (social norms).

Examples of further explanations given by respondents for choosing not to engage in Hidden Economy activity are provided in Figure 10.

FIGURE 9. Mapping the Drivers of Hidden Economy Participation Within a Region of New Zealand

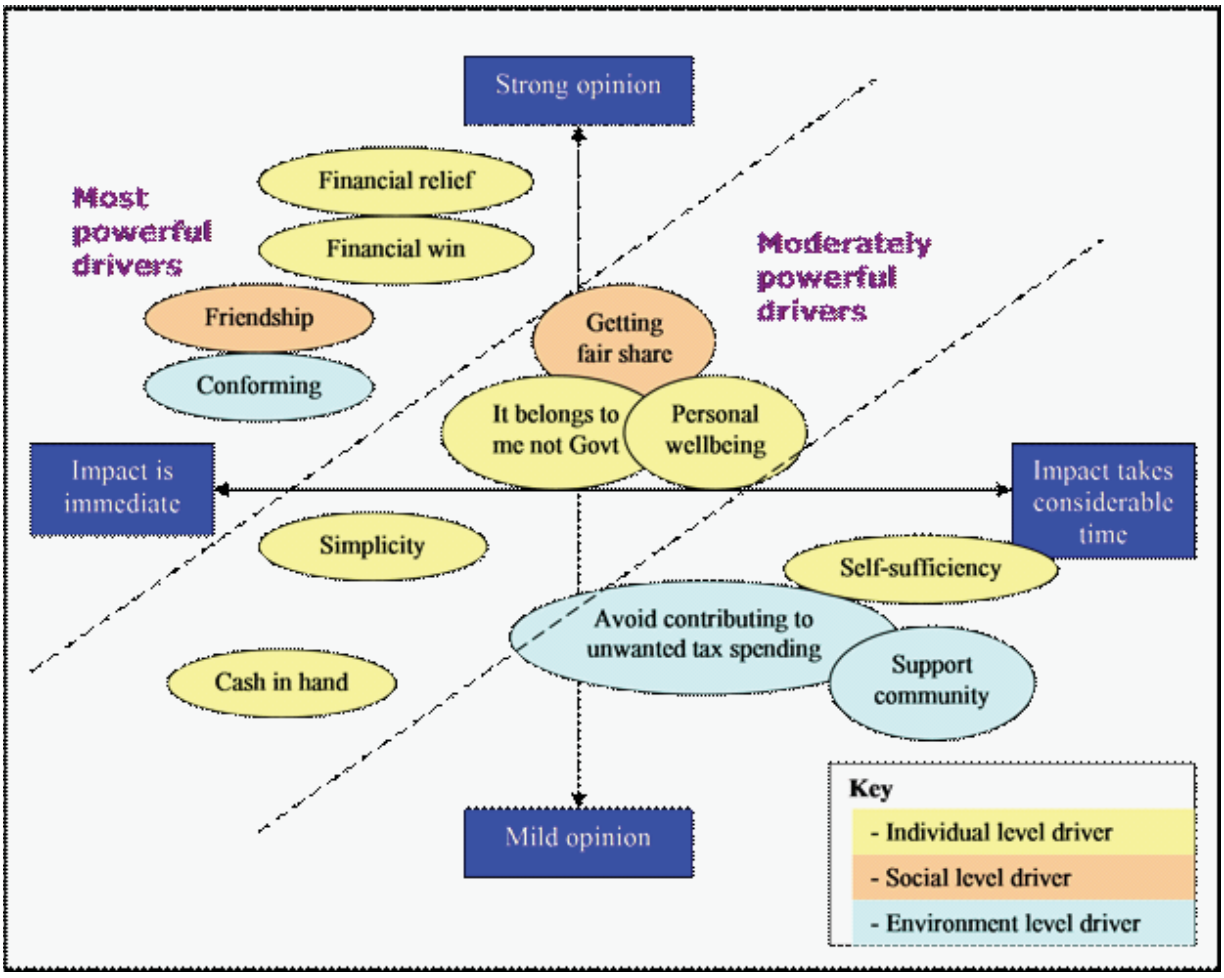
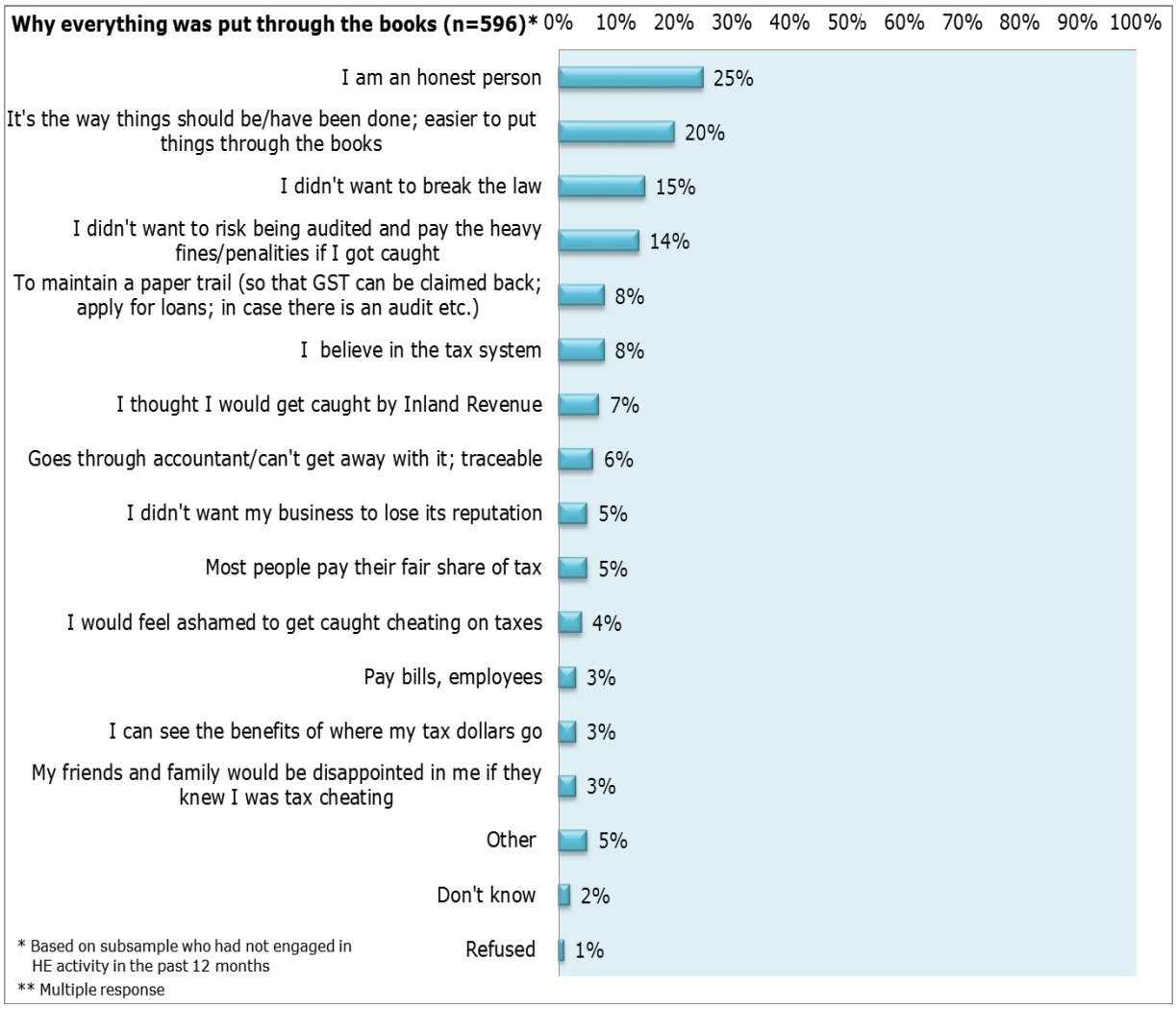


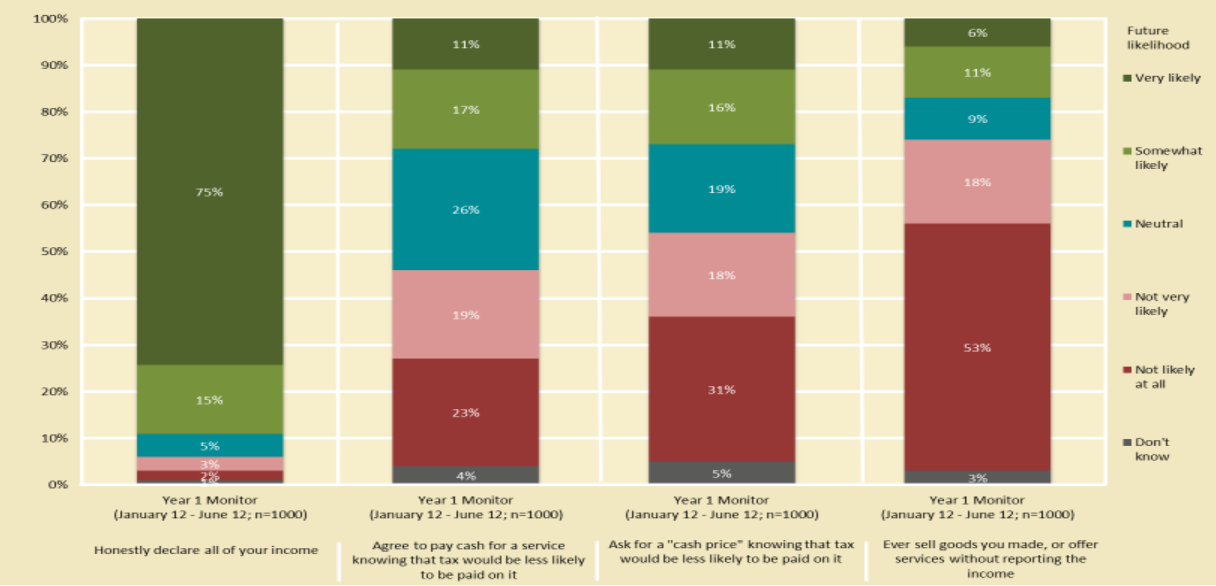
FIGURE 10. Reasons Given for Not Engaging in Hidden Economy Activity



2.3.4 Likelihood of Future Hidden Economy Behaviour

To gauge what future trends in Hidden Economy activity might hold, the general populace were asked whether they were likely to: honestly declare their income in the future; agree to pay/ask for cash jobs; and sell goods/services without declaring the income in the future. The vast majority of respondents (90%) reported that they would honestly declare all the income. A small portion also reported that they would agree to either pay for cash jobs (28%), ask for cash jobs (27%), or offer cash jobs (17%), even if tax was less likely to be paid in those circumstances (refer to Figure 11).

FIGURE 11. Future Likelihood of Engaging in Hidden Economy Activity



2.4 Lessons Learned and Implications for Inland Revenue New Zealand

2.4.1 End-to-End Campaigns/Programmes of Work

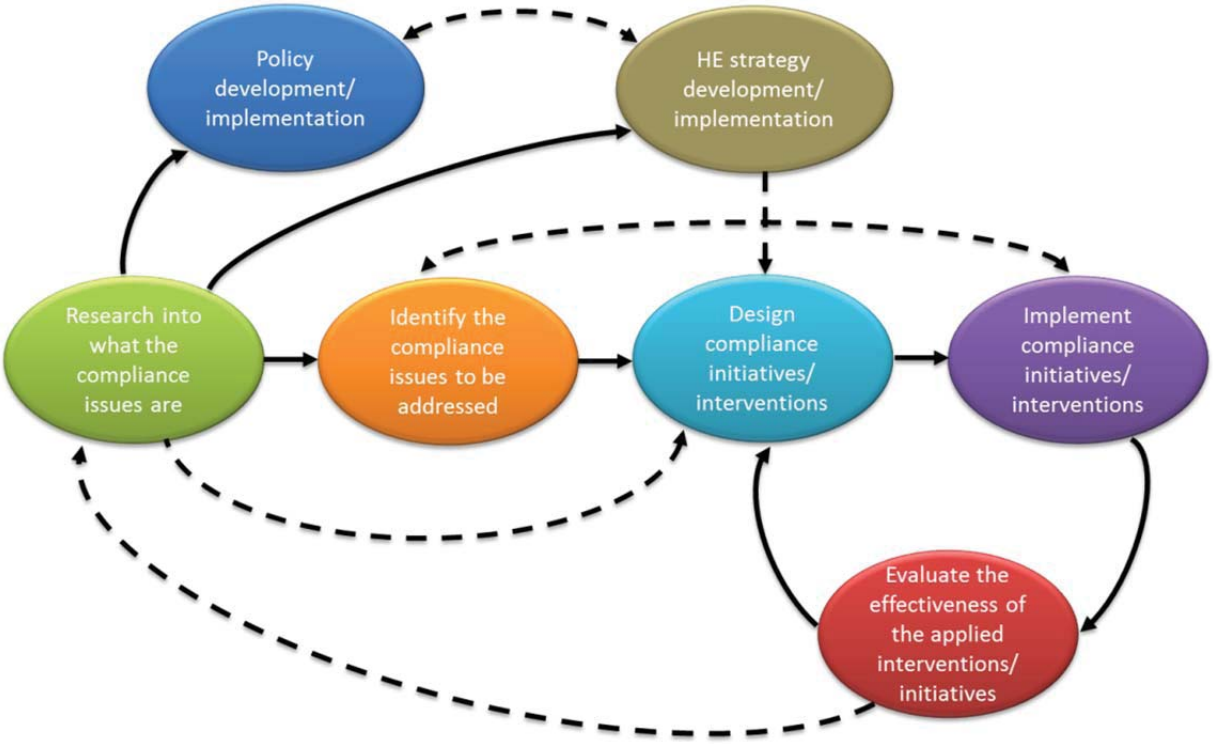
On consolidating the findings from the research and evaluation to date, they highlight some of the challenges associated with the Hidden Economy. It is apparent that holistic, "end-to-end" campaigns or programmes of work are needed to address these challenges. A simplified conceptual model of how such campaigns/programmes of work could be designed and implemented is provided below in Figure 12.

To implement such programmes of work, it is suggested that as an initial step research be conducted to develop insights into what the compliance issues are. Findings from the research and evaluation activity would then guide the development of the Hidden Economy strategy, policy development and implementation, and the design and implementation of Inland Revenue's compliance initiatives/interventions.

Subsequent to the implementation of Inland Revenue's compliance initiatives/interventions, evaluations to gauge the effectiveness of these activities are suggested. Findings from the evaluations would then feed back into the refining and redesigning of the applied interventions, to ensure that Inland Revenue remains responsive and effective in its service delivery, and that the implemented programmes of work are meeting their intended objectives.

When evaluating the effectiveness of interventions and/or programmes of work on the Hidden Economy, owing to the "hidden" nature of it, one challenge will be how to measure what is hidden, and to demonstrate what would have happened in the absence of the intervention. Consideration is currently being given into the development of "compliance indicators" that can provide measures of effectiveness of compliance activities.

FIGURE 12. Simplified Conceptual Model of How An End-to-End Campaign/Programme of Work on the Hidden Economy Could Be Implemented



2.4.2 Implications for Intervention Design and Implementation

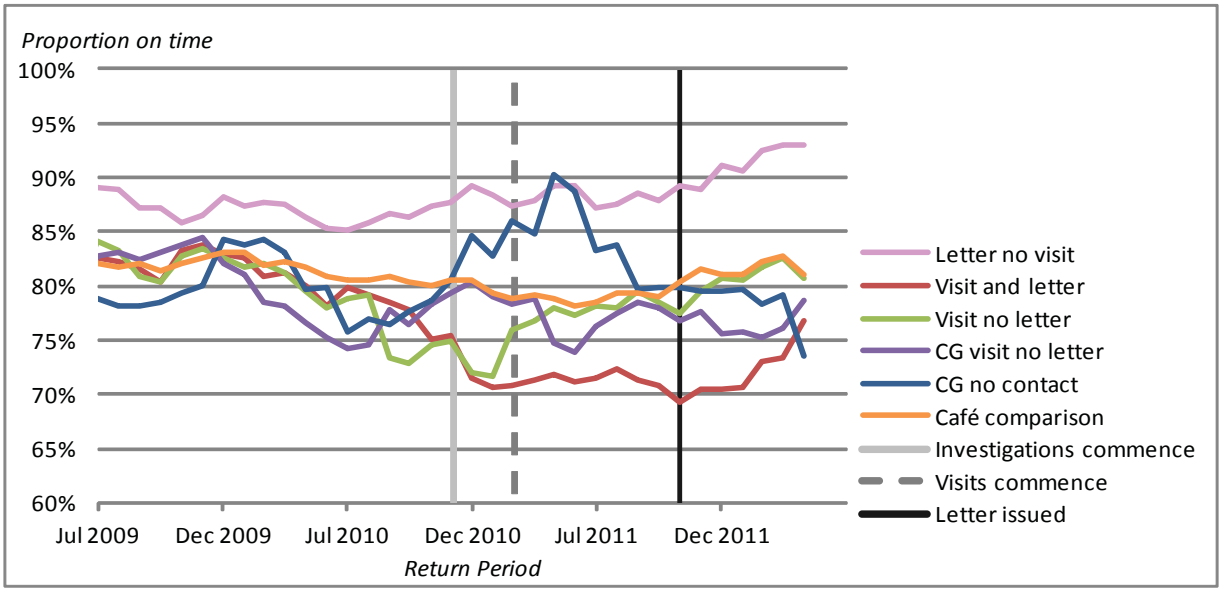
In terms of the specific design of the types of compliance initiatives or interventions that could be undertaken, findings from the Hidden Economy Research & Evaluation Programme suggest that multi-modal and multi-faceted intervention designs are needed. Therefore, in addition to the "traditional" suite of compliance activities, it is suggested that consideration also be given to:

- The application of national/regional social marketing campaigns (to help shape societal attitudes and behaviours towards the Hidden Economy and, for example, to influence social norms related to the "civic duty" of paying tax etc.);
- The increased use of social media (in response to the changing needs of customers and the growing digital environment); and
- More pronounced engagement by Inland Revenue to raise its external visibility (with stakeholders at all levels, from individual taxpayers to SMEs and large corporate organisations).

How these initiatives or interventions are implemented could also be given further consideration. As stated earlier, it may be worth exploring how randomised control trials can be built into the testing and application of these compliance initiatives or interventions.

Recent evaluation of Inland Revenue's compliance initiatives in the restaurant sector provides an example of where a combination of different interventions was trialled (refer to Figure 13). Although the findings were inconclusive, owing to the short-time frames involved in the evaluation, Figure 13 provides a graphical illustration of how future intervention design could benefit from more structured approaches to intervention design and implementation.

FIGURE 13. The Effect of Sending Letters and/or Staff Visits on the Timely Payment of GST to Businesses in the Restaurant Sector



A long-term approach to the design and implementation of initiatives and interventions, to enable continual monitoring and assessment is also suggested. Within the Hidden Economy Research & Evaluation Programme, there is currently long-term monitoring of national attitudes and behaviours towards the Hidden Economy, and a longitudinal study to explore the impact of adverse events on SME tax compliance. The benefit of interweaving longitudinal research and evaluation projects into the intervention design process is that it enables Inland Revenue to be responsive to customer needs and any compliance issues that may arise.

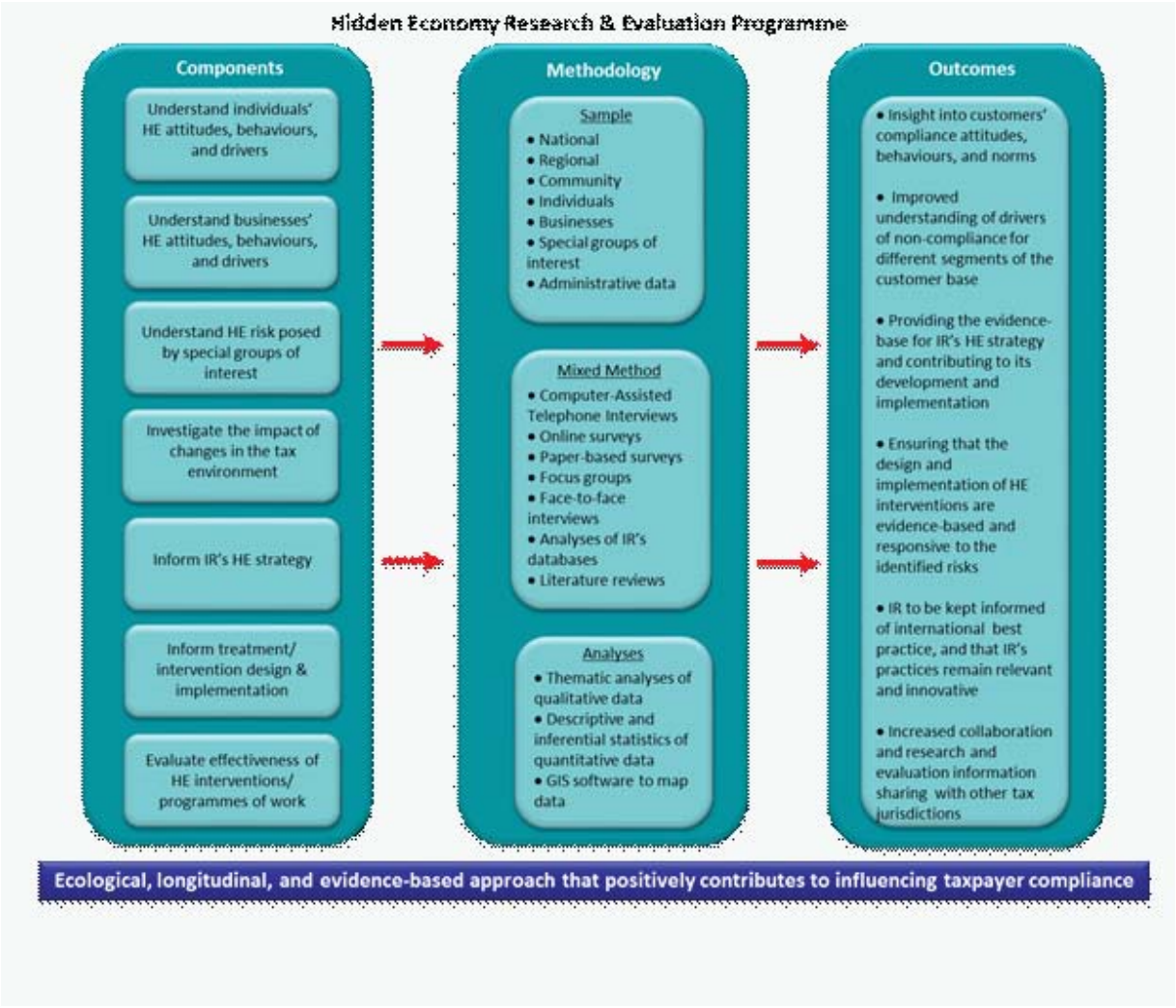
2.4.3 Long-Term Commitment

It is worth noting that the compliance issues faced in New Zealand are not unique and that Inland Revenue New Zealand shares similar compliance challenges faced by other tax administrations. Owing to the complexity of the issues faced, it may be worth contemplating how closer relationships could be fostered between New Zealand and its international colleagues to address some of the compliance challenges ahead. A long-term commitment and approach to addressing the Hidden Economy is also worth considering.

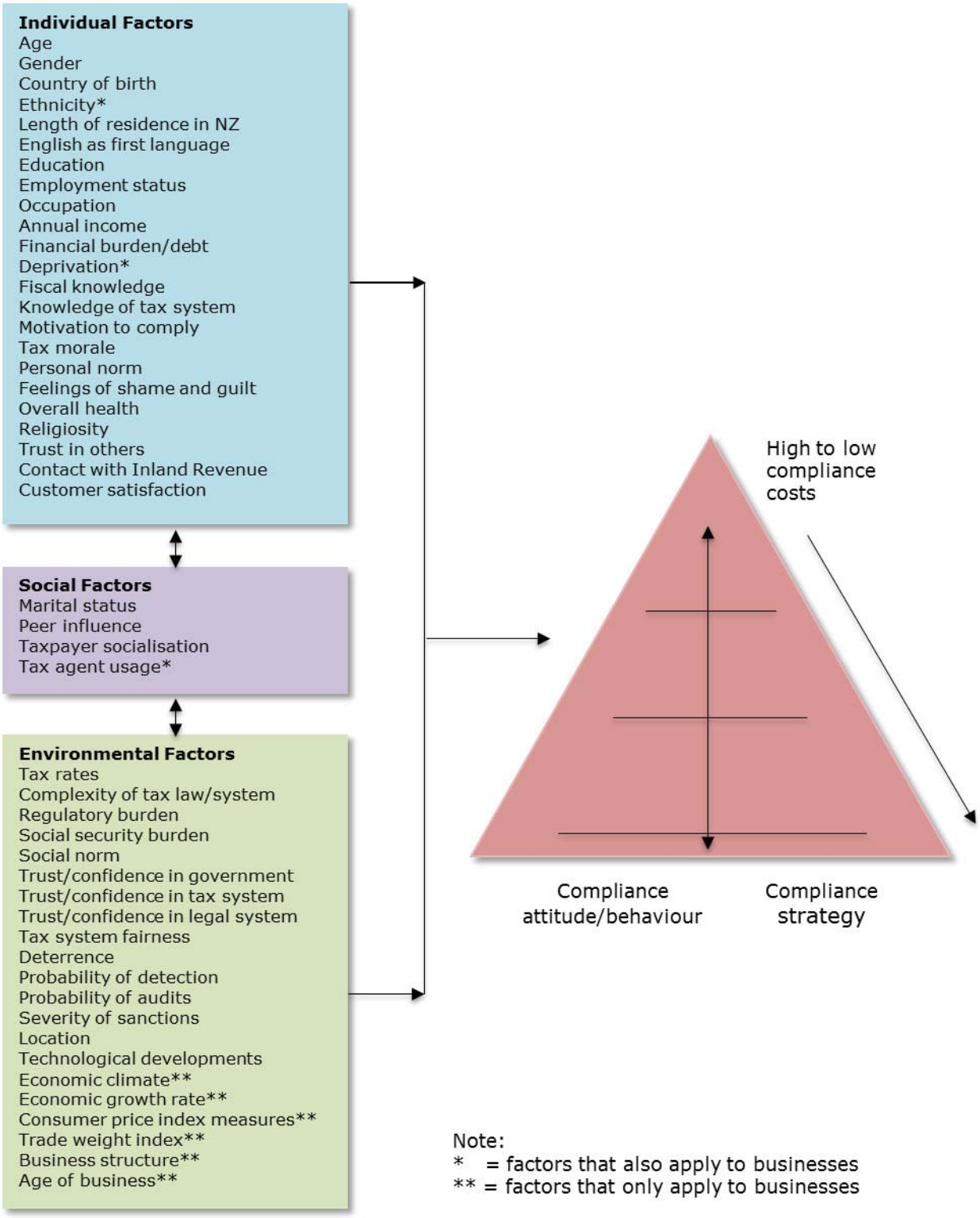
3.0 Conclusion

The Hidden Economy undermines the integrity of the tax system. This paper has outlined Inland Revenue New Zealand's programme of research and evaluation into the Hidden Economy. While still in its infancy, the Hidden Economy Research & Evaluation Programme has been able to provide insights into customer compliance attitudes and behaviours. It is hoped that this research and evaluation activity will go some way towards shedding light into New Zealand's "black hole."

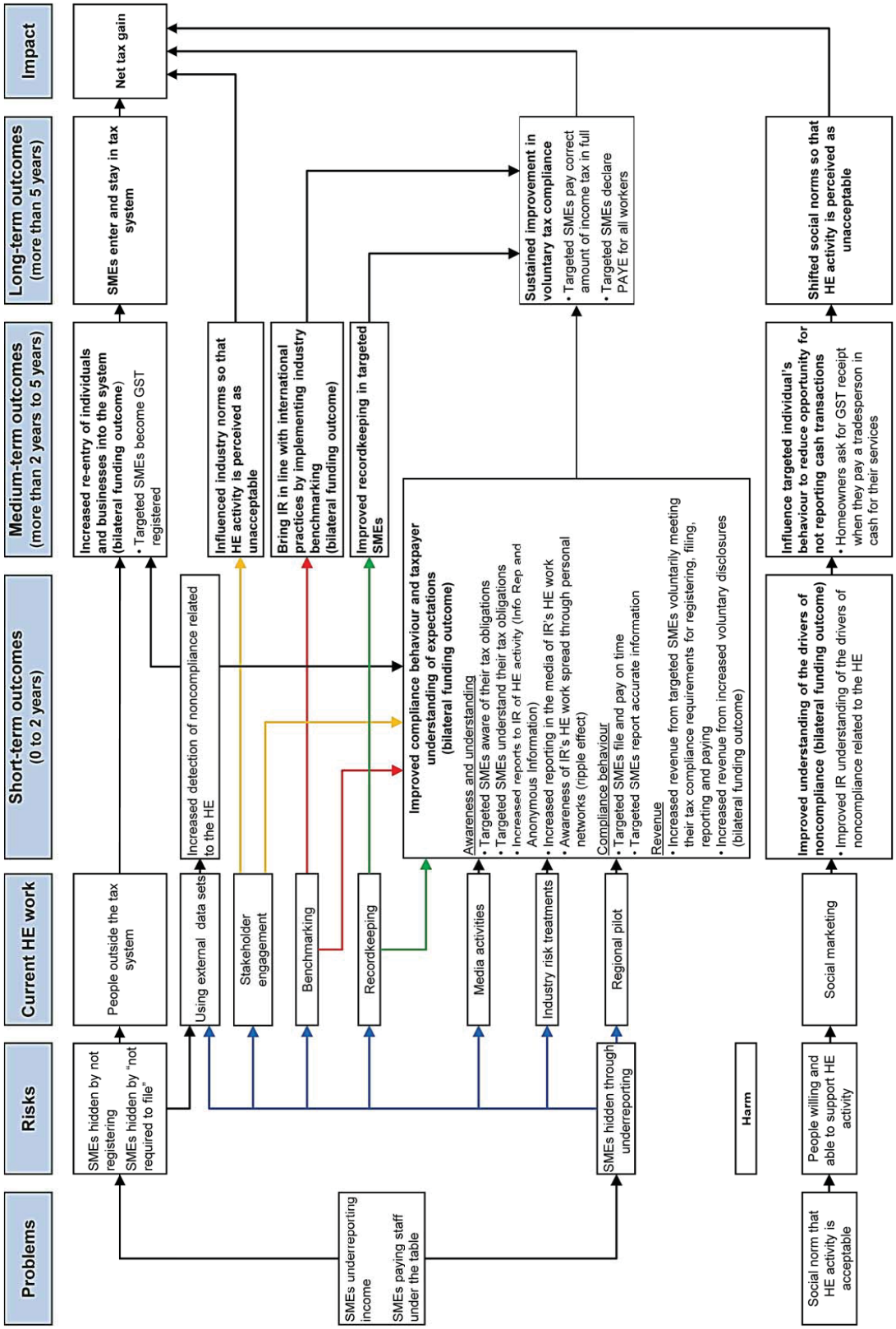
APPENDIX A
Overview of the Hidden Economy Research & Evaluation Programme



APPENDIX B
Application of An Ecological Framework to Understanding Factors Influencing Compliance Attitudes and Behaviours



APPENDIX C
Intervention Logic That Guided the Evaluation Strategy for Hidden Economy Work



References

Australian Public Service Commission (2007). Changing behaviour: A public policy perspective. Canberra: Australian Government.

Braithwaite, V, ed., (2003), Taxing Democracy: Understanding Tax Avoidance and Evasion, Ashgate Publishing Ltd, Aldershot, UK. Central Office of Information (2009). Communications and behaviour change. London: Crown.

Chand, A., & Cleland, A. (2012). Working with ethnic and migrant Small and Medium Enterprises (SMEs). Paper presented at Pathways to Metropolis in the 21st Century: Immigration issues and futures, Auckland.

European Commission (2007). Undeclared work in the European Union. EU: Author.

Halpern, D., Bates, C., Mulgan, G., Aldridge, S., Beales, G., & Heathfield, A. (2004). Personal responsibility and changing behaviour: The state of knowledge and its implications for public policy. London: Public Office.

Haynes, L., Service, O., Goldacre, B., Torgerson, D. (2012). Test, learn, adapt: Developing public policy with randomised controlled trials. London: Behavioural Insights Team, Cabinet Office.

Poppelwell, E., Sun, H., & Colmar Brunton (2012). Exploring the impact of adverse events on SME tax compliance behaviour: A longitudinal study. Findings from Year 1. Wellington: National Research & Evaluation Unit, Inland Revenue New Zealand.

Poppelwell, E., Bennett, A., Bickers, P., & Colmar Brunton (2013). Exploring the impact of adverse events on SME tax compliance behaviour: A longitudinal study on Inland Revenue's response to, and learnings from, the Christchurch earthquakes. Findings from Year 2 – Qualitative research with stakeholders and Inland Revenue staff. Wellington: National Research & Evaluation Unit, Inland Revenue New Zealand.

Statistics New Zealand (2013). 2013 Census QuickStats about national highlights. Wellington: Author.

Yuan, S., Cain, T., & Spoonley, P. (2013). Immigrant entrepreneurship and tax compliance. Wellington: National Research & Evaluation Unit, Inland Revenue New Zealand.

5

Appendix

Conference Program

Tax Administration at the Centennial: An IRS-TPC Research Conference

Urban Institute, 2100 M Street, N.W., Washington, DC • June 20, 2013

Program

8:30 – 9:00 Check-in

9:00 – 9:10 Welcome and Introductions
Eric Toder *(Co-Director, Tax Policy Center) and*
Pat McGuire *(Deputy Director, IRS Office of Research, Analysis, and Statistics)*

9:15 – 10:45 **Session 1: Individual Income Tax Dynamics**

Moderator: ***Eric Toder*** *(Tax Policy Center)*

- Older Taxpayers' Responses to Taxation of Social Security Benefits
  Len Burman *and Liu Tian (Syracuse University), Norma Coe (University of Washington), and Kevin Pierce (IRS:RAS)*
- Preparer Industry Dynamics and the Return Preparer Initiative
  Pat Langetieg, Mark Payne and ***Melissa Vigil*** *(IRS: RAS)*
- Changes in EITC Eligibility and Participation, 2005–2009
  Maggie R. Jones *(Census Bureau)*

Discussant: ***Dayanand Manoli*** *(University of Texas)*

10:45 – 11:00 Break

11:00 – 12:30 **Session 2: Business Compliance Behavior**

Moderator: ***Eric Toder*** *(Tax Policy Center)*

- Factors Influencing Voluntary Compliance by Small Businesses: Preliminary Survey Results
  Tom Beers, Eric LoPresti, and ***Eric San Juan*** *(IRS:Taxpayer Advocate Service)*
- Transfer Pricing: Strategies, Practices, and Tax Minimization
  Kenneth Klassen (University of Waterloo, Canada), ***Petro Lisowsky*** *(University of Illinois), and Devan Mescall (University of Saskatchawan, Canada)*
- Demand for Aggressive Tax Planning
  Peter Bickers*, Michael Slyuzberg, Tracey Lloyd, and Bhaskaran Nair (Inland Revenue, New Zealand)*

Discussant: ***Amy Dunbar*** *(University of Connecticut)*

1:30 – 3:00 **Session 3: Corporation Income Tax Enforcement**

Moderator: ***Javier Framiñan*** *(IRS:W&I)*

- Analysis of Ten-Year Trends in Large Business Examination Results (2001–2011)
  Dave Macias *and Kimmy Wang (IRS: LB&I)*
- The Impact of Legal Enforcement: An Analysis of Corporate Tax Aggressiveness After an Audit
  Jason DeBacker (Middle Tennessee State University), Bradley Heim and Anh Tran (Indiana University), and ***Alexander Yuskavage*** *(U.S. Treasury: Office of Tax Analysis)*
- IRS Enforcement and State Corporation Income Tax Revenues
  Margot Howard *(University of North Carolina)*

Discussants: ***Jonathan Feinstein*** *(Yale University) and* ***Brian Erard*** *(B. Erard & Associates)*

3:00 – 3:15 Break

3:15 – 4:45 **Session 4: Lessons From Other Tax Administrations**

Moderator: ***Rahul Tikekar*** *(IRS:RAS)*

- Why Evasion Under a National Sales Tax Would Explode the Tax Gap: Lessons Learned from the States
  Cary Christian *(Georgia Southern University)*
- The Influence of Tax Form Design on Use Tax Compliance
  David Merriman *(University of Illinois), and Natalie Davila and Hector M. Vielma (Illinois Department of Revenue)*
- Filling in the Black Hole: Research and Evaluation into the Hidden Economy
  Alice Cleland *(Inland Revenue, New Zealand)*

Discussant: ***Alan Plumley*** *(IRS:RAS)*

4:45 – 5:00 **Wrap-up**

Janice Hedemann *(Conference Chair, IRS:RAS)*