Recent Research on Tax Administration and Compliance

Selected Papers Given at the 2011 IRS-TPC Research Conference: New Perspectives on Tax Administration

Held at the Urban Institute
Washington, DC
June 22, 2011

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

*Prepared under the direction of Barry W. Johnson, Statistics of Income, Chief, Special Studies Branch

Foreword

This edition of the IRS Research Bulletin (Publication 1500) features selected papers from the IRS-Tax Policy Center (TPC) Research Conference "New Perspectives on Tax Administration," held at the Urban Institute in Washington, DC, on June 22, 2011. As in prior years, 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. However, this year's conference took on an exciting new format. For the first time, the IRS partnered with a nonpartisan tax research organization to host the annual research conference. More people could participate because of the TPC broadcast video of the proceedings live over the Internet and the videos were archived on their Web site. 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 of the Tax Policy Center and Rosemary Marcuss, the IRS Director of Research, Analysis, and Statistics. They were followed by comments from IRS Commissioner Doug Shulman. Mr. Shulman expressed his appreciation for the role of good research in effective tax administration and thanked the TPC for helping to host this conference. He concluded his remarks by answering a few questions from the audience.

The remainder of the conference included sessions on the impact of service on compliance, individual compliance behavior, estimating the tax gap, and new disclosure and regulation issues. The lunchtime keynote address was by David Walker, founder of the Comeback America Initiative. He outlined the seriousness of the nation's fiscal challenges and offered several principles to guide potential solutions.

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 Federal tax administration. We also hope that the research featured here will stimulate improved tax administration and additional helpful research.

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, Ted Black, John Guyton, Rahul Tikekar, Cheryl Wagner, and Leann Weyl (National Headquarters Office of Research); Melissa Laine (Office of Program Evaluation and Risk Analysis); Charles Day, Kim Henry, and Barry Johnson (Statistics of Income); Alex Turk (Small Business and Self- Employed); Dawn Fitzelle (Tax Exempt and Government Entities); Kate Burke (Wage and Investment); Carol Hatch (Taxpayer Advocate); Davy Leighton (Criminal Investigation); and Charles Boynton (Large and Midsize Business). In addition, Blake Greene from the Tax Policy Center oversaw numerous details to ensure that the conference ran smoothly.

This volume was prepared by Lisa Smith, Paul Bastuscheck, and Camille Swick (layout and graphics) and Martha Eller Gangi (editor), 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, 2011 IRS-TPC Research Conference

New Perspectives on Tax Administration: An IRS-TPC Research Conference

Contents

4. New Disclosure and Regulation Issues

5. Appendix

1

Impact of Service on Compliance

Vossler ◆ McKee ◆ Jones
Bruch ◆ Cico ◆ Mehmood
Plumley ◆ Erard ◆ Snaidauf

Some Effects of Tax Information Services Reliability and Availability on Tax Reporting Behavior[1]

Christian Vossler, University of Tennessee; Michael McKee, Appalachian State University; and Michael Jones, Bridgewater State University

To "encourage" correct tax reporting it is likely that enforcement effort, audits and penalties, will continue to be a primary tool in the tax authority's arsenal. However, many tax agencies are exploring complementary instruments of which one is the provision of information and assistance services to taxpayers. This revised paradigm recognizes that tax administrators have a role as facilitators and a provider of services to taxpayer-citizens. Almost everyone agrees that the personal income tax system in the U.S. is complex. It is counterproductive to make information resolving the complexity costly to obtain and/or unreliable.[2] Further, the "service" paradigm for tax administration fits squarely with the perspective that emphasizes the role that social norms play in tax compliance (Feld and Frey, 2002).

The effect of such service programs on tax reporting is an open empirical question but intuitively, more reliable and available service programs will be more likely to positively affect tax reporting behavior. Testing such a proposition requires an analysis of individual-level data under alternative information service settings. While some changes in the service programs have been undertaken in the past, there is not a full spectrum of such programs in existence and so field data are incomplete. Just as important, even for taxpayers that undergo a full audit, the determination of tax liability is vulnerable to audit error.

As an alternative to analyzing field data, our research utilizes controlled laboratory experiments with human decision makers and salient financial incentives. Within the laboratory, we determine (hence, know) the true tax liability, and then identify the effects of information services by systematically varying the setting across groups of players. In particular, our experimental design varies the degree of accessibility and accuracy of information services. Our design allows us to observe both the tax reporting behavior as well as the propensity of the taxpayer to obtain information by making information acquisition a (sometimes costly) choice.

Experimental Design and Treatments

Our experimental design includes many of the fundamental elements of the voluntary reporting system of the U.S. individual income tax. In each decision period of the experiment, participants earn income by performing a task and self-report their tax liability to a tax authority. Tax liability is (possibly) uncertain, and is a function of earned income, the tax rate, and tax credits applied. If an audit occurs, unreported taxes are discovered without error. If the participant has evaded taxes both the unpaid taxes and a penalty are collected. The participant's earnings for the decision period, which are denominated in "lab dollars", are her earned income less taxes paid (and penalties, if applicable). The overall earnings for the experiment are the sum of the lab dollars earned over all decision periods multiplied by a common (and known) lab to US dollar exchange rate. We discuss further the details of the experiment below.

In each period of the experiment, participants earn income based upon their performance in a simple computerized task, in which they are required to sort numbers into the correct order. Those who finish the task the fastest earn the highest income of 1500 lab dollars for the period, those who finish in the middle of pack earn 1250 lab dollars, and the slowest earn 1000 lab dollars. Participants are presented information about the distribution of group earnings to ensure that they believe the relative nature of the earnings. The earnings task is the only source of interaction and payoff interdependence; the design implements a blind setting among the participants.

After earning income, participants are presented with a screen that informs them of the earnings information and the tax policy parameters (tax rate, audit probability, and penalty rate) which are fixed throughout the experiment.[3] The decisions the participants make in the period are whether to request an information service (if one exists) and how much to claim in tax credits. Although other institutional details are embedded in the design (e.g. tax rate, taxable income, etc.), and in particular the tax form, the participant can only manipulate her tax liability through her credit reporting choice. As there are penalties for tax under reporting if audited, and foregone earnings associated with over-paying taxes, there is value to resolving any uncertainty regarding the tax credit. The expected tax credit starts at 1000 for an income of zero and declines at a rate of 0.5 for each additional dollar of income. The amount of the credit is high relative to the initial tax liability so that the credit decision is financially salient.

We implement uncertainty in the credit by placing uniform distributions around the expected credit amounts, and then randomly drawing from the distributions to determine the true credit amounts. In particular, the true credit amount can be anything in the range of the expected credit plus or minus 100%. With uncertainty, prior to making a credit choice or acquiring information (if possible), each participant sees the supports of the uniform distribution that coincides with her income. If an information service is available, participants can acquire the information with the click of a button.[4]

The participants are informed of the audit probability and the penalty rate, and know these values with certainty. In all sessions we fix the tax rate at 50% of earned income, the audit probability at 30%, and the penalty rate at 300% of unpaid taxes. Our audit rate is much higher than actual full audit rates in the United States. However, the IRS conducts a range of audits, and for many types of audits the actual rates are quite high.[5] The penalty rate is consistent with penalties imposed by the IRS for tax underreporting. Enforcement effort is held constant since the effects of enforcement efforts have been widely investigated and we only need this effort to be salient in the current setting to give value to the information that resolves tax liability uncertainty.[6] Table 1 reports the key parameters of the experiment.

Participants are able to revise their credit decision prior to filing their return, and the tax form updates their tax liability as the claimed credit is revised. Thus, they can observe the potential changes in their reported tax liability for each potential reporting strategy they investigate. A timer at the bottom of the tax form counts down the remaining time. The participants are allowed 90 seconds to file and the counter begins to flash when there are fifteen seconds remaining. Thus, the process in the lab mimics that by which a taxpayer may well conduct different calculations in the time prior to actually filing her taxes (whether he or she uses one of the available tax software programs or simply does the tax return by hand). If an information service is available, this can be requested at any time.

Audits are determined by the use of a "virtual" bingo cage that appears on the computer screen. A box with blue and white bingo balls appears on the screen following the tax filing. The ratio of blue to white balls determines the audit probability. The balls begin to bounce around in the box, and after a brief interval a door opens at the top of the box. If a blue ball exits, the participant is audited; a white ball signifies no audit. The audit applies only to the current period declarations, not to previous (or future) periods. The computer automatically deducts taxes paid and penalties (if any are owed) from participants' accounts.

When an audit occurs, the true value of the credit is used to determine taxes owed. The individual's declarations are examined. If the individual has under-reported her tax liability, she must make up for the difference as well as pay a penalty. If an individual has over-reported their tax liability no over payments are returned to the individual.[7] Tax revenues and any penalties paid are not distributed to the participants; tax collections are not used to provide a public good in order to ensure that the participants focus on the individual income disclosure decision and not on any public good provision decision. After the tax return is filed and an audit (if any) is determined, participants see one final screen that summarizes everything that happened during the period. This process is repeated for a total of 20 paid rounds, but to minimize potential end-of-game effects the number of rounds is not disclosed.

Participant Pool and Detailed Procedures

The experiments were conducted at dedicated experimental laboratories at the University of Tennessee and Appalachian State University, which both utilized the same software and experimental protocol, and have similar computer networks. The participant pools included students and non-students (university staff, mostly).[8] Student and non-students participated at separate times, and the lone difference in student versus non-student sessions is that the latter utilized a lower lab dollar to US dollar exchange rate (375 to 1 versus 750 to 1) in order to reflect the higher opportunity cost of participation. Recruiting was conducted using the Online Recruiting System for Experimental Economics (ORSEE) developed by Greiner (2004). Databases of potential participants were built using announcements sent via email to university students and staff. Registered individuals were contacted, via email, and were permitted to participate in only one tax experiment.[9] Only participants recruited specifically for a session were allowed to participate, and no participant had prior experience in this experimental setting. Methods adhere to all guidelines concerning the ethical treatment of human participants. Earnings averaged $25 for student participants and $45 for non-students. Sessions lasted between 60 and 90 minutes. A total of 730 participants took part in these sessions.

The experiment session proceeds in the following fashion. Each participant sits at a computer located in a cubicle, and is not allowed to communicate with other participants. The instructions are conveyed by a series of computer screens that the participants read at their own pace, with a printed summary sheet provided and read aloud by the experimenter (see Appendix for an example). Clarification questions are addressed after the participants have completed the instructions and two practice rounds. The participants are informed that all decisions will be private; the experimenter is unable to observe the decisions, and the experimenter does not move about the room once the session starts to emphasize the fact that the experimenter is not observing the participants' compliance decisions. This reduces, to the extent possible, peer and experimenter effects that could affect the decisions of the participants. All actions that participants take are made on their computer. After the 20 paid decision periods, participants are asked to fill out a brief questionnaire which collects basic demographics including information on tax reporting experience. Payments are made privately at the end of the session.

Treatments

We employ a between-subjects design, where the treatment variables across sessions are the presence/absence of an information service, the quality of the service if provided, and the cost of obtaining the information. These are held constant throughout a session. There are five basic treatments (see Table 2). The first (T1) is a treatment with certain tax liability, which we use as a baseline for comparison against uncertain information treatment. In this treatment, participants are automatically given information on their true credit. In the second treatment (T2), the individual's tax credit is uncertain and there is no information service available. This establishes a second baseline for comparison. In the remaining three treatments, there is an information service available. The status quo in the information service treatments, i.e. if the information service is not utilized, is identical to the uncertainty baseline.

The "perfect" information service reveals the true credit with certainty (T3). Under the other two information service types, the service is imperfect in the sense that up to two *possible* credit amounts can be provided and each amount has a 50% chance of being correct. Specifically, under the "simultaneous" information service treatment (T4) the authority simultaneously provides two credit amounts, one of which is the truth while the other is a decoy. With the "sequential" information service (T5), the participant can make up to two information requests and with each request is delivered one possible credit amount. If two requests are made, then the simultaneous and sequential services reveal the same information. However, the sequential information treatment leaves the possibility that only one credit amount is delivered, in which case it still has the same 50% chance of being the truth.

To assess the value of information services, we vary (between sessions) the cost to acquire information in the information service treatments (see Table 1). The three cost levels are $0, $50 and $100 for the perfect and the simultaneous information settings. For the sequential setting, these costs are halved and assessed separately for the two sources.

Testable Hypotheses

To derive testable implications based on economic theory, we draw heavily from the theoretical model derived in Vossler, McKee and Jones (2010). In particular, the experimental game described above represents a special case of the theory, one in which the taxpayer makes a decision on a single "line item" and the taxpayer is required to file a return. With our experimental parameters, assuming risk neutrality, when the true credit is certain the taxpayer will optimally choose to report truthfully. When the credit is uncertain, for all levels of earned income, the taxpayer will over-claim the credit. The extent of the deviation from truthful reporting increases with the level of uncertainty. As those with the lowest income have the widest range of possible credits, theory suggests the highest relative amount of over-claiming for these individuals. Point predictions from the theory have that it is optimal to over-report the tax credit by 333 lab dollars for those with earned income of 1000, over-report by 250 for those with an income of 1250, and over-report by 167 for those with an income of 1500.

Tying in the imperfect information service, when the uncertainty is reduced to two possible credit outcomes, the optimal decision is to report one of the two possible amounts. With our chosen parameters it is optimal to choose the *higher* of the two amounts. Intuitively, it is not optimal to choose something in the middle of the two amounts as, in expectation, you forego an allowable credit and pay a larger penalty. Theoretically, although the information is nevertheless valuable to the player, tax underreporting is actually *higher* with two information sources than in the case of no information (i.e. the base uncertainty situation). This is a general theoretical result that does not depend on our choice of parameters. Intuitively this is driven by the fact that under full uncertainty one's decision is driven by the expected value of the underlying credit distribution – which is the true credit—, whereas the higher (lower) of the two draws is away from the truth on average. But, to be clear here, the information is valuable to the player.

The decision of whether to request the information service(s) to resolve (some) uncertainty is driven by the value of information. Theoretically, and quite intuitively, the taxpayer's willingness-to-pay (WTP) is increasing in the level of uncertainty as well as the accuracy of the information. In the context of the experimental design, those with lower incomes face a larger range of uncertainty and, *ceteris paribus*, have a higher WTP for information.[10] Further, knowing the true credit is more valuable than receiving two possible amounts only one of which is correct.[11] In terms of point predictions, since information has value, in all situations information should be requested when it is free. At the other extreme, in all situations no information should be requested at our highest cost amount of 100 (or 50 for one imperfect information source). At the middle cost amount, those at the lowest income level should request the information (imperfect or perfect), at the middle income level it is beneficial to request perfect information, and it is not beneficial for those with high income to request information.

The main testable implications of the theory are summarized below as formal hypotheses:

Hypothesis 1. The level of tax underreporting is higher when tax liability is uncertain.

Hypothesis 2. Tax underreporting increases with the level of uncertainty (i.e. decreases with income)

Hypothesis 3. Tax underreporting decreases when information services are provided.

Hypothesis 4. Tax underreporting decreases when information service quality improves.

Results of the Data Analysis

In the analysis that follows, we largely let the data "speak" by specifying OLS regression models that simply allow the mean outcome to differ across unique experiment scenarios. As such, the purpose of the regression is largely to estimate appropriate standard errors for the means and to facilitate hypothesis testing. On this note, to control for possible heteroskedasticity and autocorrelation of unknown form, we use robust standard errors with clustering at the participant-level. Further, heteroskedasticity and autocorrelation robust t and F statistics are used when evaluating hypotheses. To more parsimoniously illustrate some key patterns in the data, we also estimate models that implement some additional structure.

Tables 3–5 present three models using the credit decision data.[12] In all models we use as the dependent variable the difference between the credits claimed on the tax form and the expected amount of the credit. Formulating the dependent variable in this way allows the model parameters to be interpreted as the average amount of tax underreporting. The expected credit depends upon treatment conditions and is measured from the perspective of the participant.[13] That is, in our baseline certainty treatment or when perfect information is obtained in Treatment 3, the expected credit is simply the true credit. In uncertainty treatments where no information is acquired, the expected credit is simply the midpoint of the uncertainty interval. When two sources of information are obtained in Treatments 4 and 5, the expected credit is the average of the two. Finally, with one (imperfect) information source, the expectation is simply the average between the information draw and the midpoint of the uncertainty interval.

Model I estimates the average level of tax underreporting separately by each income level and each treatment. Model II extends the analysis to allow the average level of underreporting to be based on whether an information service was acquired and, if so, the type of service. To accomplish this, we define four new experiment "conditions". The first includes observations from information service treatments where information was not acquired ("No Information"). The remaining three correspond to observations where information was acquired: "Perfect Information" is associated with Treatment 3; "Two Information Sources" is associated with Treatment 4, and those in Treatment 5 who sequentially requested information from both sources; and "One Information Source" is associated with Treatment 5 for those who obtained one of the two available services. Finally, the main effects of income and information cost (where relevant), are estimated by experiment condition in Model III.

One prominent effect, as evidenced by all models, is that the tax underreporting is increasing in income (i.e. decreasing with the degree of uncertainty). This is in the opposite direction predicted by theory (Hypothesis 2), but consistent with our earlier findings in a related experiment (Vossler, McKee and Jones, 2010). A second, basic implication of the theory is that the presence of uncertainty increases tax underreporting (Hypothesis 1). Comparing our certainty and uncertainty baselines, there is weak evidence of this effect. Underreporting is roughly 60 to 70 lab dollars higher for all income levels, but this difference is only marginally significant at the middle income level (p-value=0.08; based on Model I or II). Inconsistent with Hypothesis 3, levels of under reporting do not differ based on the quality of the information service. In particular, there are no statistical differences across the three information service conditions, either by income level (Model II: p=0.83 @1000; p=0.55 @1250; p=0.36 @1500) or, based on the main effects specification, we fail to reject the null hypothesis of equal intercepts, income effects and cost effects across the three conditions (Model III: $F_{6,729}$=1.39; p=0.22). Further, even with all participants pooled regardless of whether they acquired information, there are no differences across Treatments 3, 4 and 5 (Model I: p=0.78 @1000; p=0.77 @1250; p=0.41 @1500).

Our most important findings are that information services decrease tax underreporting for those that acquire the information and—even though services are acquired roughly 58% of the time—for the service treatments as a whole (i.e. Hypothesis 4 is supported). The joint finding is most important since information acquisition is a choice, and it could simply serve to sort the players into inherently compliant and non-compliant groups—with the overall effect of information being a wash. Model II illustrates that those who access the information service have the lowest levels of underreporting. In fact, those receiving information under report roughly 80%, 70% and 60% less, across the respective income levels, as compared to those in the uncertainty baseline. The differences in tax underreporting levels is statistically significant beyond the 5% level between any of the three information service conditions and the certainty baseline, uncertainty baseline or no information condition (at any income level). Also evident from Model II, participants in the information service treatments who do not acquire information (i.e. the "No Information" subgroup) tend to have reasonably high levels of underreporting, albeit similar to the levels of underreporting in the uncertainty baseline.

Model I supports the finding that there is overall less tax underreporting in the information service treatments. Comparing Treatments 3, 4 and 5 with the uncertainty baseline suggests overall tax under reporting is cut in half. In eight of the nine possible cases, underreporting is significantly different—and lower—for the information service treatment relative to the uncertainty baseline. The lone exception is when comparing

Treatment 4 and the uncertainty baseline at the high income level, where the effect is in the expected direction but marginal (t=1.62; p=0.11). We summarize the results based on our analysis of the credit decision succinctly below:

Result 1. Tax underreporting decreases with the level of uncertainty (i.e. increases with income).

Result 2. The quality, as measured by accuracy, of the information service has no effect on tax underreporting.

Result 3. Those who acquire information underreport significantly less than those who do not. Further, *unconditional* on whether the information service was acquired, the availability of information services has the *overall* effect of reducing tax underreporting.

Concluding Remarks

Our most important finding is that, as predicted by economic theory, the provision of information—even when the quality is low—significantly increases tax compliance. Other hypotheses are not fully supported by the data and undoubtedly more about taxpayer behavior will be revealed through a more intricate data analysis. We have not yet investigated subject pool effects for these treatments but other work using data from similar experimental settings suggests that observed behavior is broadly consistent across pools (Alm, Bloomquist, and McKee, 2011).[14] Further research is being undertaken with the data reported in this paper to investigate the decision to acquire information as well as the factors affecting the propensity to take a second "draw" in the sequential information setting.

When the tax system is complex taxpayers are predicted to respond positively to the provision of information services that reduce the costs of computing true tax liabilities. The results reported here demonstrate that, first, with tax uncertainty the level of underreporting increases, second, when information services are provided the level of underreporting is lowered, and third, that the aggregate level of underreporting is lowered even when only a fraction (58%) of the participants avail themselves of the information service. The experimental setting does not incorporate a cost of the service to the tax agency but the improved underreporting behavior suggests a potential for a positive return from this service. As a final observation, the participants respond to the costs of the service in a predictable fashion. While the "costs" in the experimental setting are monetary, we would expect a similar response to higher costs in the form of transaction costs, including waiting time. We are currently researching this topic.

TABLE 1. Experiment Parameters

Parameter / variable	Value(s)
Earned Income	1000, 1250 or 1500 lab dollars
Audit Probability	30%
Penalty Rate	300% on unpaid taxes
Tax Rate	50% on taxable income
Tax Deduction	50% × Earned Income (pre filled on tax form)
Tax Credit	Expected value: 1000 – (0.5×Earned Income) Range: +/- 100% of expected value
Information Cost (if service is available)	0, 50 or 100 lab dollars

TABLE 2. Experiment Treatments

Tax Liability Uncertain	Service Provided?			
	No	One Source (Complete and Correct)	Two Simultaneous Sources (One Correct)	Two Sequential Sources (One Correct)
No	T1	N/A	N/A	N/A
Yes	T2	T3 Price of Information: $0, $50, $100	T4 Price of Information: $0, $50, $100	T5 Price of Information: $0, $50, $100

TABLE 3. Credit Decision Model I

Dependent Variable: Credit claimed—(Expected) credit					
	Treatment 1 (Certainty Baseline)	Treatment 2 (Uncertainty Baseline)	Treatment 3 (Perfect Info Available)	Treatment 4 (Simultan. Info Available)	Treatment 5 (Sequential Info Available)
Income=1000	160.36** (31.38)	219.28** (29.61)	72.61** (26.13)	46.18* (26.76)	59.82** (20.68)
Income=1250	186.97** (30.52)	260.38** (27.79)	148.82** (20.04)	135.95** (18.16)	154.43** (18.57)
Income=1500	257.32** (32.32)	310.31** (35.37)	208.37** (20.81)	242.83** (21.89)	206.37** (21.59)
			N=14,594		
			R^2=0.25		
			F=30.89**		

NOTES: * and ** denotes estimates that are statistically different from zero at the 10% and 5% significance levels, respectively. Cluster-robust standard errors are in parentheses.

TABLE 4. Credit Decision Model II

Dependent Variable: Credit claimed—(Expected) credit						
	Experiment Condition					
	Certainty Baseline	Uncertainty Baseline	No Information	Perfect Information	Two Information Sources	One Information Source
Income=1000	160.36** (31.38)	219.28** (29.61)	76.67** (24.83)	46.79** (22.45)	45.21** (16.75)	22.82 (36.57)
Income=1250	186.97** (30.52)	260.38** (27.79)	218.84** (19.07)	72.20** (20.10)	98.21** (12.59)	91.50** (27.16)
Income=1500	257.32** (32.32)	310.31** (35.37)	363.36** (21.01)	111.39** (17.52)	143.04** (14.43)	117.73** (44.99)
			N=14594			
			R^2=0.29			
			F=28.79**			

NOTES: * and ** denotes estimates that are statistically different from zero at the 10% and 5% significance levels, respectively. Cluster-robust standard errors are in parentheses.

TABLE 5. Credit Decision Model III

Dependent Variable: Credit claimed—(Expected) credit						
	Experiment Condition					
	Certainty Baseline	Uncertainty Baseline	No Information	Perfect Information	Two Information Sources	One Information Source
Intercept	-47.38 (102.03)	-64.83 (94.26)	-428.11** (87.21)	-81.34 (62.03)	-135.73** (52.09)	153.81 (141.60)
Income	0.20** (0.08)	0.26** (0.08)	0.59** (0.06)	0.16** (0.05)	0.20** (0.04)	0.17** (0.11)
Information Cost			-1.02** (0.47)	-1.31** (0.39)	-0.73** (0.24)	-0.85 (0.80)
			N=14,594			
			R^2=0.29			
			F=40.08*			

NOTES: * and ** denotes estimates that are statistically different from zero at the 10% and 5% significance levels, respectively. Cluster-robust standard errors are in parentheses.

References

Alm, James, Kim Bloomquist, and Michael McKee, 2011, "On the External Validity of Tax Compliance Experiments," prepared for presentation at 2011 IRS Research Conference, Washington, DC.

Alm, James, Todd Cherry, Michael Jones, and Michael McKee, 2008. "Encouraging Filing via Tax Credits and Social Safety Nets," *The IRS Research Bulletin: Proceedings of the2008 IRS Research Conference*, pp 43–57.

_____, 2010."Taxpayer Information Assistance Services and Tax Reporting Behavior."*Journal of Economic Psychology* 31 (4): 577-586.

Alm, James, Betty R. Jackson, and Michael McKee. 1992a. "Estimating the Determinants of Taxpayer Compliance with Experimental Data." *National Tax Journal* 65 (1): 107-114.

_____. 1992b. "Institutional Uncertainty and Taxpayer Compliance," *American Economic Review*, 82 (4): 1018-1026.

Alm, James and Michael McKee. 2006. "Audit Certainty and Taxpayer Compliance." *National Tax Journal* 59 (4): 801-816.

Cadsby, C. Bram, Elizabeth Maynes, and Viswanath Umashanker Trivedi. 2006. "Tax Compliance and Obedience to Authority at Home and in the Lab: A New Experimental Approach," *Experimental Economics*, 9 (4): 343-359.

Feld, Lars and Bruno Frey. 2002. "Trust Breeds trust: How Taxpayers are Treated," *Economics of Governance*, 3: 87-99.

Gravell, Jane, 2008, "Comments," *The IRS Research Bulletin: Proceedings of the2008 IRS Research Conference*, pp 59-60.

Greiner, Ben, 2004. "The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics," Working Paper Series in Economics 10, University of Cologne, Department of Economics.

Plott, Charles R. 1987. "Dimensions of Parallelism: Some Policy Applications of Experimental Methods." In Alvin E. Roth (ed.), *Laboratory Experimentation in Economics: Six Points of View*. New York, NY: Cambridge University Press.

Smith, Vernon. 1982. "Microeconomic Systems as an Experimental Science." *The American Economic Review* 72 (5): 923-955.

Vossler, Christian, Michael McKee, Michael Jones, 2010, "The Impact of Taxpayer Information Services on Tax Reporting and Tax Filing," presented at the 2010 meetings of the Southern Economics Association, Atlanta, GA.

Appendix

Example Experiment Summary Sheet (Treatment 3, Cost of $50)

Experiment Overview

- You will be participating in a market simulation that lasts several decision "rounds". In each round, you first play an earnings game and then face a tax reporting decision.
- In the earnings game, you sort the numbers 1 through 9. Your Income earned is determined by how fast you sort the numbers relative to others. The participant in your group with the fastest time receives the highest Income earned.
- In the tax reporting stage, you fill out and file a tax form. How much you earn from the tax reporting decision depends on how much you claim in Tax Credit and whether or not you are audited. Note that the on-screen instructions do not specify the tax policy parameters (e.g. tax rate, penalty rate, etc.), but those specified below will be in effect for this experiment.
- Each round is completely independent from the others, which means your decisions in one round in no way affect the outcome of any other round.

How Your Earnings Are Determined Each Round

- On the tax form, your Initial Taxes will be calculated automatically. This amount is determined by multiplying your Income earned by a tax rate of 50%.
- You decide how much to claim in Tax Credit on the tax form. Each dollar you claim in credits reduces your Final taxes by one dollar. This amount is subtracted from the Initial Taxes to determine your Final Taxes. If Final Taxes is a negative number, this reflects a tax refund.
- You will be shown a range of tax credits (this range is highlighted in white on the left side of the decision screen), which depends on your Income earned. Each amount within the range has an equal chance of being your actual tax credit, which is the highest amount you can claim without possible penalty. You can choose to claim any amount between 0 and 1000.
- You have an information service available to you at a cost of **$50**. By clicking on the "Request Information" button you will know the *exact* amount of your actual tax credit.
- You have a 30% chance of being audited. Audits are determined completely at random and do not depend on how much you or anyone else claims in tax credits.
- If you are not audited, your earnings for the round are your Income earned minus Final taxes.
- If you are audited, but claimed *less* than or *equal* to the actual tax credit, your earnings for the round are your Income earned minus Final taxes. Know that if you under-reported the credit you will not receive additional money through the audit process.
- If you are audited, and claimed *more* than the actual tax credit, you pay back the extra tax credit you claimed and also pay a penalty.
    - The penalty is equal to 300% multiplied by the amount of *extra* tax credit you claimed. Thus, if you claimed an extra $100 your penalty is $100*300% or $300.
    - Your earnings for the round are then Income earned *minus* Final taxes *minus* the extra tax credit you claimed *minus* the penalty.

Endnotes

[1] Funding provided by IRS under TRNO – 09Z – 00019. The views expressed are those of the authors and do not reflect the opinions of the IRS or of any researchers working within the IRS. An earlier version was presented at the New Perspectives on Tax Administration: An IRS-TPC Research Conference, June 22, 2011. We are grateful to the participants and to Marsha Blumenthal in particular for comments that improved the paper.

[2] The value of the taxpayer service derives from the costs imposed on the taxpayer for noncompliance. For the payoff maximizing individual, absent enforcement effort, service that resolves tax liability uncertainty would have no value to the taxpayer. However, a taxpayer wishing to honestly report would value the information since it would enable such honesty.

[3] Our experimental setting is very contextual and the presence of the income earning task provides, we argue, for the necessary degree of "parallelism" to the naturally occurring world that is crucial to the applicability of experimental results (Smith, 1982; Plott, 1987). The experimental setting need not—and should not—attempt to capture all of the variation in the naturally occurring environment, but it should include the fundamental elements of the naturally occurring world for the results to be relevant in policy debates. In this regard, our experimental design uses tax language (which is presented via the subject interface), requires that the participants earn income in each period, and also requires that the participants disclose tax liabilities in the same manner as in the typical tax form. As in the naturally occurring setting, there is a time limit on the filing of income. A clock at the bottom of the screen reminds the participants of the time remaining, and there is a penalty for failing to file on time set equal in all sessions to 10 percent of taxes owed; also, the individual is automatically audited if he or she fails to file on time, so that the participant pays the non-compliance fine as well.

[4] Such information reduces the cognitive burden of computing tax liabilities. The issue of tax liability uncertainty differs from enforcement uncertainty. As Alm, Jackson, and McKee (1992b) demonstrate, the tax authority may use enforcement uncertainty to increase compliance. Theory predicts that uncertain penalties increase compliance by risk-averse agents and this is borne out in the data from a set of experiments. Alm and McKee (2006) extend this and report on the compliance effects of informing the taxpayer their return will be audited with certainty.

[5] While overall audit rates are quite low, among certain income and occupation classes they are more frequent. The oft-reported IRS audit rate (currently less than one percent) is somewhat of an understatement. This reported rate usually refers to full audits. In fact, the IRS conducts a wide range of audit-type activities, including line matching and requests for information, and these activities are much more frequent. For example, in 2005 only 1.2 million individual returns (or less than one percent of the 131 million individual returns filed) were actually audited. However, in that year the IRS sent 3.1 million "math error notices" and received from third parties nearly 1.5 billion "information returns", which are used to verify items reported on individual income tax returns.

[6] See Alm, Jackson, and McKee, 1992a.

[7] Certain errors on the part of the taxpayer may not be easily verified in the event of an audit. For example, failure to claim a deduction for a charitable contribution because the taxpayer was uncertain of the status (e.g., 501c(3) status) of the organization may not be observed by the tax agency even in the event of an audit.

[8] An individual session included only students or non-student participants—they were not mixed in a session.

[9] Other experimental projects were ongoing at the time and participants may have participated in other types of experiments.

[10] This may be partially offset by the income effect since information is expected to be a normal good.

[11] This stems from the adage that "if a person has one clock she always knows what time it is but if she has two clocks she is never quite sure."

[12] The analysis excludes the 6 rounds out of 14,600 (730 participants x 20 rounds) where the tax form "timed out".

[13] This formulation is consistent with the theory, which is also from the perspective of the taxpayer. However, since the information services are unbiased, and given a large number of random credit draws are accumulated over participants and rounds, if we instead use the actual level of tax under reporting as the dependent variable this should only lead to trivial differences in results.

[14] Further, as noted above, Alm, Bloomquist, and McKee (2011) demonstrate the external validity of the experimental setting through a series of comparisons with field data results. This effectively addresses the criticisms of some who have questioned the use of lab experiments in tax compliance research (see Gravelle, 2008 (commenting on Alm et.al., 2008); Cadsby, Maynes, and Trivedi, 2006). Recall, for the current experiments we have conducted sessions at 2 institutions and with 2 pools (students and non-students) at each. Thus we have several ways the pool effects could be analyzed.

2009 Multi City Study of the Effect of Assistance on Compliance

Tiffanie N. Bruch, David C. Cico, and Saima S. Mehmood

The 2009 Multi City Study of the Effect of Assistance on Compliance was designed to examine Internal Revenue Service (IRS) service usage and the relationship between IRS service and compliance in a controlled environment. This research effort was modeled after a 1989 Price Waterhouse study conducted to measure the effect of assistance on voluntary compliance in which participants completed hypothetical tax situations.[1] Use of assistance in the Price Waterhouse study was high with a usage rate of 65 percent. Results from the Price Waterhouse study indicated that participants with assistance available had lower absolute error (i.e., commit fewer errors and/or increase tax revenue collections) than those without assistance and revealed no significant difference between types of assistance.

In July and August of 2006, Wage & Investment Research & Analysis (WIRA) conducted a pilot study in Atlanta, GA with 176 participants. Utilizing an experimental design in which participants completed a mock tax return similar in content to their own tax situation, the pilot study tested the compliance impact of the following IRS service channels: 1) telephone, 2) walk-in assistance, and 3) IRS.gov internet assistance. The pilot resulted in low service use (24 percent), low overall accuracy (20 percent), and the research was unable to establish a positive relationship between service and compliance. This current research effort expands and improves on the Price Waterhouse study and WIRA pilot by:

- Increasing motivation by changing the incentive structure from a flat rate compensation to a flat rate plus bonus for accuracy.
- Instituting higher quality recruiting by using participants who self-prepared their own tax return at least once in the last three years. For the pilot, participants were required to have completed their own tax return only once in the last five years.
- Increasing the number of participants and conducting the study in cities across the country to more thoroughly investigate and ascertain potential regional differences in the relationship between IRS service usage and compliance and to ensure accurate representation of the W&I taxpayer population.

Objectives

The objective of the research study was to quantify and measure customer preference for IRS service channels and the relationship between IRS service and compliance in a controlled environment for five types of tax scenarios: Taxable Social Security, Earned Income Credit (EIC), Itemized Deductions, Deductions that Could Not Be Itemized, and Standard Deduction scenarios.

Analysis of the findings was conducted in four phases, beginning first with overall analysis of all Multi City participants,[2] following with analysis of only Taxable Social Security participants,[3] then of EIC and CTC participants,[4] and finally of Schedule A participants.[5] The current report follows a similar structure, beginning first with overall analysis of all Multi City participants' use of service and accuracy on tax returns, followed by Taxable Social Security, EIC and CTC, and Schedule A participants' use of service and accuracy on tax returns.

Research Methodology

The research study utilized an experimental design in which participants, during two hour long sessions, completed a mock tax return similar in content to their own tax situation. All participants were screened for eligibility and asked about their personal tax situations prior to being assigned to a scenario group (see Appendix

A: Screener). During the study, participants were provided with an orientation to the study, a packet with instructions, mock tax scenarios, calculators, and pencils (see Appendix B: Session Instructions).

The IRS service channels tested included telephone assistance, which involved speaking with an IRS representative via 1-800-829-1040; walk-in assistance, which was modeled after IRS Taxpayer Assistance Centers using trained research staff members to provide assistance on site; and internet assistance, which was restricted to the use of www.IRS.gov only. All participants were provided with IRS forms and publications applicable to their scenarios. Participants were randomly assigned to one of five groups 1) walk-in assistance, 2) telephone assistance, 3) internet assistance, 4) assistance from any or all of the previously mentioned channels, or 5) no assistance.[6] All participants had access to forms and publications applicable to their tax scenarios. Participants' interactions with service were recorded in order to better understand their needs. The recordings were analyzed to determine which questions or topics from the mock tax scenarios taxpayers asked and how those questions and answers related to taxpayer errors.

The study design included a variable honorarium rate—between $60 and $100. Participants were instructed during their study orientation that more accurate responses would earn the higher honorarium amount. The rationale for the variable honorarium rate was based on the pilot study's unexpectedly low accuracy rates. Since real-life taxpayers are highly motivated to represent their tax liability accurately because a clear financial incentive exists, it was hypothesized that a variable incentive level could potentially increase accuracy rates to more closely approximate real-life motivation levels. In practice, all participants who demonstrated an effort to complete their forms with a reasonable degree of accuracy received the full $100 stipend.

Lastly, upon completion of the mock tax scenario, participants were given a debrief questionnaire to complete. The objective of the debrief questionnaire was to facilitate an understanding of the participant experience based on three phases:

- Systematic reflection and analysis of the Multi City experiment.
- Strengthening and personalization of their experience with the experiment to their own tax situation.
- Generalization and application of their tax situations to their broader financial situations.

More specifically, the debrief questionnaire addressed topics such as ability to complete the tax scenario, satisfaction with the provided IRS publications and forms, confidence in the accuracy of assistance received, confidence in the accuracy of tax return completed, past resources and/or services used to complete tax return, and attitudes concerning their financial situations (see Appendix C: Multi City Study Debrief Survey).

Sample Design

The population for this study consisted of taxpayers over the age of 18 who completed their own Federal income tax return with the form 1040 (1040EZ, 1040A, or 1040) series in the past three years. Participants were targeted based on a mix of demographic characteristics including gender, age, income, and internet access and use to ensure the taxpayer population was broadly represented. To the extent possible, recruitment also focused on representing participants as broadly as possible with respect to education, ethnicity, and tax filing status.

Testing sessions for the research study were conducted during the period of March to July 2009 in four geographically dispersed cities across the country (see Appendix D: Testing Locations). Selection was limited to cities with IRS facilities or IRS-approved federal facilities that could accommodate the following requirements:

- The availability to host 25-30 participants at each 2-hour session, with at least two sessions being held in the early evening.
- The ability to provide a minimum of six separate rooms for a) phones; b) computers; c) walk-in assistors; d) completion of mock tax scenarios; e) waiting; and f) greeting, debriefing, and provision of incentives.
- A minimum of three outgoing phone lines with a minimum of three phones.
- Internet access for three different computers.

PHASE 1: Overall Analysis of Multi City Participants

Participant Demographics

Participants were given one of five different scenario types which included: taxable Social Security, Earned Income Credit (EIC), itemized deductions, standard deductions, and deductions that could not be itemized. After adjusting for anomalies,[7] there were a total of 1,027 individuals who participated in the study. The following is a breakdown of participants by testing location:

- Atlanta, GA: 223 participants
- St. Louis, MO: 293 participants
- Boston, MA: 272 participants
- Seattle, WA: 239 participants

The sample consisted of an equal proportion of males and females. Participants were a majority Caucasian (71 percent), followed by 22 percent African American, and 7 percent of participants classified as some other race.

Most of the participants reported having completed an Associate's Degree or higher. Table 1 shows the distribution of participants reported education level.

TABLE 1: Participant Reported Education

Reported Education Level	Percentage
Bachelor's Degree	37%
Advanced Degree	25%
Some College, No Degree	18%
Associate's Degree	10%
High School Diploma/GED	9%
Trade/Vocational School Certificate	1%
Some High School	<1%

Over half of the participants reported working either full time or part time. Table 2 shows the distribution of participants' reported employment status.

TABLE 2: Participant Reported Employment Status

Reported Employment Status	Percentage
Employed Full Time	39%
Not Employed, but Looking	18%
Retired, Not Employed	17%
Employed Part Time, Not a Student	15%
Other	3%
Not Employed, Not Looking	3%
Retired, Employed Part Time	3%
Full Time Student, Not Employed	1%
Full Time Student, Employed Part Time	1%
Part Time Student, Not Employed	<1%
Part Time Student, Employed Part Time	<1%

The majority of participants (84 percent) reported having internet available in their home.[8] Ninety-five percent of participants said they use the internet at least occasionally and 93 percent of participants reported using email at least occasionally.

Participant Tax History

Forty-seven percent of participants electronically filed their Tax Year (TY) 2008 return using software, 30 percent filed using hand-prepared paper forms, nine percent filed v-coded returns,[9] seven percent had their return checked by a professional, and seven percent had a professional complete their 2008 return.[10] With respect to the type of Form 1040 participants used for their TY 2008 return, 70 percent filed a Form 1040 (70 percent), 16 percent filed a Form 1040A, and 14 percent filed a Form 1040EZ. Additionally, as reported in the debrief questionnaire, the majority (96 percent) of participants reported being the person in their household who was most familiar with tax preparation.

Participant Use of Service

Participants were randomly assigned to one of five service conditions: walk-in assistance (n=200), telephone assistance (n=208), internet assistance (n=197), assistance from any or all of the channels (n=209), and no assistance (n=213). All participants had access to IRS forms and publications applicable to their scenarios.

In total, 814 participants were eligible to use service. A total of 217 (27 percent) participants were recorded using service 283 times, with 43 of these participants using service more than once. Walk-in assistance was the most popular channel for seeking assistance; participants used this channel 186 times. Telephone assistance was used by participants 59 times, and internet assistance was used by participants 35 times. Among participants who had access to all three IRS service channels, walk-in assistance was again the most popular channel. When considering this groups' initial instance of service used, 68 percent chose to use walk-in assistance, 22 percent used telephone assistance, and 10 percent used internet assistance.

Of the 43 participants who sought service more than once,[11] 36 participants used walk-in assistance, five participants used telephone assistance, and two participants used internet assistance. Although 15 of these participants had the option to use any of the service channels, only five switched to a different service channel.[12] Of the five participants who were recorded using multiple service channels, three switched from telephone assistance to walk-in assistance, one switched from walk-in assistance to internet assistance, and one switched from walk-in assistance to telephone assistance.

Among those eligible to receive assistance but reported not using any assistance on their debrief form, 93 percent stated that they did not use assistance because they did not need the help. One percent reported that the wait time was too long, one percent reported not having help available to them, two percent said they did not know how to get help, and three percent gave some other reason for not seeking help. Nearly all participants (98 percent) reported using the IRS forms and publications that were provided to them in their scenario packages.

Service Questions and Issues by Scenario Type

When participants received service, their main question or issue was recorded by one of the research staff. Since questions differed by scenario type, Table 3 shows the top five questions/issues by scenario type.

Among Taxable Social Security participants, the most frequently asked question understandably related to Social Security benefits. For those given an Earned Income Credit (EIC) scenario, the most frequently asked questions pertained to Child Tax Credit (CTC) or EIC. For participants with Itemized Deductions and Could Not Itemize scenarios, the top questions or issues related to itemized deductions. Among participants who received a scenario with standard deductions, questions most often related to interest income, such as taxable interest on a banking account.

TABLE 3: Top Service Questions or Issues by Scenario Type

Taxable Social Security	EIC	Itemized Deductions	Could Not Itemize[13]	Standard Deduction
Social Security Benefits (n=18)	Child Tax Credit (n=10)	Itemized Deductions (n=20)	Itemized Deductions (n=33)	Interest Income (n=11)
Assistance with 1040 (n=10)	EIC (n=10)	CTC (n=10)	Standard Deduction (n=7)	Assistance with 1040A (n=9)
Standard Deductions (n=10)	Miscellaneous n=6)	ACTC (n=7)	ESP (n=6)	ESP (n=7)
Taxable Income (n=8)	Assistance with 1040 (n=5)	Advanced EIC (n=6)	Assistance with 1040 (n=5)	Tax Tables (n=7)
Miscellaneous (n=7)	Advanced EIC (n=5)	Interest Income (n=4)	CTC (n=4)	Miscellaneous (n=6)
	ACTC (n=5)			
Total Issues ***(n=87)***	Total Issues ***(n=75)***	Total Issues ***(n=71)***	Total Issues ***(n=75)***	Total Issues ***(n=66)***

NOTE: ACTC refers to Additional Child Tax Credit, CTC refers to Child Tax Credit, EIC refers to Earned Income Credit, and ESP refers to Economic Stimulus Payment.

Participant Confidence in the Accuracy of Assistance

Participants were asked to rate their confidence in the accuracy of the assistance they received on a scale of 1 to 8, where 1 is not at all confident and 8 is very confident.[14] Overall, participants reported being confident in the accuracy of the service they received. Among all participants who reported using service, the confidence in the service received had a mean of 6.8. However, there existed a statistically significant difference in the reported confidence level by the type of service first used. For those participants who used walk-in assistance, the mean confidence in the accuracy of the service received was 7.1. Participants who used telephone assistance on average rated their confidence in the accuracy of the service received as 6.6 while those participants using internet assistance reported a mean of 5.7. Consequently, participants who used either walk-in assistance or telephone assistance were significantly more confident in the accuracy of the assistance they received when compared to participants who used internet assistance.

Participant Accuracy on Tax Returns

Ninety-eight percent of participants reported on their debrief form that they were able to complete their scenario. Participants' completed scenarios were evaluated based on five critical lines of Form 1040 or Form 1040A including adjusted gross income (AGI), taxable income, total tax, total payments, and overpaid or amount owed. If the amount calculated on the critical line was within $1, it was considered correct to account for rounding.

Most participants (80 percent) correctly calculated AGI. Table 4 shows the overall accuracy rate for each critical line as well as the accuracy rate for each line when the graded line preceding it was correct.

Participant Accuracy Rates by Scenario

Accuracy rates varied depending on the scenario type (i.e., Taxable Social Security, EIC, Itemized Deductions, Could Not Itemize, or Standard Deductions). Table 5 shows the accuracy rates for the five critical lines by scenario type. Overall, participants who completed a Standard Deduction scenario were the most accurate, with 46 percent correctly completing all five critical lines on their tax returns. Additionally, Itemized Deductions participants were significantly more accurate than Taxable Social Security, EIC, and Could Not Itemize participants.

TABLE 4: Accuracy Rates by Critical Line on Tax Form

Critical Line	Percentage Correct	Participant Percentage Correct when Previous Line was Correct
Adjusted Gross Income	80%	N/A
Taxable Income	44%	53%
Total Tax	35%	73%
Total Payments	76%	76%
Overpaid/Amount Owed	28%	36%*

**Percentage correct when total payments was calculated correctly.*

TABLE 5: Accuracy Rates by Scenario Type

Scenario Type	Correct AGI	Correct Taxable Income	Correct Total Tax	Correct Total Payments	Correct Refund/ Balance Due	100% Accurate
	Percentage Correct					
Itemized Deductions	89%	60%	45%	91%	45%	40%
Standard Deductions	92%	64%	51%	90%	49%	46%
Could Not Itemize	88%	19%	16%	92%	16%	14%
EIC	87%	66%	59%	18%	17%	13%
Taxable Social Security	42%	15%	10%	72%	8%	7%
All Scenario Types	**80%**	**44%**	**35%**	**76%**	**28%**	**25%**

In addition to scenario type effects, there were also significant differences in accuracy within each scenario type. Table 6 shows accuracy rates by scenario. Some of the fictional tax scenarios within scenario types appeared to have been more difficult to complete compared to other scenarios. The Madison, Jackson, and Grant scenarios had significantly higher accuracy rates compared to other scenarios in their respective categories (see Appendix E: Multi City Scenarios for a complete description of research scenarios by scenario name).

Results also show that participants with fictional scenarios without dependents completed more accurate returns than those with fictional scenarios with dependents. Of those participants with scenarios that had no dependents, 34 percent correctly completed their tax return, compared to 16 percent of participants with scenarios that had dependents.

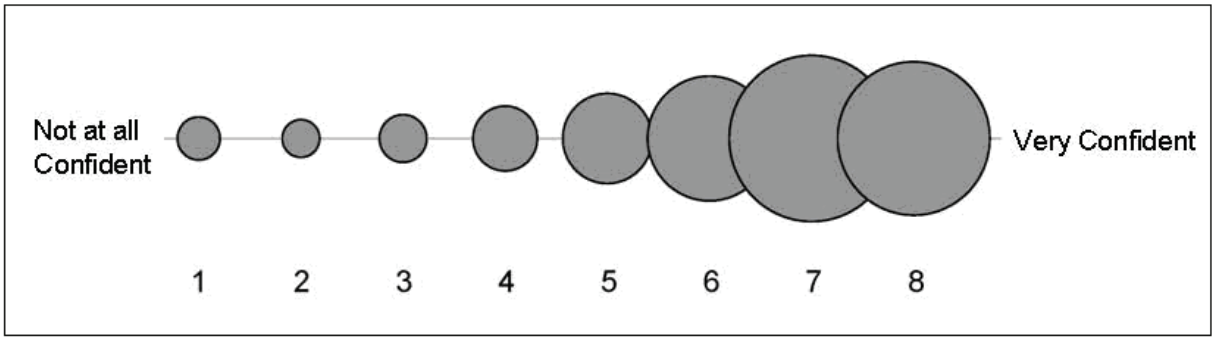
Participant Confidence in Return Accuracy

Although only approximately one in four participants correctly completed their fictional tax return (i.e., correctly calculated all five critical lines), the majority participants reported feeling somewhat or very confident in the accuracy of their return. Eighty-eight percent of participants rated their confidence as a 5 or higher, somewhat confident to very confident, and 60 percent of participants reported being very confident in the accuracy of their return (i.e., 7 or 8 on the scale). Refer to Figure 1 on the next page.

Overall Effect of Assistance on Compliance

After controlling for race, city, employment status, scenario type, and education effects, the use of IRS services was not a significant predictor of accuracy.[15] This is true for both overall accuracy rates as well as the accuracy of individual critical line items (see Appendix F: Effect of Service on Return Accuracy). Additionally, group assignments to a particular service channel were not predictors of accuracy.

FIGURE 1: Participant Confidence in Return Accuracy



Participants who reported on their debrief form using IRS forms and instruction booklets to complete their most recent tax return did significantly better on correctly completing their tax form when compared to participants who reported not using forms and instructions. Among those taxpayers who reported using IRS forms and instructions for TY 2008 to complete their tax return, 31 percent correctly completed their scenario compared to 19 percent of those participants who reported not using IRS forms and instructions. This result suggests that prior familiarity with IRS forms and publications may yield better accuracy rates among taxpayers. Additionally, participants who reported using www.IRS.gov to complete their previous tax return had higher accuracy rates compared to participants who did not report using www.IRS.gov; 31 percent compared to 20 percent accurately completed their scenarios respectively.

TABLE 6: Accuracy Rates by Scenario

Scenario Type	Scenario Name	Percentage Completely Accurate	Number of Participants
Taxable Social Security	McCook	9%	128
	Thornton	7%	14
	Brown	2%	62
EIC	Madison	31%	49
	Hood	4%	70
	Harrison	6%	33
	Adams	*	*
Itemized Deductions	Jackson	46%	104
	Washington	38%	99
	Pierce	19%	21
	Truman	*	*
Standard Deduction	Grant	54%	147
	Wilson	24%	37
	Taft	22%	18
	Fillmore	*	*
Could Not Itemize	Hayes	15%	103
	Chapman	13%	103
	Polk	5%	19
	Tyler	*	*

** Number of participants with scenario type equal to or less than 10 and not reported.*

PHASE 2: Taxable Social Security Scenarios

Twenty percent of the Multi City population completed a scenario with taxable social security benefits. In total, 204 participants completed one of three taxable social security scenarios: 63 percent completed scenario McCook, 30 percent completed Brown, and 7 percent completed scenario Thornton. Most of these participants (83 percent) indicated to the screener before the experiment that their most recent tax return included taxable social security.

Seventy-one percent of Taxable Social Security participants were 65 years of age or older. This figure is significantly higher than the percentage of all Multi City participants of age 65 or higher (21 percent). Additionally, 58 percent of Taxable Social Security participants reported being retired and not employed, compared to 17 percent of all Multi City participants who reported the same employment status.

Taxable Social Security Participant Accuracy on Tax Returns

As stated previously, Taxable Social Security participants were least likely to complete an accurate return compared to other participants in the study; only seven percent of Taxable Social Security respondents completed an accurate return.

With respect to line 20b or 14b for taxable social security on Form 1040 or Form 1040A, 41 percent of Taxable Social Security participants entered the correct value. One common mistake among these participants was that they incorrectly recorded that either all or none of their social security income was taxable.

Although accuracy rates were very low for Taxable Social Security participants, most participants felt that they had completed an accurate return. The majority of participants (83 percent) rated their confidence in the return they completed as a 5 or higher, somewhat confident to very confident, and 51 percent of participants reported being very confident in the accuracy of their return (i.e., 7 or 8 on the scale).

Social Security Benefits Worksheet

One hundred and fifty-seven Taxable Social Security participants (77 percent) were recorded using the Social Security Benefits Worksheet that was available to them in the 1040 and 1040A instruction booklets (i.e., they made at least one entry on the worksheet; see Appendix G: Social Security Benefits Worksheet). Of these 157 participants, 82 percent completed lines 1 and 18 of the worksheet, indicating that they had completed the entire worksheet.

Of the participants who completed the Social Security Benefits Worksheet, 57 percent recorded a correct amount on lines 20b or 14b of Form 1040 or Form 1040A, respectively, the lines for taxable social security benefits. However, from the worksheet data we find that 60 percent of participants who completed the worksheet calculated the last line of the worksheet correctly. The difference between the two is attributable to errors when participants transferred the numbers from worksheet to 1040 or 1040A, and from participants recording the taxable social security amount on another line on the 1040 or 1040A. Table 7 shows the results of the line-by-line analysis of the Social Security Benefits Worksheet. In general, the errors made by participants can be grouped into two categories: calculation errors and decisions errors.

Accuracy rates for line 5 of the worksheet, which should be the sum of lines 1 through 4, drop significantly due to participants either incorrectly carrying down the value from line 1, 2, or 3 without adding them, or adding incorrect amounts together. At line 16, another addition line, there is another drop in accuracy.

For those participants who did not complete the worksheet provided, only 16 percent correctly calculated their taxable Social Security. Results from a chi-squared analysis showed a significant difference with regard to accuracy of taxable Social Security between those who completed the worksheet and those who did not complete the worksheet, in that participants who completed the worksheet were significantly more likely to calculate the correct taxable Social Security amount.

TABLE 7: Common Errors on the Social Security Benefits Worksheet

Line #	Percent Correct	Common Mistakes
1	90%	Only entered one SSA-1099
2	88%	Carryover mistakes from line 1
3	94%	Incorrect amounts from other lines on 1040/1040A
4	100%	
5	78%	Carryover mistakes from previous lines, brought down line 2 or 3, and calculation error on line 5
6	98%	Entered adjustment amounts that did not exist
7	78%	Same mistakes as line 5
8	96%	Took wrong amount or put a zero
9	77%	Carryover mistakes from lines 7 & 8
10	98%	Took wrong amount
11	75%	Subtracted line 10 from 9 incorrectly
12	80%	Carryover calculation mistakes from previous lines
13	79%	Carryover calculation mistakes from previous line
14	80%	Carryover calculation mistakes from previous line
15	70%	Entered zero on line 15, carryover mistakes from line 11
16	65%	Carryover calculation mistakes from lines 14 and 15
17	72%	Carryover error from line 1
18	60%	Carryover mistakes from line 1 and line 11

Taxable Social Security Participant Use of Service

Taxable Social Security participants were randomly assigned to one of five groups; walk-in assistance (n=44), telephone assistance (n=34), internet assistance (n=46), assistance from any or all of the channels (n=47), and no assistance (n=33). All participants had access to IRS forms and publications applicable to their scenarios.

In total, 171 Taxable Social Security participants were eligible to use service. A total of 47 participants (27 percent) were recorded using service 66 times, with 12 of these participants using service more than once. Walk-in assistance was the most popular channel for seeking assistance; participants used this channel 49 times. Telephone assistance was used by participants 11 times and internet assistance was used by participants six times. Among participants who had access to all three IRS service channels, walk-in assistance was again the most popular channel. When considering participants in this group who used service, 75 percent chose to use walk-in assistance, 17 percent used telephone assistance, and nine percent used internet assistance.[16]

As shown previously in Table 3, the most frequently asked question among Taxable Social Security participants pertained to taxable Social Security benefits. Other popular topics included questions about standard deductions and taxable income. The majority of Taxable Social Security participants (84 percent) reported being somewhat to very confident in the accuracy of the service they received.

Among those eligible to receive assistance who reported not using any assistance on their debrief form, 93 percent reported that they did not use assistance because they did not need the help. Two percent said they did not know how to get help, and five percent stated another reason for not seeking help. All participants reported using the IRS forms and publications that were provided to them in their scenario packages.

Analysis showed no significant difference between those who used service and those who did not use service with respect to correctly computing taxable Social Security. Among those who used service, 42 percent correctly computed the Social Security amount. For those who did not use service, 45 percent recorded the correct taxable Social Security amount.

Recommendations for Taxable Social Security

Although use of IRS service channels did not have a significant effect on compliance for Taxable Social Security participants, those participants who used the Social Security Benefits Worksheet were significantly more likely to compute the taxable portion of their Social Security. This finding highlights the importance of the Social Security Benefits Worksheet in completing a more accurate return. A potential factor for participants not using the worksheet could have been that participants did not recognize or follow the instructions listed on the Form 1040 or Form 1040A directing them to calculate the taxable portion of their benefits. Additionally, participants may have felt that the instructions did not apply to them or to the completion of their scenario. Therefore, WIRA recommended identifying methods to increase taxpayer awareness and use of the Social Security Benefits Worksheet.

One method for increasing awareness and use of the worksheet would be to partner with the Social Security Administration (SSA) to mail the Social Security Benefits Worksheet with the SSA-1099 statement that taxpayers receive. Additionally, WIRA suggested partnering with Stakeholder Partnerships, Education, and Communication (SPEC) to develop workshops on how to compute portions of social security benefits. By leveraging partnerships with groups such as AARP, SPEC could raise awareness of the Social Security Benefits Worksheet through these workshops.

Furthermore, since the line-by-line analysis indicated drops in accuracy at certain lines, WIRA suggested giving taxpayers visual cues to prompt expected taxpayer behavior, such as adding or subtracting. The worksheet currently has stop signs which alert taxpayers that a decision must be made at these junctions. By including additional visual cues near the lines to the right of the form, taxpayers may be more aware of what steps are next required to accurately complete the worksheet.

Finally, since the majority of taxpayers who use the Social Security Benefits Worksheet are 65 years of age or older, WIRA suggested using larger print so that these taxpayers have an easier time reading and understanding the worksheet.

PHASE 3: Earned Income Credit (EIC) Scenarios

A total of 155 participants (15 percent of all Multi City participants) were given an EIC scenario type. These participants were provided with one of four different scenarios, each with set fictional tax data with which to prepare the returns. The fictional individuals whose tax returns were being prepared were eligible for EIC, Child Tax Credit (CTC), or Additional Child Tax Credit (ACTC), or had collected Advance Earned Income Credit (AEIC) payments. See Table 8.

TABLE 8: Scenario Breakdown of EIC Participants

Scenario Name	Percentage	Credit Eligibility/Payments Made
Madison	32%	EIC
Harrison	21%	EIC, ACTC
Hood	45%	EIC, AEIC, CTC, ACTC
Adams	2%	AEIC, CTC, ACTC

Of these participants, 71 percent indicated to the screener that they had claimed EIC payments on their most recent tax return. Additionally, 26 percent of participants reported having claimed CTC on their 2008 return.

The majority of EIC participants (46 percent) indicated that their filing status for TY 2008 was Head of Household. Thirty-one percent indicated their filing status as Single, 21 percent as Married Filing Jointly, and two percent as Married Filing Separately.

With respect to employment, the majority of EIC participants were employed full-time (39 percent), employed part-time and not a student (26 percent), or not employed but looking for work (22 percent). The latter two of these three percentages differ significantly from those for the entire Multi City population; 39 percent of

all Multi City participants reported being employed full-time, 15 percent reported being employed part-time and not a student, and 18 percent were not employed but looking for work.

EIC Participant Accuracy on Tax Returns

Based on the accuracy of the five critical lines mentioned previously, only 13 percent of EIC participants completed their returns correctly. Accuracy was especially low for total payments and refund/balance due: 18 percent and 17 percent respectively.

The significantly higher accuracy rates for EIC participants on the lines for AGI, taxable income, and total tax compared to non-EIC participants can likely be attributed to the EIC participants' lack of any adjustments to total income and to lack of any taxes paid, other than AEIC for eligible scenarios. Conversely, EIC participants' significantly lower accuracy rates for total payments and overpaid/amount owed can likely be attributed to failure to correctly calculate EIC and ACTC; these two credits are entered on two of 11 lines that are used to calculate total payments and overpaid/amount owed.

Despite low return accuracy rate, an overwhelming majority of EIC participants rated confidence in the return they completed as somewhat to very high. Based on a scale of 1 to 8, only 15 percent of participants rated their confidence as somewhat to very low (1 through 4), while 85 percent marked 5 through 8 on the scale to indicate somewhat to very confident. Fifty percent of EIC participants rated their confidence as very high (7 or 8).

Earned Income Credit

Of the 155 EIC participants, 152 were eligible to claim EIC on their fictitious return. Of these participants, 48 percent entered the correct value of this credit on their tax return, 26 percent entered an incorrect value, and the remaining 26 percent entered "0" on this line or left it blank. The percentage of EIC participants in the Multi City study who claimed any amount of the credit on their fictional return (74 percent) is similar to actual filing estimates of the number of taxpayers who receive EIC payments as compared to the number of all taxpayers who are eligible to receive payments; eligible taxpayer participation rate for EIC is estimated to be 75 percent and appears to be relatively stable over time.[17]

Instructions for determining EIC eligibility could be found within standard 1040 instruction booklets that were distributed to all participants (see Appendix H: EIC Instructions). Of all EIC participants, 44 percent were recorded having used the instructions, and of this sample, 75 percent were recorded as having entered the correct EIC value on their Form 1040 or Form 1040A. Of the 56 percent of EIC participants who did not use the instructions, only 27 percent correctly calculated EIC. Table 9 provides the distribution of EIC accuracy between groups that did and did not use the instructions for those participants who were eligible to receive the credit.

TABLE 9: EIC Participant Use of Instructions and EIC Accuracy

EIC Instructions Used	Percentage	EIC Calculation	Percentage
Yes	44%	Correct	75%
		Incorrect	25%
No	56%	Correct	27%
		Incorrect	73%
Total	100%	Correct	48%
		Incorrect	52%

Results from statistical analyses again show that those participants who used the EIC instructions were significantly more likely to correctly calculate EIC than those participants who did not use the worksheet.

EIC Worksheet A, consisting of six questions, was included in the instruction booklet that was distributed to all participants as well (see Appendix I: EIC Worksheet A). Of those participants who were eligible to claim EIC on their return, 40 percent were recorded as having used Worksheet A, and of this sample, 80 percent

correctly entered the EIC value on their tax return. Conversely, 27 percent of participants who did not use the worksheet to determine their EIC correctly calculated the amount of the credit (see Table 10).

TABLE 10: EIC Participant Use of Worksheet A and EIC Accuracy

EIC Worksheet A Used	Percentage	EIC Calculation	Percentage
Yes	40%	Correct	80%
		Incorrect	20%
No	60%	Correct	27%
		Incorrect	73%
Total	100%	Correct	48%
		Incorrect	52%

Results from statistical analyses again showed that those participants who used Worksheet A were significantly more likely to correctly calculate EIC than those participants who did not use the worksheet.

In order to claim EIC with qualifying children, taxpayers must fill out and submit Schedule EIC with their returns. Two of the three scenarios eligible for EIC required this schedule to be completed and attached to the tax return. Ninety-five percent of participants in these two scenarios were recorded as having filled out this worksheet; that is, these participants were recorded as having entered any value on any line of the worksheet.

Advance Earned Income Credit

Taxpayers who qualify for EIC can elect to receive a portion of the credit in advance as a part of the paycheck that is issued by their employer, instead of receiving all of the credit at once as a part of their income tax refund. The amount of AEIC received by the employee is reported on his or her Form W-2, in box 9, and this value must be entered onto line 60 of Form 1040 or line 36 of Form 1040A.

Two scenarios from the present study, Hood and Adams, received AEIC payments, the values of which were recorded on their fictional Forms W-2. Seventy-one percent of these participants correctly entered the AEIC value on their tax returns; the remaining 29 percent either entered "0" or left this line blank on their return.

Child Tax Credit

Within the EIC scenario type, 47 percent of participants with two scenarios, Hood and Adams, were eligible to claim CTC on their tax returns. Of these eligible participants, 22 percent claimed the correct amount of CTC, 34 percent claimed the incorrect amount of CTC, and 44 percent did not claim the credit on their return. Eight ineligible EIC participants claimed the CTC credit in error.

Five of the 16 Multi City scenarios other than those included in the EIC scenario type were eligible to claim CTC: Thornton, McCook, Washington, Chapman, and Taft. Of the 362 eligible participants in these five scenarios, 56 percent entered the correct value for CTC on their Form 1040 or Form 1040A, nine percent entered the incorrect value for CTC, and 36 percent failed to claim the credit on their return.[18] Among all Multi City participants who were not eligible, four percent (22 participants) claimed the CTC credit in error.

The CTC Worksheet, included in the instruction booklet that was available to all participants, helps taxpayers to determine both the amount of CTC they can claim on their return as well as whether they may be able to claim the Additional Child Tax Credit (ACTC) on their return (see Appendix J: CTC Worksheet).

Of EIC participants who were eligible to claim CTC on their returns, 80 percent used the worksheet. However, overall accuracy of these participants in calculating CTC was low: only 26 percent of participants who used the worksheet entered the correct CTC value on their Form 1040 or Form 1040A. This low accuracy rate can likely be attributed to a failure to correctly add total tax on their Form 1040 or Form 1040A, as this value was also to be entered on line 2 of the CTC worksheet.

Among non-EIC participants who were eligible to claim CTC, 74 percent used the worksheet to calculate the amount of credit for which they were eligible, and of these participants, 71 percent claimed the correct amount of CTC on their returns. All eligible non-EIC participants were allowed to claim exactly $1,000 for CTC, whereas eligible participants within the EIC scenario type were allowed to claim $613 and $900, respectively. The additional calculation to determine total tax for eligible EIC participants may have contributed to lower accuracy rates for this group as compared to accuracy rates for non-EIC participants.

Additional Child Tax Credit

In total, 106 (68 percent) of the 155 EIC participants were eligible to claim ACTC. Thirteen percent of participants who were eligible to claim ACTC on their returns entered the correct value for this credit on either line 66 of Form 1040 or line 41 of line 1040A, 17 percent claimed an incorrect amount of this credit, and 70 percent of eligible participants failed to claim this credit on their returns.[19]

Taxpayers may be eligible to claim ACTC payments despite being ineligible to claim CTC payments on their returns. In order to determine how much, if any, of the credit they are able to claim, taxpayers with at least one qualifying child are invited to use the ACTC worksheet, Form 8812 (see Appendix K: Form 8812).

Of those participants who were eligible to claim ACTC, 73 percent used Form 8812. Accuracy rates, however, for these participants were low; 18 percent of those who used the worksheet entered the correct ACTC value on their tax returns, while none of the participants who did not use the worksheet entered the correct amount. Table 11 shows line-by-line accuracy and likely reasons for errors made by those who completed Form 8812. Lines 4b and 7 through 12 are not included in the table as these lines were not applicable to any of the scenarios in the study.

TABLE 11: Line-by-line Accuracy for Form 8812, ACTC Worksheet

Line Number	Percentage Correct	Likely Reason for Error
1	34%	Incorrect number of qualifying children
2	43%	Incorrect calculation of CTC
3	20%	Math error (subtraction of line 2 from line 1)
4a	31%	Incorrect calculation of earned income
5	20%	Math error (subtraction of $8,500 from line 4a)
6	20%	Carryover error from line 5
13	17%	Decision error

Earned Income Credit Participant Use of Service

Participants with EIC scenarios were randomly distributed across one of the five service groups: walk-in assistance (23 percent), telephone assistance (21 percent), internet assistance (14 percent), assistance from all three channels (19 percent), or no assistance (23 percent). All participants were provided access to IRS publications and forms that were applicable to their scenarios.

Of those with EIC scenarios, 77 percent were eligible to use service. Thirty-seven of these individuals (31 percent) were recorded using service 52 times, with nine participants electing to use service more than once. Participants were most likely to use walk-in assistance, as 40 out of the 52 service uses (77 percent) were through this channel. Telephone assistance was used six times (12 percent), as was internet assistance.[20] Walk-in assistance was the most popular channel for those participants who had access to all three service types, and it was the only channel for those who used service for a second, third, or fourth time.

Among EIC participants who were recorded as having used service, 88 percent rated their confidence in the accuracy of assistance received as somewhat to very high, with 68 percent of participants rating confidence in assistance as very high. These percentages are comparable to the entire Multi City population's confidence levels in accuracy of assistance; of those participants who were recorded as having used service, 84 percent rated their confidence in accuracy as somewhat to very high, with 55 percent of participants rating confidence as very high.

As stated earlier, 31 percent of the EIC sample that was eligible to use service was recorded as having used one of the three service channels. This percentage is significantly higher than that of Multi City participants not in the EIC scenario that used service; of the 692 participants in the other four scenarios who were eligible to use service, 180 participants (26 percent) elected to use service.

Participants who were eligible to receive assistance but indicated that they did not use any of the channels on their debrief form cited "Did not need help" as the reason for not using service 86 percent of the time. Nine percent stated that they did not use service due to "Other" reasons. Three percent did not use service because they "Did not know how," and two percent did not use service due to "Wait time too long/too many other people in line." Ninety-seven percent of EIC participants cited that they used IRS publications or instruction booklets to complete their returns.

Consistent with findings from the Taxable Social Security participants, use of service failed to denote an indicator of return accuracy. No significant difference emerged between participants who used service and those who did not use service with respect to entering correct EIC values on tax returns and to completing an overall more accurate return.[21]

Recommendations for Earned Income Credit and Child Tax Credit

Although use of walk-in, internet, and telephone assistance did not significantly impact compliance for EIC participants, the findings highlight the importance of completing the EIC instructions and worksheets. A potential factor for participants not using the instructions and worksheets could be that participants were not aware of the existence of these two tools within the 1040 instruction booklets that were provided to them. Additionally, participants may have felt that the instructions did not apply to them or to the completion of their scenario. As the cost of increasing awareness is less than the cost of processing errors and amended returns, WIRA recommended identifying methods to increase taxpayers' awareness and use of both the EIC instructions and Worksheet A. Preemptive messages in instruction booklets and on forms should be direct enough to prompt taxpayers to use worksheets that apply to their situations. Emphasis should also be placed on outreach to individuals filing as Single and with no children, as individuals in this scenario were particularly unaware of their eligibility for claiming EIC.[22]

Simplification of the steps needed to accurately claim EIC is also advisable. As having to complete several steps in the EIC instructions, Worksheet A, and if applicable, Schedule EIC, in order to claim this credit is burdensome to taxpayers, WIRA recommended exploring options to reduce taxpayer burden associated with calculating EIC. For example, two different types of EIC instructions can be included in 1040 and 1040A instructions booklets: one for taxpayers with qualifying children, and one for taxpayers with no qualifying children. The third step of the 2008 EIC instructions, which let taxpayers know whether they have qualifying children, can be a preliminary step that would lead taxpayers to either one of the two types of instructions. A breakdown of this type could potentially increase the proportion of taxpayers with no qualifying children who claim this credit.

Because sample sizes were not large enough to conduct statistical analysis of the accuracy rates of the CTC and ACTC worksheets, WIRA recommended comprehension testing to examine the effectiveness of these worksheets and the value of different types of preemptive messages through publications, phone scripts, and online scripts. Comprehension testing would involve participants filling out worksheets and being exposed to these messages, and subsequently completing debriefs to explain their thought processes. These responses would allow for the determination of how to increase accuracy and to alleviate difficulties that taxpayers face when trying to understand instructions.

Lastly, since several participants filled out unnecessary forms when completing their scenarios, WIRA suggested having a check sheet for paper filers to avoid undue burden when completing a tax return. Much like tax software does, this check sheet would list life events and situations, and would prompt the taxpayer that they may be eligible for a certain credit and give information on steps to complete the process. This type of check sheet could potentially reduce taxpayer burden by informing taxpayers up front which forms or worksheets apply to their tax situation.

PHASE 4: Schedule A Scenarios

A total of 460 participants (45 percent of the entire sample) were tasked with one of two scenario types that required the use of Schedule A to correctly determine deductions (refer to Appendix L: Schedule A).[23] Descriptions of the fictional individuals in these scenario types included paragraphs stating that although these individuals had not itemized their deductions last year, they may be able to itemize this year. This narrative was followed by a list of potentially applicable receipts for expenses that are deductible on Schedule A. If a fictional individual's itemized deduction was less than his or her standard deduction, the participant should have entered the standard deduction on his or her return instead of the itemized deduction that was calculated using Schedule A. See Table 12 for a breakdown of Schedule A scenario types.

TABLE 12: Schedule A Scenarios

Scenario Type	Scenario Name	Percentage of Schedule A Participants
Itemized Deductions	Jackson	23%
	Washington	22%
	Pierce	5%
	Truman	1%
Could Not Itemize	Chapman	22%
	Hayes	22%
	Polk	4%
	Tyler	1%
Total		100%

When asked if they had itemized deductions on their most recent tax return, 49 percent of Schedule A participants indicated to the screener that they had itemized.[24] Conversely, when asked if they had taken a standard deduction on their most recent tax return, 30 percent of Schedule A participants affirmed that they had. Twenty-one percent of Schedule A participants did not indicate what type of deduction they had taken the previous year, and less than one percent indicated that they did not know what type of deductions they had taken on their most recent return.

Seventy-one percent of Schedule A participants reported having completed a Bachelor's Degree or higher. In comparison, 52 percent of participants in non-Schedule A scenarios (Standard Deductions, Taxable Social Security, and EITC) reported having completed a Bachelor's degree or higher. Also, while 50 percent and nine percent of Schedule A participants reported being employed full-time and retired/not employed, respectively, 30 percent and 24 percent of non-Schedule A participants reported being employed full-time and retired/not employed, respectively.

Schedule A Participant Accuracy on Tax Returns

Overall, participants with Itemized Deductions scenarios were significantly more likely to complete an accurate return when compared to Could Not Itemize participants. See Table 13 for Schedule A participants' accuracy by the five critical lines.

Accuracy for Could Not Itemize participants fell significantly for taxable income, total tax, and overpaid/amount owed when compared to Itemized Deductions participants. This difference can likely be attributed to the significantly lower accuracy of deductions for Could Not Itemize participants. Sixty-eight percent of participants with the Itemized Deductions scenario type took the correct deduction on line 40 of Form 1040, while only 22 percent of participants in the Could Not Itemize scenario type took the correct deduction on either line 40 of Form 1040 or line 24 of Form 1040A.[25] The value of deductions subsequently affects the amount of taxable income, total tax, and overpaid/amount owed that participants calculated.

TABLE 13: Schedule A Participant Accuracy on Tax Returns

Critical Line	Percentage Correct	
	Itemized Deductions	Could Not Itemize
Adjusted Gross Income (AGI)	89%	88%
Taxable Income	60%	19%
Total Tax	45%	16%
Total Payments	91%	92%
Overpaid/Amount Owed	45%	16%
Completely Accurate	40%	14%

An overwhelming majority of participants in both scenario types rated confidence in the accuracy of the return they had prepared as somewhat to very high on an eight-point scale, despite the especially low return accuracy rates for participants with Could Not Itemize scenarios. Only five percent of Itemized Deductions participants and Could Not Itemized participants rated their confidence as somewhat to very low (1 through 4), while 95 percent rated their confidence as somewhat to very high (5 through 8).

Itemized Deductions on Schedule A

Schedule A, a worksheet allowing taxpayers to calculate the amount of their itemized deductions, is divided into eight sections: medical and dental expenses; taxes paid; interest paid; gifts to charity; casualty and theft losses; job expenses and certain miscellaneous deductions; other miscellaneous deductions; and total itemized deductions. If the amount of itemized deductions on Schedule A is higher than the taxpayer's standard deduction, he or she can deduct the itemized amount on Form 1040, line 40.

Ninety-seven percent of all Schedule A participants were recorded as having used Schedule A; that is, these participants made any entry on any of the lines on the form. The percentage of participants in both scenarios who correctly calculated critical lines[26] from each of the eight sections on the worksheet can be seen in Table 14.

TABLE 14: Line-by-line Accuracy for Schedule A

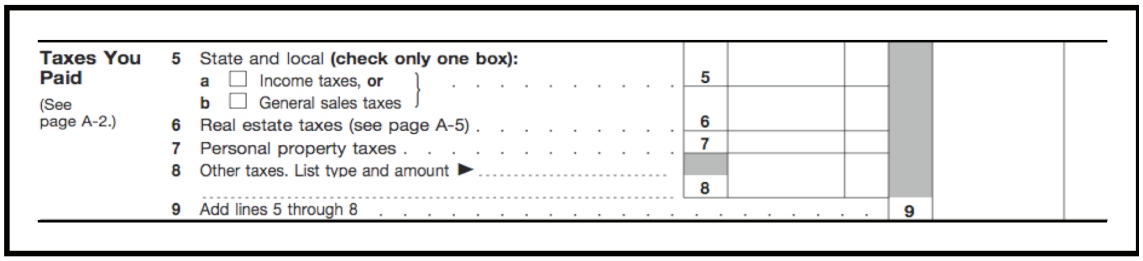
Line Number	Section Description	Percentage Correct
4	Medical and dental expenses	90%
9	Taxes paid	83%
15	Interest paid	97%
19	Gifts to charity	97%
20	Casualty and theft losses	100%
27	Job expenses and certain miscellaneous deductions	97%
28	Other miscellaneous deductions	97%
29	Total itemized deductions (sum of lines 4, 9, 15, 19, 20, 27, and 28)	74%

While accuracy for critical lines on Schedule A was relatively high, participants made a number of common errors worth noting. First, two of the eight scenarios involved fictional individuals with receipts for medical premiums that had been payroll deducted pre-tax. Because these premiums had been deducted pre-tax from the individuals' paychecks, these premiums should not have been deducted again as a medical expense on Schedule A. However, 49 of the 111 participants (44 percent) with these two scenarios made the error of adding these medical premiums on line 1 of Schedule A, which is the line for medical and dental expenses.

In the second section of Schedule A, "Taxes You Paid," 14 percent of participants incorrectly entered line 7, the line for personal property taxes. Typically, taxes paid on motor vehicles can be entered on this line, and taxes paid on home properties can be entered on line 6, which is the line for real estate taxes.[27] However, par-

ticipants' confusion regarding lines 6 and 7 caused 42 percent of those who incorrectly entered line 7 to enter real estate taxes here (see Figure 2 below).

FIGURE 2: Taxes You Paid on 2008 Schedule A



For line 9 of Schedule A, the sum of deductible taxes paid, 17 percent of participants incorrectly entered this value. The most common error among these participants was that 35 percent did not add motor vehicle taxes into this sum, an error carrying over from line 7.

Line 29, the line with the sum of all itemized deductions, had a significantly lower accuracy rate than that of the previous seven critical lines[28]; 74 percent of participants who filled in line 29 entered the correct amount. This lower accuracy rate in comparison to accuracy on previous lines may be attributed to the situation in which a participant correctly calculated five or six of the previous seven lines, with his or her final sum on line 29 being incorrect because of the one or two miscalculated lines. In fact, 73 percent of participants who used Schedule A had correctly entered all seven critical lines, and 95 percent of these participants also correctly entered line 29. Twenty-three percent of participants who used Schedule A correctly calculated six of the seven critical lines, with 25 percent of these participants correctly entering line 29. Lastly, five percent of participants who used Schedule A correctly entered five or fewer of the seven critical lines, and just two of these participants correctly entered line 29.

Schedule A Participant Use of Worksheets and Deduction Accuracy

Of participants with Itemized Deductions scenarios, 98 percent used Schedule A. Table 15 shows the accuracy of participants with Itemized Deductions scenarios who did and did not use Schedule A.

Of participants with Could Not Itemize scenarios, use of Schedule A had no effect on accuracy of the standard deduction entered on tax returns (see Table 16).

Ninety-five percent of participants with Could Not Itemize scenarios were tasked with filing the return of a fictional individual who had paid real estate taxes. These participants should have used the Standard Deduction Worksheet in the instruction booklet to account for real estate taxes paid that were added into their standard deductions (see Appendix M: Standard Deduction Worksheet). However, only 23 percent of these participants used the Standard Deduction Worksheet, and these participants were significantly more likely to calculate the correct standard deduction than those who did not use the worksheet.[29] See Table 17 for accuracy of participants with scenarios in which the fictional individuals paid real estate taxes.

Schedule A Participant Use of Service

Participants with Schedule A scenarios were randomly assigned to one of the five service conditions: walk-in assistance (18 percent), telephone assistance (21 percent), internet assistance (19 percent), assistance from all three channels (20 percent), and no assistance (22 percent). All participants were provided with IRS instruction booklets and forms that were applicable to their scenarios.

TABLE 15: Use of Schedule A and Itemized Deduction Accuracy

Schedule A Used	Percentage	Itemized Deduction on Form 1040	Percentage
Yes	98%	Correct	69%
		Incorrect	31%
No	2%	Correct	0%
		Incorrect	100%
Total	100%	Correct	68%
		Incorrect	32%

TABLE 16: Use of Schedule A and Standard Deduction Accuracy

Schedule A Used	Percentage	Standard Deduction on Form 1040 or 1040A	Percentage
Yes	96%	Correct	22%
		Incorrect	78%
No	4%	Correct	22%
		Incorrect	78%
Total	100%	Correct	22%
		Incorrect	78%

TABLE 17: Use of Standard Deduction Worksheet and Standard Deduction Accuracy

Standard Deduction Worksheet Used	Percentage	Standard Deduction on Form 1040 or 1040A	Percentage
Yes	23%	Correct	69%
		Incorrect	31%
No	77%	Correct	6%
		Incorrect	94%
Total	100%	Correct	21%
		Incorrect	79%

Of the 78 percent of all Schedule A participants who were eligible to receive assistance, 76 percent chose to not use service, and the remaining 24 percent used service 101 times. Fifteen percent of the individuals who used service elected to use service more than once. The modest incidence of multiple contacts implies that participants who used service felt confident in applying the information they received, even though accuracy did not increase for these participants.

As stated previously, 24 percent of Schedule A participants who were eligible to use service took advantage of one or more of the three service channels. In contrast, 29 percent of participants with other scenario types (Standard Deduction, Taxable Social Security, and EITC) who were eligible to receive assistance took advantage of available service.

Participants were most likely to use walk-in assistance, as 58 of the 101 (57 percent) service uses were through this channel. Telephone assistance was used 30 percent of the time, and internet assistance was used 13 percent of the time. Of participants who were eligible to receive assistance but did not, 82 percent cited "Did not need help" as the reason for not using service.

The majority of Schedule A participants (89 percent) who used service rated confidence in the accuracy of the assistance they received as somewhat to very high on an eight-point scale, with 88 percent of these participants rating confidence in accuracy of assistance as very high.

Consistent with results from Taxable Social Security and EIC participants, use of walk-in, telephone, and internet assistance failed to be an indicator of accuracy for Schedule A participants. No significant difference emerged between those participants who used service and those who did not use service (see Table 18).

Table 18: Use of IRS Service Channels and Deduction Accuracy

Scenario Type	IRS Service Used	Percentage with Correct Deduction
Itemized Deduction Participants	Yes	70%
	No	68%
Could Not Itemize Participants	Yes	17%
	No	23%
All Schedule A Participants	Yes	42%
	No	46%

Schedule A and Non-Schedule A Participant Financial Situations

With respect to personal financial situations, participants rated six statements on the debrief form according to an eight-point scale, with 1 indicating "Strongly disagree" and 8 indicating "Strongly agree." Table 19 compares Schedule A participant responses to non-Schedule A participant (Standard Deductions, EITC, and Taxable Social Security participants) responses for statements.

TABLE 19: Comparison of Schedule A and Non-Schedule A Participant Financial Situations

Statement	Percentage of Participants who Somewhat to Strongly Agree*	
	Schedule A Participants	Non-Schedule A Participants
I enjoy managing my household finances.	88%	85%
I usually pay my bills on time.	95%	90%
I usually use an online bill paying service(s).	63%	51%
I wish I had a better understanding of my finances.	36%	39%
I feel confident in my ability to solve financial problems that come up in my everyday life.	92%	88%
I wish I had better English reading and writing skills.	10%	13%

**"Somewhat to Strongly Agree" indicates that these participants marked 5 through 8 on the scale.*

Schedule A participants consistently rated agreement higher than non-Schedule A participants on all dimensions, except for the two statements describing participants' wish for a better understanding of finances and English skills; for both of these statements, Schedule A participants rated agreement lower than non-Schedule A participants. Additionally, for the two statements "I wish I had a better understanding of my finances" and "I wish I had better English reading and writing skills," Schedule A participants' agreement was lowest overall on these in comparison to both groups' agreement on all of the other statements. These results indicate that Schedule A participants in general have higher confidence in their ability to manage financial situations than participants in other scenarios.

Recommendations for Schedule A

Participants with Itemized Deductions scenarios who completed Schedule A and participants with Could Not Itemize scenarios who completed the Standard Deduction Worksheet were significantly more likely to take the correct deduction. This finding is similar to results for participants in other scenarios, in which participants who completed supplemental worksheets applicable to their scenarios were significantly more likely to prepare their returns accurately than those who did not utilize these service resources. Again, these findings highlight the importance of both increasing awareness of these worksheets and of making the worksheets easier to understand.

With respect to the first line of Schedule A, "Medical and dental expenses," 111 participants had a scenario that involved a fictional individual who had medical premiums that were deducted from their paycheck pre-tax; 44 percent incorrectly deducted these premiums again on Schedule A. Due to the lack of clarification relating to medical premiums that have been payroll deducted pre-tax, WIRA recommended adding a bullet

detailing this type of payment under the section "Examples of Medical and Dental Payments You Cannot Deduct" in Appendix A of the 1040 instruction booklet. Further, above line 1 on Schedule A is a "Caution" statement with the description: "Do not include expenses reimbursed or paid by others." WIRA recommended adding a clause to this "Caution" statement that warns against including expenses that have been payroll deducted pre-tax.

Alerting taxpayers with pre-tax medical premiums that they cannot deduct these expenses on Schedule A via a more rigorous outreach plan may or may not be beneficial to the IRS in terms of cost if only small proportion of taxpayers face this scenario. For this reason, WIRA recommended initiating a research study in association with the National Research Program (NRP) with the goal of determining whether the benefits of a pre-tax outreach plan that enables the IRS to prevent this type of noncompliance will outweigh the cost of this outreach plan. By measuring line-item compliance of line 1 on Schedule A, the IRS can better understand the percentage of taxpayers who make the error of incorrectly deducting this expense on Schedule A. Answers for the study's questions can be pre-defined to avoid ambiguous narrative responses.

The ability to increase standard deductions by the state and local real estate taxes paid by up to $500 ($1,000 if married filing jointly) was new for TY 2008. Since awareness was low for this additional deduction, a over half of participants with this fictional situation (54 percent) made the error of not including real estate taxes paid with their deductions. This indicates that the instructions and forms were not effective enough to make taxpayers aware of this new clause and prevent them from incorrectly calculating their standard deduction. Here, an opportunity exists to improve the manner in which new clauses are highlighted in instruction booklets and on tax forms. WIRA therefore recommended that a greater effort be made to highlight the new clauses of the tax law in instruction booklets and on tax forms to effectively alert the relevant segment of taxpayers.

With respect to providing outreach before and during filing season to taxpayers who will likely be affected by new clauses in the tax law, WIRA recommends the development of more rigorous and far-reaching pre-tax communication materials via the Wage & Investment (W&I) Communications & Liaison (C&L) office for intended audiences. Communication material should ideally follow the "what, why, how" marketing communication approach that has been proposed for the Stakeholder Partnerships, Education, & Communications (SPEC) outreach model[30] and should be distributed through multiple channels, such as www.IRS.gov and SPEC for partners to distribute to taxpayers.

Lastly, an implication of the difference between Schedule A and non-Schedule A participants' financial situations is that Schedule A participants have invested more resources to maximize the utility of their finances. Accordingly, these taxpayers may adapt to complex tax issues more readily than taxpayers who are less likely to itemize deductions. Additionally, Schedule A participants were more likely to use an online bill paying service, to have access to the internet, and to use the internet and email more frequently than non-Schedule A participants. Therefore, taxpayers who are more likely to itemize their deductions can potentially be a targeted population for transferring their primary communication channel with the IRS from more traditional, expensive channels, such as mail and telephone, to more cost-effective, interactive channels, such as web platforms on IRS.gov.

Conclusion of Multi City Study

Evaluating Effect of Assistance on Compliance

In assessing the lack of impact that telephone, internet, and walk-in assistance had on accuracy rates for participants in all scenario types, a number of factors emerge as possible explanations.

Completing an accurate return according to the five critical lines on Form 1040 and Form 1040A was dependent on a number of factors, such as use of IRS publications, use of supplemental worksheets that were applicable to each scenario, and correctly calculating lines that required mathematical computation. Because accuracy, evaluated as a single measure, was dependent on multiple factors, measuring the effectiveness of an isolated factor has a diluted impact. This inference carries over to the impact of telephone, internet, and walk-in service on accuracy; since use of each service channel is an isolated variable, the ability to effectively

measure the impact of just one of these variables on accuracy is improbable given the effect of numerous other factors on accuracy.

The nature of the participant's inquiry and his or her application of the assistor's response while using one of the three service channels pose another challenge in evaluating impact of service use. Assuming the inquiry was correctly stated, it cannot be automatically presumed that the participant accurately applied the response. Also, although unlikely, professionally-trained IRS staff may have misunderstood the participant's question, consequently giving misguided information and leading the participant to enter incorrect information on his or her tax return.

Confounding variables, interpretation and application of information, and potentially misguided assistance can likely be mitigated by the use of tax preparation software. Software can mitigate the problem of confounding variables by relating all of the extraneous variables, such as completing supplemental worksheets and mathematical computations, together through a probe-and-response application for the taxpayers. Tax software would also be useful to provide the correct information and application of data that the taxpayer provides, and would thereby eliminate the need for third-party assistance. Again, the probe-and-response feature of the software essentially limits the taxpayer's role to posing questions, such as "Am I eligible?", and supplying the required information to answer the question. Therefore, future research on the impact of service on accuracy should consider the application of tax preparation software to alleviate the difficulties that arise when evaluating the two educational models.

With respect to service channel preference, the majority of Multi City participant initial service uses were through the walk-in channel (62 percent of all service uses), followed by telephone assistance (23 percent), and lastly internet assistance (15 percent). The lower incidence of telephone and internet assistance in the Multi City study can likely be attributed to the convenience of walk-in assistance at the time, for which participants simply walked to an adjacent room and waited for five minutes or less to speak with a research staff member. In fact, results from the Taxpayer Experience Survey showed that 56 percent of individuals from a sample of local IRS office visitors would be somewhat or very likely to continue to wait for service if wait time increased by a half-hour. This number dropped to 43 percent when wait time increased to an hour. Additionally, 53 percent stated they would use a computer or phone at the local IRS office instead of talking with a representative.[31] This percentage is significantly higher than that of Multi City participants who initially used a computer or phone for their service needs (38 percent).

Multi City Participant Confidence in the Accuracy of Tax Returns

Overall, results from the four phases of the Multi City study shed light on the significant challenge of making taxpayers aware that, for more common errors on tax returns, their confidence in accuracy may be misplaced. As stated previously, of all Multi City participants, 88 percent rated confidence in the accuracy of their return as somewhat to very high. See Table 20 for mean confidence scores of participants who scored poorly, moderately, and exceptionally well on their fictional returns.

TABLE 20: Multi City Participant Confidence in Accuracy of Tax Return

Overall Performance on Tax Return	Number of Correct Critical Lines*	Mean Confidence Score**
Poor	Zero or one	5.7
Moderate	Two or three	6.4
Exceptional	Four or five	7.0

**"Critical Lines" refers to the five critical lines that were evaluated to determine overall accuracy of tax returns: adjusted gross income (AGI), taxable income, total tax, total payments, and overpaid/amount owed.*

***Confidence in accuracy of return was rated on an eight-point scale, with 1 indicating "Not at all confident" and 8 indicating "Very confident."*

Participants who scored poorly had significantly lower confidence in the accuracy of their return than participants who scored moderately and exceptionally well. However, these participants' mean confidence score was in the "somewhat confident" range (a rating of 5 or 6 on the eight-point scale), signifying inflated confidence relative to these participants' actual performance.

Research from psychological literature suggests that this type of inflated confidence is also apparent among people in general who assess themselves in other domains, such as their health, education, and careers. According to Dunning, Heath, and Suls (2004), the correlation between people's self-assessment of skills and actual performance in several domains is moderate to meager. The following quote provides insight into this predicament:

> People overrate themselves. On average, people say that they are "above average" in skill (a conclusion that defies statistical probability), overestimate the likelihood that they will engage in desirable behaviors and achieve favorable outcomes … and reach judgments with too much confidence.[32]

To curtail the inflated confidence taxpayers have in the accuracy of their returns, the initiation of preemptive communication can draw taxpayers' attention to areas with high error potential and the increased risk of inaccuracy associated with these errors. Clear options for resolving uncertainty and obtaining accurate information must be provided preemptively via web platforms, telephone communication, media outreach, and publications. Additionally, outreach pilots should be designed to test the value and effectiveness of these new communication approaches and to identify additional ways to continuously improve the model.

Merit of the Multi City Study

The merit of the Multi City study lies in the potential for significant operational change of outreach strategies, forms, and publications based on recommendations from each group of participants: Taxable Social Security, EIC, and Schedule A. Understanding the unique barriers faced by taxpayers with different situations will allow the IRS to increase voluntary compliance for those who self-prepare their tax returns by targeting and alleviating these challenges, thereby improving the preparation experience for the over 56 million taxpayers who prepare their own returns each year.

Endnotes

1 Source: *Study of the Effect of Taxpayer Assistance on Voluntary Compliance*, IRS Management Briefing and Executive Summary. Price Waterhouse, July 28, 1989.

2 Source: *Phase 1: 2009 Multi City Study of the Effect of Assistance on Compliance*, W&I Research & Analysis Group 4, Project Number 3-08-07-S-032T, April 2010.

3 Source: *Phase 2: 2009 Multi City Study of the Effect of Assistance on Compliance—Taxable Social Security*. W&I Research & Analysis Group 4, Project Number 3-08-07-S-032T, June 2010.

4 Source: *Phase 3: 2009 Multi City Study of the Effect of Assistance on Compliance—Earned Income Credit and Child Tax Credit*. W&I Research & Analysis Group 4, Project Number 4-10-09-S-052, September 2010.

5 Source: *Phase 4: 2009 Multi City Study of the Effect of Assistance on Compliance—Schedule A*. W&I Research & Analysis Group 4, Project Number 4-10-09-S-065, December 2010.

6 WIRA acknowledges that the experimental design does not address taxpayer burden when choosing a service channel (i.e., channel preference). The experiment was designed to test the compliance impact of the IRS service channels if all service was equally available (i.e., service usage).

7 Examples of anomalies included participants who said to a research staff member that they had never done their own taxes or used tax preparers, and participants who did not complete the back of their Form 1040 or Form 1040A.

8 Seventy-seven percent of taxpayers who participated in the 2009 Benchmark Survey indicated that they had internet access at home. Source: *2009 Benchmark Survey, Phase 1—Analysis of Generational and Income Segments*, May 2010, page 18.

9 V-coded tax returns are returns that were prepared using tax software, and were subsequently printed and mailed in to the IRS.

[10] For TY 2008, 40 percent of taxpayers self-prepared their returns. Of those who self-prepared, 60 percent filed electronically, 23 percent filed hand-prepared paper forms, and 18 percent filed v-coded paper returns. Source: *Compliance Data Warehouse (CDW) Electronic Tax Administration Marketing Database (ETA MDB).*

[11] Seventeen participants were recorded using service three times, while three participants were recorded using service four times. No participants were recorded using service more than four times.

[12] Results from the Taxpayer Experience Survey showed that 63 percent of respondents who contacted the IRS more than once used the same service channel for multiple contacts. Source: *2009 W&I Taxpayer Experience National Report, Tax Year 2008* (also known as *Market Segment Survey*), July 2010, Slide 14.

[13] Similar to participants with Itemized Deductions scenarios, participants with Could Not Itemize scenarios were given a list of potential deductions that could be itemized. However, if they had correctly completed Schedule A, Could Not Itemize participants would have found their total itemized deduction amount to be less than their standard deduction amount.

[14] On the Taxpayer Experience Survey, respondents who had contacted the IRS in the past 12 months were asked to rate their confidence in the completeness and accuracy of information and assistance received from the IRS. Ninety-one percent of respondents rated confidence as somewhat to very high, with 66 percent of these respondents rating confidence as very high. Source*: 2009 W&I Taxpayer Experience National Report, Tax Year 2008* (also known as *Market Segment Survey*), July 2010, Slide 66.

[15] Exploratory descriptive statistics reveal the existence of demographic effects on accuracy. Applying chi-squared analysis, the variables city, race, education, employment, and scenario type each resulted in a statistically significant difference in accuracy rates. To isolate whether the use of IRS services was a significant predictor of accuracy, these variables were controlled.

[16] Totals more than 100 percent due to rounding.

[17] Source: *Earned Income Tax Credit Participation Rate for Tax Year 2005*, Dean Plueger, Recent Research on Tax Administration and Compliance: Selected Papers Given at the 2009 IRS Research Conference, pages 151-195.

[18] Stakeholder Partnership, Education, and Communications (SPEC) shopping review results through April of 2009 found that 78 percent of scenarios eligible for CTC that were prepared by volunteers had the correct amount of CTC entered on the tax returns. Source: *Ensuring the Quality Assurance Processes Are Consistently Followed Remains a Significant Challenge for the Volunteer Program, Treasury Inspector General for Tax Administration,* September 15, 2009.

[19] SPEC shopping review results through April 2009 show that 78 percent of taxpayers who elected a volunteer to prepare their returns claimed the correct amount of ACTC on their returns. Source: *Ensuring the Quality Assurance Processes Are Consistently Followed Remains a Significant Challenge for the Volunteer Program, Treasury Inspector General for Tax Administration,* September 15, 2009.

[20] Percentages add to 101 percent due to rounding.

[21] With respect to accuracy of CTC and ACTC, differences between groups that did and did not use service could not be statistically analyzed due to the low number of participants who were eligible for these credits and who used service.

[22] Fifty-three percent of participants with scenario Madison, a single fictional individual with no children, failed to claim any amount of EIC on their returns. In contrast, 40 percent and six percent of participants with scenarios Harrison and Hood, respectively, failed to claim any amount of EIC.

[23] For the purpose of this report, Schedule A participants will refer to participants tasked with either an Itemized Deductions or Could Not Itemize scenario. Both scenario types required participants to use Schedule A to determine whether or not they could itemize deductions on their mock tax returns.

[24] Of all Multi City participants, 34 percent indicated to the screener that they had itemized deductions for the previous tax year. This figure is in line with IRS estimates that, on average, one in three taxpayers itemizes their deductions.

[25] Itemized deductions can be taken only on Form 1040. Standard deductions, however, can be taken on either Form 1040 or Form 1040A. For this study, participants with an Itemized Deductions scenario must have used Form 1040 to correctly complete their scenario, while participants with a Could Not Itemize scenario could have used either a 1040 or 1040A to correctly complete their scenario.

[26] Critical lines are the last lines of the eight sections on Schedule A, which contain the sum of deductions from previous lines of the section. Line 29 is the sum of all previous critical lines on Schedule A.

[27] Each of the eight fictional scenarios that required the use of Schedule A involved individuals who paid motor vehicle taxes, and six of the eight scenarios involved individuals who paid home property taxes.

[28] The seven critical lines refer to lines 4, 9, 15, 19, 20, 27, and 28.

[29] The debrief form that was completed after the study did not include questions to determine why or why not participants used worksheets applicable to their scenario.

[30] Source: *SPEC Outreach Study: Final Report*, Project # 4-09-01-S-019, W&I Research & Analysis Group 4, October 2010.

[31] Source: *2009 W&I Taxpayer Experience National Report, Tax Year 2008* (also known as *Market Segment Survey*), July 2010, Slide 13.

[32] Source: *Flawed Self-Assessment: Implications for Health, Education, and the Workplace*, David Dunning, Chip Heath, and Jerry Suls from Cornell University, Stanford University, and University of Iowa, respectively; American Psychological Society, Volume 5, 2004

Appendix A: Screener

[SCRIPT FOR ANSWERING MACHINES]

Hello, this message is for _____, my name is _____, calling from Development Associates on behalf of the Internal Revenue Service regarding the research study in [city name].

I wanted to take a few minutes to ask you some questions to see if you might qualify. Please call me back at 1-800-443-5696. We appreciate your interest in this important study. Thank you.

Goodbye

[WHEN REACH POTENTIAL RESPONDENTS]

1. There is a confirmation code on the front of the postcard near your name and address. It would help me if you could read it to me. The number will start with STL, and be followed by something like G-E-N, I-N-C, or AG-E.

Please use the blank space to write your answers.

Confirm Code (Ex: STL-GEN1-99999-A):

2. May I please have your contact information, including your full legal name?

[INTERVIEWER INSTRUCTION—we need the legal name (no nicknames).]

Please use the blank space to write your answers.

First Name:
Last Name:
Address 1:
Address 2:
City:
State:
Zip:
Phone:
E-Mail:

3. How did you hear about this study [MARK ALL THAT APPLY]?

[INTERVIEWER INSTRUCTION—If you already know the answer, do not ask this question again. Fill in the response from your paper form.]

Please check all that apply and/or add your own variant.

Postcard
Web (Craigslist)
Word of Mouth
Print Advertisement
Radio Advertisement
Flyer
Other (Please Specify):

4. Are you over age 18?

Please pick one of the answers below.

YES
NO

5. What is your age? [Enter -99 if refuses to answer]

Please use the blank space to write your answers.

IF RESPONDENT DOES NOT GIVE EXACT AGE, READ OR PROVIDE THE FOLLOWING: If you would like, I could read some age categories; just stop me when I reach the one closest to your age.

Please pick one of the answers below.

18 to 24
25 to 34
35 to 44
45 to 54
55 to 64
65 to 74
75 to 84
85 and above
REFUSED

6. Gender of respondent

[INTERVIEWER INSTRUCTION: Unless there is some ambiguity, please complete without asking]

[SUGGESTED PHRASING: My apologies for asking the next question, but what is your gender?]

Please pick one of the answers below.

Male
Female

7. This question is about the language spoken in your home. Would you say that English is the primary language.

A language other than English is the primary language.
Refused
Other (Please specify):

8. Are you of Hispanic, Latino or Spanish origin?

Please pick one of the answers below.

Yes
No
Refused

9. I apologize for asking, but what is your race?

Are you: [MARK ALL THAT APPLY]

INTERVIEWER INSTRUCTION: If they answer Hispanic/Latino, please follow-up with "Would you categorize yourself as White, Black, or American Indian?"

Please check all that apply and/or add your own variant.

White or Caucasian
Black or African American
American Indian or Alaska Native
Asian
Native Hawaiian or Other Pacific Islander
Refused
Other (Please specify):

10. Are you currently employed?

Please pick one of the answers below or add your own.

Employed full-time
Employed part-time, not a student
Not employed, but looking for work
Not employed, and not looking for work
Full-time student, not working
Full-time Student, working
Part-time student, working
Part-time student, not working
Retired, working part-time
Retired, not employed
Other (Please specify):

11. I have seven categories for your total annual household income. If you feel comfortable, please stop me when I read the one that best describes your total annual household income.

Please pick one of the answers below or add your own.

Under $17,000
$17,000 to $25,000
$25,000 to $35,000
$35,000 to $50,000
$50,000 to $75,000
$75,000 to $100,000
$100,000 and over
No Income
Refused
Other (Please specify):

12. What is the highest level of education you have completed?

[INTERVIEWER INSTRUCTION—If someone has multiple degrees and you can determine which degree is higher, then mark that category; if not, mark both.]

Please check all that apply and/or add your own variant.

Grade School
Some High School
High School Diploma/GED

Trade/Vocational School Certificate
Some College, No Degree
Associate's Degree
Bachelor's Degree
An advanced degree (INCLUDES MASTERS, DOCTORAL, AND PROFESSIONAL DEGREES)
Other (Please specify):

13. Do you have Internet access at home?

Please pick one of the answers below.

Yes
No
Don't Know

14. Do you use the internet, at least occasionally?

[INTERVIEWER INSTRUCTION—if person asks for clarification, ask them to "just use your best judgment."]

Please pick one of the answers below.

Yes
No

15. Do you send or receive email, at least occasionally?

[INTERVIEWER INSTRUCTION—if person asks for clarification, ask them to "just use your best judgment."]

Please pick one of the answers below.

Yes
No

16. Are you the adult in your household who is most familiar with preparing and filling out your federal income tax return?

[INTERVIEWER INSTRUCTION: Pause for response. Then tell respondent]

Please pick one of the answers below or add your own.

Yes, most familiar
Equally familiar
Neither most familiar nor equally familiar
Other (Please specify):

The computer is matching your responses with our targeted demographic to ensure we have a representative U.S. Population, if you can give me one moment.

17. [INTERVIEWER INSTRUCTION: Does this person still qualify based upon their demographic information?]

Please pick one of the answers below.

Yes
No

All right, the next set of questions are about how you filed your United States federal tax forms in each of the last three years. I'm going to ask you to think back to the last three tax years for which you filed. Most people will file their tax forms for tax year 2008 by April 15th of this year. So, when I say tax year 2008, I mean this year. When I say tax year 2007, I mean last year.

18. What method did you use or will you use to file your taxes this year--tax year 2008? We are interested in knowing whether you used tax software, paper forms, or a tax professional, for example. How about last year—2007? And tax year 2006?

[INTERVIEWER INSTRUCTION—probe to make sure you are able to place them correctly.]

Please fill in the answers in the table below (mark appropriate circles and squares and fill in the blank spaces).

2008 2007 2006

1. Filled out the paper forms without tax preparation software
2. Used tax preparation software, but mailed IRS paper forms
3. Used tax preparation software, and filed electronically
4. Hired a tax professional to check my work and make suggestions
5. Hired a tax professional who took all my tax information and completed my form
6. Not sure
7. Not required to file

19. The next thing we're interested in is your filing status for the last three tax years. I'm going to read some sample filing statuses.

Please let me know which one best describes you for this year—tax year 2008? And 2007? And 2006?

Please fill in the answers in the table below (mark appropriate circles and squares and fill in the blank spaces).

2008 2007 2006

1. Single
2. Single Head of Household
3. Married Filing Jointly
4. Married Filing Separately
5. Qualifying Widow(er) with dependent child
6. Not sure
7. Not required to file

20. Next, I'm going to read off some of the standard forms that are often filed. We'd like to know which forms you used or will use when you file your tax return this year—tax year 2008? And 2007? And 2006?

Please fill in the answers in the table below (mark appropriate circles and squares and fill in the blank spaces).

2008 2007 2006

1. Short form 1040EZ
2. Short form 1040A
3. Long form 1040 without other forms or schedules
4. Long form 1040, with other forms or schedules
5. Some other form
6. Not sure
7. Not required to file

21. You stated that you used or will use the 1040EZ form. Would you be comfortable completing the 1040A Short Form?

[INTERVIEWER INSTRUCTION—If respondent states "I don't know", mark "no"]

Please pick one of the answers below.

Yes
No

22. I have just one additional set of questions about the most recent federal tax return you filed. We will use this information to help us place you into the right study group. I'm going to read a list of items that include things like tax deductions and tax credits. Did your most recent tax return include any of the following?

[INTERVIEWER INSTRUCTION: Please reference q18 to find the respondent's most recently filed tax year.]

Please fill in the answers in the table below (mark appropriate circles and squares and fill in the blank spaces).

	Yes	No	Don't Know
Earned Income Credit (EIC)			
Social Security			
Benefits paid to you that were taxable			
Child Tax Credit			
Itemized Deductions			
Standard Deductions			
Dependents			

The computer is going through your responses to see if you are eligible for the study. If you can bear with me for just a second . . .

23. Congratulations, you are eligible to participate! As you may know, we are assisting the IRS with its research to better understand how taxpayers complete their tax forms and what help people need. The study takes about two hours and you receive an honorarium of $60 for filling in the form and up to an additional $40 based upon the reasonable accuracy of your responses—so potentially $100 total.

Would you like to participate?
Please pick one of the answers below or add your own.

Yes
No
Not Sure (Please specify):

24. The study will take place on DAY, MONTH, at TIME am/pm. Will that work for you?

[INTERVIEWER INSTRUCTION—Offer 5:15 PM sessions as a last resort. Note that the participant needs to show up 15 minutes beforehand.]

[INTERVIEWER INSTRUCTION—when you pick a session, write it down on a piece of paper so that you can read it back to the R at the end of the interview.]

Please pick one of the answers below.

TUES MAY 5—08:30 AM - 10:30 AM (8:15 AM)
TUES MAY 5—11:00 AM - 01:00 PM (10:45 AM)

TUES MAY 5—03:00 PM - 05:00 PM (2:45 PM)
TUES MAY 5—05:15 PM - 07:15 PM (5:00 PM)
WED MAY 6—08:30 AM - 10:30 AM (8:15 AM)
WED MAY 6—11:00 AM - 01:00 PM (10:45 AM)
WED MAY 6—03:00 PM - 05:00 PM (2:45 PM)
WED MAY 6—05:15 PM - 07:15 PM (5:00 PM)
THURS MAY 7—08:30 AM - 10:30 AM (8:15 AM)
THURS MAY 7—11:00 PM - 01:00 PM (10:45 AM)
THURS MAY 7—03:00 PM - 05:00 PM (2:45 PM)
THURS MAY 7—05:15 PM - 7:15 PM (5:00 PM)
No Time will work

25. We will send you a confirmation letter with the place and time of your session. Is your mailing address the same as your contact address?

Please pick one of the answers below.

Yes
No

26. Mailing Information:

Please use the blank space to write your answers.

ADDRESS 1:
ADDRESS 2:
CITY:
STATE:
ZIP:

The tax situation for the family or individual in your scenario will be very much like yours, so it should be like doing taxes for someone like yourself. To participate, you will need to come to your session to complete the hypothetical tax scenario. This session is located at [location]. We will include directions to the session in the materials you will receive in the mail.

[INTERVIEWER INSTRUCTION: If respondent insists on having directions while on the phone, please ask for their method of transportation, and reference the following:

Driving Directions: Plug in their address to either Mapquest or Google Maps, then relay the directions over the phone.

Public Transportation: Metrolink (subway) to location]

Now, let me tell you a little bit about what will happen when you arrive for your session …

First, an employee from Information Experts will greet you in the lobby. He or she will check you in and out, and pay you the honorarium when you are finished. Only he or she will have your name. While you are there, you will be given a badge with a number on it. You will use this number during the study so that your name will not be associated with any of your responses. Your participation will be completely anonymous, and no one at the IRS will ever have your name. When you come, you will need to bring two things with you. One, a government-issued photo ID, such as a driver's license. This is so we can get you into the building. Two, your confirmation letter that we will mail to you one week prior to your session. This letter will contain your participant ID number, so it is important that you bring it with you. We will need these items, both photo ID and confirmation letter, so that we can pay you. We will also remind you to bring these items when we send

the confirmation letter in the mail. If you do not receive your confirmation letter by Monday, May 4th, please give us a call on our 1-800 number.

Also, we want to mention that if you need reading glasses, you will need to remember to bring them. We also want to mention that the temperature of the building in which the session is located varies, so please bring appropriate clothing. In addition, neither the IRS nor Information Experts will provide any childcare for any duration of the study. Please do not bring anything related to your own personal income taxes. Everything you need to complete the hypothetical tax scenario will be ready for you.
Do you have any questions?

We're almost done, but we're required by law to read the following message to you.

The Paperwork Reduction Act requires that the IRS display an OMB control number on all public information requests. The OMB Approval Number for this study is 1545-1349. Also, if you have any comments associated with this study or suggestions on making this process simpler, I have an address you can write to. Would you like that address?

I.R.S.
Tax Products Coordinating Committee
1111 Constitution Ave. NW
Washington, DC 20224

Thank you for participating in this important study!

We look forward to seeing you on [DAY, MONTH, at TIME] am/pm. [END INTERVIEW]

I'm sorry. You are not eligible for this study. We appreciate you taking the time to see if you qualify. Thank you for your interest.

Ok. We appreciate your time. And thank you for your interest.

27. This survey had some issues that need to be clarified / researched.

Please use the blank space to write your answers.

Needs a follow-up call
Problem present, not resolved
Evening session desired, re-contact
Referral Information present

28. Interviewer comments:

Please write your answer in the space below.

29. Primary Interviewer:

Please pick one of the answers below.

[Names of interviewers]

This survey is complete.

Don't forget to press SUBMIT!

Appendix B: Session Instructions

I want to start by thanking you all for coming in today to help evaluate some tax forms. Our goal is to better understand how people complete their tax returns and how the IRS affects that process.

There are several other staff members here who are wearing badges indicating they are project staff. Please refer any questions you have to these staff members only.

When you came in, you each should have received a packet. There are several different packets, so please make sure the number on your badge matches the number on the packet. In your packet you will find a copy of the instructions for your participation in this project. Each packet has a brief description of the person you are going to pretend to be today for the purpose of completing a tax return. Each packet also contains the tax documents for that person, tax forms, and instruction booklets. We have a calculator and some pencils for each of you. If you need more pencils, just let me know.

All the necessary tax information and relevant tax documents for your person are in the written description and/or the packet. So, if your scenario does not mention childcare expenses, you should assume your person did not have any childcare expenses. Some packets may contain information that is not necessary for completing the person's tax return. Use what you need.

One thing we are looking at today is how people use help to complete their tax form. Each of you will have access to a certain type of help, if you need it. The type of help you can get is based on the color on your badge.

- Blue is for in person assistance,
- Red is for web,
- Yellow is for telephone,
- Green is a wild card for any or all of the three: in person assistance, web and telephone; and
- White is for instructions and forms only.

Now I'll explain how to request the different types of services you may be eligible for and what to expect for each type of service.

For In-Person Assistance: You may receive help by talking to an IRS customer service representative face-to-face. We call this person the "walk-in assistor." If you have a question you would like to ask a walk-in assistor about your person's tax return, please tell one of the staff members you would like walk in or face-to-face assistance. Before you see the walk in assistor, you will be given a piece of paper. Please write your participant number and the line number or topic that you want help with. Remember when you talk to the assistor to use your person's name, not yours. Please do not give the assistor your real name.

For Web: If you have a question about your person's taxes, you can use the website, called " IRS.gov." To use the website, please tell the escort you would like to use the website. At each computer you will find a piece of paper where you will need to enter your participant number and the line number or topic that you want help with. Please do not visit any other websites. A staff member will be at the computers to help you get started and to help you with the computers. Remember this person cannot answer any tax questions. The computers you will be using are set up to record where you go within IRS.gov to help us better understand how people navigate the IRS.gov website.

For Telephone: You can get help by calling the IRS toll-free line. If you have a question you would like to ask the IRS toll-free line, please tell one of the staff members you would like toll-free assistance. At each telephone you will find a piece of paper with the IRS toll-free assistance number on it. Please write your participant number on that piece of paper and the line number or topic that you want help with. A staff member will be

at the phones to help you complete this form and call the IRS toll-free line. When you call, the assistor should give you his or her ID number and maybe their last name. Please write the assistors ID number and last name on the sheet of paper also. Remember when you call to use the name of your person. Please do not give the assistor your real name. Also, do not mention that you are participating in a research study.

For Instructions and Forms: Your packet contains IRS forms and instructions that have been included to provide you with assistance in completing your person's tax return. Feel free to write on the forms and in the instruction booklets.

When you have completed the tax form, please tell one of the staff members that you are finished. You will receive a short questionnaire to give us some feedback on your experience with your person's taxes and you will also receive your honorarium. While you are completing the questionnaire, we will score your work to see how you did and determine if you earned the bonus.

With your packet, you were also given a consent form. Please take a few minutes now to read and sign the consent form. Let me know when you are done.

Finally, we have just a few rules.

1. If you need to use the restroom, just let one of the staff members know.
2. Please turn off all cell phones.
3. Please do not talk to the other participants.
4. If you have the option of having the IRS calculate a line on the form, do NOT exercise that option. We want YOU to complete the entire form.

Any questions?

Let's begin.

Appendix C: Multi City Study Debrief Survey

1040 Tax Forms Study Debriefing Questions

Please mark your answer to each question.

1. Were you able to complete your tax scenario?

_____ Yes _____ No Why not?

2. Did the packet give you enough information to complete the tax form?

_____ Yes _____ No Why not?

3. Did you use the IRS publications or instruction booklets that came with the tax forms?

_____ Yes **[Go to Question 4]** _____ No **[Skip to Question 5]**

4. Please rate your level of satisfaction or dissatisfaction with the IRS publications or instruction booklets you used today. On a scale of 1 to 8, where 1 is very dissatisfied and 8 is very satisfied, how satisfied or dissatisfied were you with each of the following?

Circle One Answer for Each Statement									
		Very dissatisfied							Very satisfied
a	Ease of finding answers	1	2	3	4	5	6	7	8
b	Ease of understanding instructions	1	2	3	4	5	6	7	8
c	Completeness of instructions	1	2	3	4	5	6	7	8
d	The explanation of what procedures are necessary to complete the forms	1	2	3	4	5	6	7	8

—Over—

5. Did you use any help other than the IRS publications or instruction booklets?

_____ Yes _____ No → Why not? **[Check All That Apply]**

a. ___ Wait time too long/too many other people in line

b. ___ Did not know how
c. ___ Did not need help → **[Skip to Question 8b]**
d. ___ Nothing else was available to me → **[Skip to Question 8b]**
f. ___ Other _____

6. Did you get help more than once?

_____ Yes _____ No

7. What question or issue caused you to seek help from something other than the IRS publications and instruction booklets?

8. On a scale of 1 to 8, where 1 is not at all confident and 8 is very confident, how confident are you with the **accuracy** of:

		Circle One Answer							
		Not at all confident							Very confident
a	The help you received	1	2	3	4	5	6	7	8
b	**The tax return you completed**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**

Thank you for answering the previous questions about the tax scenario you just completed. We would now like you to tell us a bit about yourself.

9. There are many **IRS** resources and services people can use to help them complete their taxes. Which of the following **IRS** resources and services did you use to complete **your tax return this year and the previous two years?**

[Check All That Apply]

IRS Services	**Used this year**	**Used during previous two years**
IRS forms and instruction booklets		
IRS website (www.irs.gov)		
IRS Tax Assistance Centers (walk-in sites)		
Automated IRS phone system		
IRS phone representatives		
E-mail with the IRS		
Written correspondence with the **IRS** (other than e-mail)		
I did not use any IRS resources or services.		

10. There are many **other** resources and services people can use to help them complete their taxes. Which of the following other resources and services did you use to complete **your tax return this year and the previous two years?**

[Check All That Apply]

Other Services	Used this year	Used during previous two years
Volunteer tax preparation clinics		
Tax preparation company		
Non-IRS books and publications		
Your own accountant or bookkeeper		
Tax preparation software		
Internet websites (other than the IRS website)		
I did not use any other resources or services.		

- Over -

11. Other than the types of IRS assistance listed above, are there any other types of assistance you think the IRS should offer taxpayers?

_____

12. Have you ever filed your federal tax return electronically?

_____ Yes _____ No → Why not?

13. Did you hire a professional tax preparer or did you prepare your taxes yourself this year?

_____ Professional tax preparer _____ Self-prepare

14. Why did you choose to hire a professional tax preparer or to self-prepare?

_____

15. The following are some general statements about the way you might feel about your own personal situation. On a scale of 1 to 8, where 1 is strongly disagree and 8 is strongly agree, to what extent do you agree with each of the following statements?

Circle One Answer for Each Statement									
		Strongly disagree							Strongly agree
a	I enjoy managing my household finances.	1	2	3	4	5	6	7	8
b	I usually pay my bills on time.	1	2	3	4	5	6	7	8
c	I usually use an online bill paying service(s).	1	2	3	4	5	6	7	8
d	I wish I had a better understanding of my finances.	1	2	3	4	5	6	7	8
e	I feel confident in my ability to solve financial problems that come up in my everyday life.	1	2	3	4	5	6	7	8
f	I wish I had better English reading and writing skills.	1	2	3	4	5	6	7	8

Thank you for helping with this important study.

Your participation will help the IRS improve its tax forms and instructions.

Please see an IE employee to check out and receive your honorarium.

Appendix D: Testing Locations

The specific locations and dates for the testing sessions were as follows:

Atlanta, GA Date: March 2-6, 2009	Cambridge Building 2965 Flowers Road South Atlanta, GA 30341
St. Louis, MO Date: May 4-8, 2009	Robert A. Young Federal Building 1222 Spruce Street St. Louis, MO 63103
Boston, MA Date: June 8-12, 2009	JFK High Rise 15 New Sudbury Street Boston, MA 02203
Seattle, WA Date: July 6-10, 2009	Henry Jackson Federal Building 915 Second Avenue Seattle, WA 98174

Appendix E: Multi City Scenarios

Name: Cameron Adams
DOB: 08/15/1952
SS#: 886-00-4805
Employment: Teacher Assistant
Marital Status: Legally separated
Spouse's name (if any): Mackenzie Hood

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Mackenzie Adams	813-00-5076	03/05/1955	Wife
Isabel Adams	834-00-7406	05/03/1990	Daughter
Joshua Adams	834-00-4638	02/15/1992	Son

You work as a teacher assistant at your local elementary school. Even though you and your wife, Mackenzie, are legally married, due to previous financial problems, you refuse to file a joint tax return with your wife. Your children, Isabel and Joshua, live with you full time in the family home and you provide over 50% of their support. You elected to receive advance EIC this year. You and your children lived in the state of Georgia for the entire year and are U.S. Citizens. Georgia has a state income tax.

Name: Sally Brown
DOB: 09/12/1935
SS#: 876-00-4532
Employment: Retired
Marital Status: Single
Spouse's name (if any): None

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year: NONE

You are retired. You receive income from pensions and social security and you do not work. You have lived in the state of Georgia in the same house for the last 20 years. You are a U.S. Citizen. Georgia has a state income tax.

Name: Ella Chapman
DOB: 11/22/1969
SS#: 827-00-1774
Employment: Marketing Manager
Marital Status: Married
Spouse's name (if any): Derrick Chapman

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Derrick Chapman	842-00-6518	01/14/1963	Husband
Jasmine Chapman	827-00-1667	11/03/1988	Daughter
Ryan Chapman	827-00-1370	06/15/2001	Son

You are a marketing manager for a local company, your husband, Derrick, is a social worker. You have a savings account at your credit union which earns interest.

Your daughter, Jasmine, is attending the state college on a full tax-free scholarship. Jasmine lives at home and you provide all of her support. Your son, Ryan, also lives with you full time and you provide all of his support. You and your family have lived in the state of Georgia for 7 years. Georgia has a state income tax. You and your family are U.S. Citizens. Last year you received the full amount of the economic stimulus payment for your filing status and number of children.

You did not itemize your deductions last year, but you think you might save money if you itemize this year. The following is a summary of the applicable receipts and records you collected for itemizing your deductions:

Medical Expenses	$1,750
Real Estate Taxes on your primary residence	$2,846.73
Motor Vehicle Tax (value based)	$92.19
Home Mortgage Interest	See 1098
Charitable contributions:	
Checks to local charity (you have a statement from the charity)	$420
Clothing donation receipts fair market value	$156
State income tax	See w-2

Name: Sierra Fillmore
DOB: 10/22/1961
SS#: 823-00-7621
Employment: Head Teller
Marital Status: Married
Spouse's name (if any): Jonathan Fillmore

Name	SS#	DOB	Relationship
Jonathan Fillmore	891-00-3750	01/06/1958	Husband

You are a head teller for a local bank and your husband, Jonathan, is a part-time waiter. You have a savings account at your local bank that earns interest.

You and your husband both lived in the state of Georgia for the entire tax year and are U.S. Citizens, but you have not lived together for the past year and you have not supported your husband in the past year. You and Jonathan have decided to file your taxes separately this year. Jonathan has not filed his taxes yet, but will claim himself when he does. Georgia has a state income tax.

Name: Dana Grant
DOB: 05/26/1949
SS#: 840-00-9570
Employment: Management Assistant
Marital Status: Single

You are single and have no children. You are a management assistant for a vice president of a national paper company. You have a savings account at your local bank that earns interest.

You lived in the state of Georgia for the past 10 years. You are a U.S. Citizen. Georgia has a state income tax.

Name: Jeremy Green
DOB: 11/20/1934
SS#: 832-00-9384
Employment: Retired
Marital Status: Married
Spouse's name (if any): Jessica Green

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Jessica Green	895-00-5193	05/13/1939	Wife

You are a retired autoworker and you receive income from your pension, your social security, and interest on your savings account. You are married to Jessica Green. You are both U.S. citizens and you live in the United States. You have no children. Because of financial problems caused by your first wife, you refuse to file a joint return with anyone, including Jessica. You will file your own tax return and claim an exemption for yourself. Jessica will file her own tax return and claim an exemption for herself. You are a resident of the state of Georgia which has a state income tax and a U.S. Citizen.

Name: Steve J. Harrison
Employment: Service Technician
DOB: 10/18/1966
SS#: 809-00-9273
Marital Status: Married
Spouse's name (if any): Ruth Harrison

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Ruth Harrison	811-00-7881	06/28/1964	Spouse
Lydia Rogers	845-00-2630	05/03/1990	Daughter
Jacob Harrison	824-00-3450	02/15/1992	Son

You are a service technician for the local telephone company and your wife Ruth works part time at the book store. You have a small savings account at your credit union that earns interest. Lydia, Ruth's daughter, lives with you fill time and you provide all of her support. Jacob, your son, also lives with you full time in the family home and you provide all of his support. You, your children and your spouse are all Georgia residents and U.S. Citizens. Georgia has a state income tax.

Name: David Hayes
DOB: 05/26/1949
SS#: 840-00-6874
Employment: Claims Adjuster
Marital Status: Single

You are single and have no children. You are a claims adjuster for a national insurance company. You have a savings account at your local bank that earns interest.

You lived in the state of Georgia for the past 10 years. You are a U.S. Citizen. Georgia has a state income tax. You did not itemize your deductions last year, but think you maybe able to this year. The following is a summary of the applicable receipts and records you collected for itemizing your deductions

Medical and Dental Expenses	$985
Medical Premiums (deducted from your pay check pretax)	$1,872
Motor Vehicle Tax (value based)	$105
Home Property Tax	$609
Mortgage Interest and Points	See 1098
State income tax	See W-2

Name: Mary J. Hood
DOB: 12/12/1966
SS#: 895-00-9015
Employment: Operator
Marital Status: Divorced
Spouse's name (if any): None

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Lauren Salem	824-00-3571	05/03/1991	Daughter
William Hood	816-00-2643	02/15/1993	Son

You are employed as an operator at Bluefield Telecommunications, and this is your only source of income. Both of your children, Lauren and William, lived with you full time in the family home for the entire tax year. You are divorced and provide all of your children's support. You and your children lived in the state of Georgia all year and are U.S. Citizens. Georgia has a state income tax.

Name: Emily Jackson
DOB: 08/19/1974
SS#: 867-00-4371
Employment: Contracts Administrator
Marital Status: Single

You are single and have no children. You are a contracts administrator for a national transportation company. You have a savings account that earns interest at your local bank.

You lived in the state of Georgia for the entire tax year and are U.S. Citizens. Georgia has a state income tax. You have had long distance on your home phone for at least the last 10 years.

You did not itemize your deductions last year, but think you might be able to this year because you bought a condo for your primary residence. The following is a summary of the applicable receipts and records you collected for itemizing your deductions:

Medical and Dental Expenses	$1,530
Medical Premiums (not payroll deducted pre-tax)	$1,872
Motor Vehicle Tax (value based)	$225
Home Property Tax	$709
Mortgage Interest and points	See 1098
Charitable contributions:	
Checks to local charity (you have receipts)	$ 575
State income tax	See W-2

Name: Ashley Madison
DOB: 02/16/1978
SS#: 898-00-4308
Employment: Student and Waitress
Marital Status: Single
Spouse's name (if any): None

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year: None

You are a full-time student at Georgia State and receive tax-free scholarships to pay for your school and living expenses. You work part-time as a waitress for extra money. Your checking account with the credit union earns interest. You do not live with your parents and they do not provide any of your support. You lived in the state of Georgia for the entire year and are a U.S. Citizen. Georgia has a state income tax.

Name: Troy H. McCook
DOB: 03/12/1935
SS#: 876-00-6251
Employment: Retired
Marital Status: Married
Spouse's name (if any): Yvonne A. Smith

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Yvonne A. Smith	853-00-2894	10/30/1940	Spouse
Ashley Fergus	867-00-7521	04/05/1995	Granddaughter

You and your wife, Yvonne, are both retired. You and your wife receive income from pensions and social security. Your granddaughter moved in with you in May of 2005 and you provide all of her support. You and your family lived in the state of Georgia for the entire year and are U.S. Citizens. Georgia has a state income tax.

Name: Kevin Pierce
DOB: 06/28/1967
SS#: 861-00-2460
Employment: Property Manager
Marital Status: Single (divorced)
Spouse's name (if any): None

Name	SS#	DOB	Relationship
Jonathan Pierce	861-00-0657	05/11/1990	Son

You are a property manager for a local condominium community. You have a savings account at your bank that earns interest.

You are divorced and have full custody of your son who lives with full time. Your wife does not pay you child support or alimony. You and your son both lived in the state of Georgia for the entire tax year and are U.S. Citizens. Georgia has a state income tax.

You did not itemize your deductions last year. The following is a summary of the applicable receipts and records you collected for itemizing your deductions.

Medical and Dental Expenses	$1,970
Medical Premiums (not payroll deducted pre-tax)	$2,640
Prescription eyewear	$320
Motor Vehicle Tax (value based)	$270.62
Real Estate Taxes (value based)	$1,500
Mortgage Interest	See 1098
State income tax	See W-2

Name: Nicholas Polk
DOB: 04/16/1969
SS#: 817-00-4376
Employment: Firefighter
Marital Status: Single (divorced)
Spouse's name (if any): None

Name	SS#	DOB	Relationship
Megan Polk	817-00-9670	09/29/1990	Daughter

You are a firefighter for your local county. You have a savings account at your local bank which earns interest. You are divorced and have full custody of your daughter who lived with you for the entire year of 2006. Your wife does not pay you child support or alimony. You and your daughter have lived in the state of Georgia for the past 5 years and are U.S. Citizens. Georgia has a state income tax.

You did not itemize your deductions last year, but you think you may be able to itemize this year. The following is a summary of the applicable receipts and records you collected for itemizing your deductions:

Medical and Dental Expenses	$843
Medical Premiums (deducted from your paycheck pre-tax)	$2,640
Prescription eyewear	$120
Motor Vehicle Tax (value based)	$298
Real Estate Taxes (value based)	$659
Mortgage Interest	See 1098
State income tax	See W-2

Name: Everett Taft
DOB: 11/22/1969
SS#: 827-00-6978
Employment: Warehouse Manager
Marital Status: Married

Spouse's name (if any): Sasha Taft
People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Sasha Taft	842-00-6500	01/14/1963	Wife
Arabella Taft	827-00-1668	11/03/1988	Daughter
Lucas Taft	827-00-1371	06/15/2001	Son

You are a warehouse manager for a local company, your wife, Sasha, is a pastry chef. You have a savings account at your credit union which earns interest.

Your daughter, Arabella, is attending the state college on a full tax-free scholarship. Arabella lives at home and you provide all of her support. Your son, Lucas, also lives with you full time and you provide all of his support. You and your family have lived in the state of Georgia for the 7 years. Georgia has a state income tax. You and your family are U.S. Citizens.

Name: Andrew Thornton
DOB: 03/12/1935
SS#: 876-00-6987
Employment: Retired
Marital Status: Single
Spouse's name (if any): None

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Elizabeth Monroe	876-00-2001	04/05/1995	Granddaughter

You are retired. You receive income from pensions and social security. Your granddaughter moved in with you in May of 2005 and you provide all of her support. You have lived in the state of Georgia for the past 40 years and are a U.S. Citizen. Georgia has a state income tax.

Name: Parker Truman
DOB: 09/17/1970
SS#: 849-00-9507
Employment: Assistant Manager
Marital Status: Married
Spouse's name (if any): Jessica Truman

Name	SS#	DOB	Relationship
Jessica Truman	882-00-6671	06/23/1965	Wife

You are an assistant manager for a grocery store and your wife, Jessica, is a human resources manger. You have a savings account at your credit union which earns interest.

You and your wife both lived in the state of Georgia, for the entire tax year and are U.S. Citizens, but you have not lived together for the past year and you have not supported her. Jessica has already filed her taxes she claimed herself and itemized her deductions. Georgia has a state income tax. Last year you received the full amount of the economic stimulus payment for your filing status and number of children.

You did not itemize your deductions last year. The following is a summary of the applicable receipts and records you collected for itemizing your deductions:

Medical and Dental Expenses	$2,350
Medical Premiums (not payroll deducted pre-tax)	$2,448
Prescription eyewear	$320
Motor Vehicle Tax (value based)	$309
Charitable contributions:	
Checks to local charity (you have a statement from the charity)	$875
State income tax	See W-2

Name: Carla Tyler
DOB: 10/22/1959
SS#: 823-00-1267
Employment: Loan Processor
Marital Status: Married
Spouse's name (if any): Brent Tyler

Name	SS#	DOB	Relationship
Brent Tyler	891-00-3746	01/06/1956	Husband

You are a loan processor for a local bank and your husband, Brent, is a part-time waiter. You have a savings account at your local bank that earns interest.

You and your husband both lived in the state of Georgia for the entire tax year and are U.S. Citizens, but you have not lived together for the past year and you have not supported your husband in the past year. You and Brent have decided to file your taxes separately this year. Brent has not filed his taxes yet, but will claim himself when he does. Georgia has a state income tax. Last year you received the full amount of the economic stimulus payment for your filing status.

You did not itemize your deductions last year, but think you may be able to this year. The following is a summary of the applicable receipts and records you collected for itemizing your deductions:

Medical and Dental Expenses	$1,475
Medical Premiums (not payroll deducted pre-tax)	$936
Motor Vehicle Tax (value based)	$187.26
Charitable contributions:	
Checks to local charity (you have a statement from the charity)	$450
State income tax	See W-2

Name: John T. Washington
DOB: 11/22/1965
SS#: 837-00-5631
Employment: Plumber
Marital Status: Married
Spouse's name (if any): Brenda Washington

People who lived in the house with you and anyone living outside of your home that you or your spouse (if any) supported during the tax year:

Name	SS#	DOB	Relationship
Brenda Washington	813-00-4537	01/14/1967	Wife
Lydia Washington	805-00-1379	11/03/1987	Daughter
Bryce Washington	805-00-5136	06/15/2000	Son

You are a plumber for a local company, your wife Brenda is a stay at home wife. You have a checking account at your credit union and your wife has a money market account at a local bank, both accounts earn interest.

Your daughter, Lydia, is attending the local community college on a full tax-free scholarship. Lydia lives at home and you provide all of her support. Your son, Bryce, also lives with you full time and you provide all of his support. You, your wife and your children lived in the state of Georgia for the entire tax year and are U.S. Citizens. Georgia has a state income tax. Last year you received the full amount of the economic stimulus payment for your filing status and number of children.

You did not itemize your deductions last year, but you think you will save money if you itemize this year. The following is a summary of the applicable receipts and records you collected for itemizing your deductions:

Medical Expenses	$1,750
Real Estate Taxes on your primary residence	$3,750.69
Motor Vehicle Tax (value based)	$92.19
Home Mortgage Interest	See 1098
Charitable contributions:	
Checks to local charity (you have a statement from the charity)	$360
Clothing donation receipts fair market value	$327
State income tax	See W-2

Name: Russell Wilson
DOB: 04/16/1969
SS#: 817-00-6734
Employment: Recruiter
Marital Status: Single (divorced)
Spouse's name (if any): None

Name	SS#	DOB	Relationship
Robin Wilson	817-00-9680	09/29/1990	Daughter

You are a recruiter for a local company. You have a savings account at your local bank which earns interest. You are divorced and have full custody of your daughter who lived with you for the entire year of 2006. Your wife does not pay you child support or alimony. You and your daughter have lived in the state of Georgia for the past 5 years and are U.S. Citizens. Georgia has a state income tax.

Appendix F: Effect of Service on Return Accuracy

Overall Effect of Assistance on Compliance for All Scenario Types *			
	Service Used	**Service Not Used**	**Total**
Correct Return **	47	207	254
Incorrect Return	170	603	773
Total	217	810	1,027
p = .14 ***			

** Fisher's exact tests were conducted to determine the effect of service use through contingency tables.*

*** A tax return with all five critical lines (AGI, Taxable Income, Total Tax, Total Payments, and Overpaid / Amount Owed) correct.*

**** Difference was considered significant if p < .05.*

Analysis of Absolute Error for Service Used vs. Service Not Used for All Scenario Types *			
		Service Used	**Service Not Used**
AGI	*N* **	215	808
	Average Absolute Error	$1,045	$1,169
	Variance	$4,149	$4,675
	Sig.	*p = .53* ***	
Taxable Income	*N*	215	801
	Average Absolute Error	$3,203	$2,963
	Variance	$6,324	$5,749
	Sig.	*p = .38*	
Total Tax	*N*	198	730
	Average Absolute Error	$885	$1,185
	Variance	$2,088	$3,808
	Sig.	*p = .09*	
Total Payments	*N*	217	808
	Average Absolute Error	$298	$276
	Variance	$803	$861
	Sig.	*p = .64*	
Overpaid / Amount Owed	*N*	171	613
	Average Absolute Error	$748	$766
	Variance	$1,742	$1,936
	Sig.	p = .87	

** Independent samples t-tests were conducted to test the difference between absolute error for participants who used any service type, versus those who did not use any service type, regardless of eligibility for service (walk-in only, telephone only, internet only, all service types, or no service type).*

*** Number of participants with an entry for this line. Frequencies vary within groups due to participants leaving certain lines empty on their returns.*

**** Differences were considered significant if p < .05.*

Appendix G: 2008 Social Security Benefits Worksheet

Form 1040—Lines 20a and 20b

Social Security Benefits Worksheet—Lines 20a and 20b

Keep for Your Records

Before you begin:

- ✓ Complete Form 1040, lines 21 and 23 through 32, if they apply to you.
- ✓ Figure any write-in adjustments to be entered on the dotted line next to line 36 (see the instructions for line 36 on page 34).
- ✓ If you are married filing separately and you lived apart from your spouse for all of 2008, enter "D" to the right of the word "benefits" on line 20a. If you do not, you may get a math error notice from the IRS.
- ✓ Be sure you have read the **Exception** on page 26 to see if you can use this worksheet instead of a publication to find out if any of your benefits are taxable.

1. Enter the total amount from **box 5** of **all** your **Forms SSA-1099** and **Forms RRB-1099.** Also, enter this amount on Form 1040, line 20a . . . . . **1.** _____

2. Enter one-half of line 1 . . . . . **2.** _____

3. Enter the total of the amounts from Form 1040, lines 7, 8a, 9a, 10 through 14, 15b, 16b, 17 through 19, and 21 . . . . . **3.** _____

4. Enter the amount, if any, from Form 1040, line 8b . . . . . **4.** _____

5. Add lines 2, 3, and 4 . . . . . **5.** _____

6. Enter the total of the amounts from Form 1040, lines 23 through 32, plus any write-in adjustments you entered on the dotted line next to line 36 . . . . . **6.** _____

7. Is the amount on line 6 less than the amount on line 5?

☐ **No.** STOP None of your social security benefits are taxable. Enter -0- on Form 1040, line 20b.

☐ **Yes.** Subtract line 6 from line 5 . . . . . **7.** _____

8. If you are:

- Married filing jointly, enter $32,000
- Single, head of household, qualifying widow(er), or married filing separately and you **lived apart** from your spouse for all of 2008, enter $25,000
- Married filing separately and you lived with your spouse at any time in 2008, skip lines 8 through 15; multiply line 7 by 85% (.85) and enter the result on line 16. Then go to line 17

. . . . . **8.** _____

9. Is the amount on line 8 less than the amount on line 7?

☐ **No.** STOP None of your social security benefits are taxable. Enter -0- on Form 1040, line 20b. If you are married filing separately and you **lived apart** from your spouse for all of 2008, be sure you entered "D" to the right of the word "benefits" on line 20a.

☐ **Yes.** Subtract line 8 from line 7 . . . . . **9.** _____

10. Enter: $12,000 if married filing jointly; $9,000 if single, head of household, qualifying widow(er), or married filing separately and you **lived apart** from your spouse for all of 2008 . . **10.** _____

11. Subtract line 10 from line 9. If zero or less, enter -0- . . . . . **11.** _____

12. Enter the **smaller** of line 9 or line 10 . . . . . **12.** _____

13. Enter one-half of line 12 . . . . . **13.** _____

14. Enter the **smaller** of line 2 or line 13 . . . . . **14.** _____

15. Multiply line 11 by 85% (.85). If line 11 is zero, enter -0- . . . . . **15.** _____

16. Add lines 14 and 15 . . . . . **16.** _____

17. Multiply line 1 by 85% (.85) . . . . . **17.** _____

18. **Taxable social security benefits.** Enter the **smaller** of line 16 or line 17. Also enter this amount on Form 1040, line 20b . . . . . **18.** _____

TIP *If any of your benefits are taxable for 2008* ***and*** *they include a lump-sum benefit payment that was for an earlier year, you may be able to reduce the taxable amount. See Pub. 915 for details.*

Appendix H: 2008 EIC Instructions

Form 1040—Lines 64a and 64b

Lines 64a and 64b—Earned Income Credit (EIC)

What Is the EIC?

The EIC is a credit for certain people who work. The credit may give you a refund even if you do not owe any tax.

TIP *Special rules may apply for people who had to relocate because of the storms, tornadoes, or flooding in a Midwestern disaster area. For details, see Pub. 4492-B.*

To Take the EIC:

- Follow the steps below.
- Complete the worksheet that applies to you or let the IRS figure the credit for you.
- If you have a qualifying child, complete and attach Schedule EIC.

For help in determining if you are eligible for the EIC, go to *www.irs.gov/eitc* and click on "EITC Assistant." This service is available in English and Spanish.

CAUTION *If you take the EIC even though you are not eligible and it is determined that your error is due to reckless or intentional disregard of the EIC rules, you will not be allowed to take the credit for 2 years even if you are otherwise eligible to do so. If you fraudulently take the EIC, you will not be allowed to take the credit for 10 years. See* Form 8862, who must file, *that begins on page 50. You may also have to pay penalties.*

Step 1 All Filers

1. If, in 2009:
   - 3 or more children lived with you, is the amount on Form 1040, line 38, less than $43,279 ($48,279 if married filing jointly)?
   - 2 children lived with you, is the amount on Form 1040, line 38, less than $40,295 ($45,295 if married filing jointly)?
   - 1 child lived with you, is the amount on Form 1040, line 38, less than $35,463 ($40,463 if married filing jointly)?
   - No children lived with you, is the amount on Form 1040, line 38, less than $13,440 ($18,440 if married filing jointly)?

   ☐ **Yes.** Continue ☐ **No.** STOP You cannot take the credit.

2. Do you, and your spouse if filing a joint return, have a social security number that allows you to work or is valid for EIC purposes (see page 51)?

   ☐ **Yes.** Continue ☐ **No.** STOP You cannot take the credit. Enter "No" on the dotted line next to line 64a.

3. Is your filing status married filing separately?

   ☐ **Yes.** STOP You cannot take the credit. ☐ **No.** Continue

4. Are you filing Form 2555 or 2555-EZ (relating to foreign earned income)?

   ☐ **Yes.** STOP You cannot take the credit. ☐ **No.** Continue

5. Were you or your spouse a nonresident alien for any part of 2009?

   ☐ **Yes.** See *Nonresident aliens* on page 51. ☐ **No.** Go to Step 2.

Step 2 Investment Income

1. Add the amounts from Form 1040:

   Line 8a _____
   Line 8b + _____
   Line 9a + _____
   Line 13* + _____

   Investment Income = ☐

   *If line 13 is a loss, enter -0-.

2. Is your investment income more than $3,100?

   ☐ **Yes.** Continue ☐ **No.** Skip question 3; go to question 4.

3. Are you filing Form 4797 (relating to sales of business property)?

   ☐ **Yes.** See *Form 4797 filers* on page 50. ☐ **No.** STOP You cannot take the credit.

4. Do any of the following apply for 2009?
   - You are filing Schedule E.
   - You are a member of a qualified joint venture that is a passive activity reporting rental real estate income not subject to self-employment tax on Schedule C or C-EZ.
   - You are reporting income from the rental of personal property not used in a trade or business.
   - You are reporting income on Form 1040, line 21, from Form 8814 (relating to election to report child's interest and dividends).

   ☐ **Yes.** You must use Worksheet 1 in Pub. 596 to see if you can take the credit. ☐ **No.** Go to Step 3.

Continued from page 48

Step 3 Qualifying Child

A qualifying child for the EIC is a child who is your...

Son, daughter, stepchild, foster child, brother, sister, stepbrother, stepsister, or a descendant of any of them (for example, your grandchild, niece, or nephew)

AND

was ...

Under age 19 at the end of 2009 and younger than you (or your spouse, if filing jointly)

or

Under age 24 at the end of 2009, a student (see page 51), and younger than you (or your spouse, if filing jointly)

or

Any age and permanently and totally disabled (see page 51)

AND

Who is not filing a joint return for 2009 (or is filing a joint return for 2009 only as a claim for refund)

AND

Who lived with you in the United States for more than half of 2009.
If the child did not live with you for the required time, see *Exception to time lived with you* on page 50.

CAUTION *If the child meets the conditions to be a qualifying child of any other person (other than your spouse if filing a joint return) for 2009, or the child was married, see page 51.*

1. Do you have at least one child who meets the conditions to be your qualifying child?
   - ☐ **Yes.** The child must have a valid social security number (SSN) as defined on page 51 unless the child was born and died in 2009. If at least one qualifying child has a valid SSN (or was born or died in 2009), go to question 2. Otherwise, you cannot take the credit.
   - ☐ **No.** Skip question 2; go to Step 4.

2. Could you, or your spouse if filing a joint return, be a qualifying child of another person in 2009?
   - ☐ **Yes.** STOP You cannot take the credit. Enter "No" on the dotted line next to line 64a.
   - ☐ **No.** Skip Step 4; go to Step 5 on page 50.

Step 4 Filers Without a Qualifying Child

1. Is the amount on Form 1040, line 38, less than $13,440 ($18,440 if married filing jointly)?
   - ☐ **Yes.** Continue
   - ☐ **No.** STOP You cannot take the credit.

2. Could you, or your spouse if filing a joint return, be a qualifying child of another person in 2009?
   - ☐ **Yes.** STOP You cannot take the credit. Enter "No" on the dotted line next to line 64a.
   - ☐ **No.** Continue

3. Can you, or your spouse if filing a joint return, be claimed as a dependent on someone else's 2009 tax return?
   - ☐ **Yes.** STOP You cannot take the credit.
   - ☐ **No.** Continue

4. Were you, or your spouse if filing a joint return, at least age 25 but under age 65 at the end of 2009? If your spouse died in 2009, see Pub. 596 before you answer.
   - ☐ **Yes.** Continue
   - ☐ **No.** STOP You cannot take the credit.

5. Was your home, and your spouse's if filing a joint return, in the United States for more than half of 2009? Members of the military stationed outside the United States, see page 51 before you answer.
   - ☐ **Yes.** Go to Step 5 on page 50.
   - ☐ **No.** STOP You cannot take the credit. Enter "No" on the dotted line next to line 64a.

Need more information or forms? See page 96.

Form 1040—Lines 64a and 64b

Continued from page 49

Step 5 Earned Income

1. Are you filing Schedule SE because you were a member of the clergy or you had church employee income of $108.28 or more?

   ☐ **Yes.** See *Clergy* or *Church employees*, whichever applies, on this page. ☐ **No.** Continue

2. Figure earned income:

   Form 1040, line 7 _____

   Subtract, if included on line 7, any:
   - Taxable scholarship or fellowship grant not reported on a Form W-2.
   - Amount received for work performed while an inmate in a penal institution (enter "PRI" and the amount subtracted on the dotted line next to Form 1040, line 7).
   - Amount received as a pension or annuity from a nonqualified deferred compensation plan or a nongovernmental section 457 plan (enter "DFC" and the amount subtracted on the dotted line next to Form 1040, line 7). This amount may be shown in box 11 of Form W-2. If you received such an amount but box 11 is blank, contact your employer for the amount received as a pension or annuity.

   – _____

   Add all of your nontaxable combat pay if you elect to include it in earned income. Also enter this amount on Form 1040, line 64b. See *Combat pay, nontaxable* on this page.

   + _____

   CAUTION *Electing to include nontaxable combat pay may increase or decrease your EIC. Figure the credit with and without your nontaxable combat pay before making the election.*

   Earned Income = ☐

3. Were you self-employed at any time in 2009, or are you filing Schedule SE because you were a member of the clergy or you had church employee income, or are you filing Schedule C or C-EZ as a statutory employee?

   ☐ **Yes.** Skip question 4 and Step 6; go to Worksheet B on page 53. ☐ **No.** Continue

4. If you have:
   - 3 or more qualifying children, is your earned income less than $43,279 ($48,279 if married filing jointly)?
   - 2 qualifying children, is your earned income less than $40,295 ($45,295 if married filing jointly)?
   - 1 qualifying child, is your earned income less than $35,463 ($40,463 if married filing jointly)?
   - No qualifying children, is your earned income less than $13,440 ($18,440 if married filing jointly)?

   ☐ **Yes.** Go to Step 6. ☐ **No.** STOP You cannot take the credit.

Step 6 How To Figure the Credit

1. Do you want the IRS to figure the credit for you?

   ☐ **Yes.** See *Credit figured by the IRS* on this page. ☐ **No.** Go to Worksheet A on page 52.

Definitions and Special Rules

Adopted child. An adopted child is always treated as your own child. An adopted child includes a child lawfully placed with you for legal adoption.

Church employees. Determine how much of the amount on Form 1040, line 7, was also reported on Schedule SE, line 5a. Subtract that amount from the amount on Form 1040, line 7, and enter the result in the first space of Step 5, line 2. Be sure to answer "Yes" to question 3 in Step 5.

Clergy. The following instructions apply to ministers, members of religious orders who have not taken a vow of poverty, and Christian Science practitioners. If you are filing Schedule SE and the amount on line 2 of that schedule includes an amount that was also reported on Form 1040, line 7:

1. Enter "Clergy" on the dotted line next to Form 1040, line 64a.
2. Determine how much of the amount on Form 1040, line 7, was also reported on Schedule SE, line 2.
3. Subtract that amount from the amount on Form 1040, line 7. Enter the result in the first space of Step 5, line 2.
4. Be sure to answer "Yes" to question 3 in Step 5.

Combat pay, nontaxable. If you were a member of the U.S. Armed Forces who served in a combat zone, certain pay is excluded from your income. See *Combat Zone Exclusion* in Pub. 3. You can elect to include this pay in your earned income when figuring the EIC. The amount of your nontaxable combat pay should be shown in box 12 of Form(s) W-2 with code Q. If you are filing a joint return and both you and your spouse received nontaxable combat pay, you can each make your own election.

Credit figured by the IRS. To have the IRS figure your EIC:

1. Enter "EIC" on the dotted line next to Form 1040, line 64a.
2. Be sure you enter the nontaxable combat pay you elect to include in earned income on Form 1040, line 64b. See *Combat pay, nontaxable* above.
3. If you have a qualifying child, complete and attach Schedule EIC. If your EIC for a year after 1996 was reduced or disallowed, see *Form 8862, who must file* below.

Exception to time lived with you. Temporary absences by you or the child for special circumstances, such as school, vacation, business, medical care, military service, or detention in a juvenile facility, count as time the child lived with you. Also see *Kidnapped child* on page 19 or *Members of the military* on page 51. A child is considered to have lived with you for all of 2009 if the child was born or died in 2009 and your home was this child's home for the entire time he or she was alive in 2009.

Form 4797 filers. If the amount on Form 1040, line 13, includes an amount from Form 4797, you must use Worksheet 1 in Pub. 596 to see if you can take the EIC. Otherwise, stop; you cannot take the EIC.

Form 8862, who must file. You must file Form 8862 if your EIC for a year after 1996 was reduced or disallowed for any reason other

Appendix I: 2008 EIC Worksheet A

Worksheet A—Earned Income Credit (EIC)—Lines 64a and 64b *Keep for Your Records*

Before you begin: √ Be sure you are using the correct worksheet. Use this worksheet only if you answered "No" to Step 5, question 3, on page 50. Otherwise, use Worksheet B that begins on page 53.

Part 1
All Filers Using Worksheet A

1. Enter your earned income from Step 5 on page 50. [1]

2. Look up the amount on line 1 above in the EIC Table on pages 55–71 to find the credit. Be sure you use the correct column for your filing status and the number of children you have. Enter the credit here. [2]

   If line 2 is zero, STOP You cannot take the credit. Enter "No" on the dotted line next to line 64a.

3. Enter the amount from Form 1040, line 38. [3]

4. Are the amounts on lines 3 and 1 the same?

   ☐ **Yes.** Skip line 5; enter the amount from line 2 on line 6.

   ☐ **No.** Go to line 5.

Part 2
Filers Who Answered "No" on Line 4

5. If you have:
   - No qualifying children, is the amount on line 3 less than $7,500 ($12,500 if married filing jointly)?
   - 1 or more qualifying children, is the amount on line 3 less than $16,450 ($21,450 if married filing jointly)?

   ☐ **Yes.** Leave line 5 blank; enter the amount from line 2 on line 6.

   ☐ **No.** Look up the amount on line 3 in the EIC Table on pages 55–71 to find the credit. Be sure you use the correct column for your filing status and the number of children you have. Enter the credit here. [5]

   Look at the amounts on lines 5 and 2. Then, enter the **smaller** amount on line 6.

Part 3
Your Earned Income Credit

6. **This is your earned income credit.** [6]

   Enter this amount on Form 1040, line 64a.

Reminder—

√ If you have a qualifying child, complete and attach Schedule EIC.

CAUTION *If your EIC for a year after 1996 was reduced or disallowed, see page 50 to find out if you must file Form 8862 to take the credit for 2009.*

Appendix J: 2008 CTC Worksheet

Child Tax Credit Worksheet—Line 52 *Keep for Your Records*

CAUTION

- To be a qualifying child for the child tax credit, the child must be **under age 17** at the end of 2008 and meet the other requirements listed on page 17.
- **Do not** use this worksheet if you answered "Yes" to question 1, 2, or 3 on page 42. Instead, use Pub. 972.

1. Number of qualifying children: _____ × $1,000. Enter the result. | 1 |

2. Enter the amount from Form 1040, line 46. | 2 |

3. Add the amounts from Form 1040:

Line 47 _____

Line 48 + _____

Line 49 + _____

Line 50 + _____

Line 51 + _____ Enter the total. | 3 |

4. Are the amounts on lines 2 and 3 the same?

☐ **Yes.** STOP

You cannot take this credit because there is no tax to reduce. However, you may be able to take the **additional child tax credit.** See the **TIP** below.

☐ **No.** Subtract line 3 from line 2. | 4 |

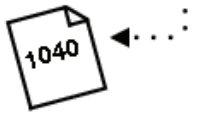
5. Is the amount on line 1 more than the amount on line 4?

☐ **Yes.** Enter the amount from line 4. Also, you may be able to take the **additional child tax credit.** See the **TIP** below.

☐ **No.** Enter the amount from line 1.

This is your child tax credit. | 5 |

Enter this amount on Form 1040, line 52.



TIP

You may be able to take the **additional child tax credit** on Form 1040, line 66, if you answered "Yes" on line 4 **or** line 5 above.

- First, complete your Form 1040 through line 65.
- Then, use Form 8812 to figure any additional child tax credit.

Appendix K: 2008 Form 8812, Additional Child Tax Credit

Form **8812**

Department of the Treasury
Internal Revenue Service (99)

Additional Child Tax Credit

1040 1040A 1040NR 8812

Complete and attach to Form 1040, Form 1040A, or Form 1040NR.

OMB No. 1545-0074

2008

Attachment Sequence No. **47**

Name(s) shown on return | Your social security number

Part I All Filers

1	Enter the amount from line 1 of your Child Tax Credit Worksheet on page 43 of the Form 1040 instructions, page 38 of the Form 1040A instructions, or page 19 of the Form 1040NR instructions. If you used Pub. 972, enter the amount from line 8 of the worksheet on page 4 of the publication		1
2	Enter the amount from Form 1040, line 52, Form 1040A, line 33, or Form 1040NR, line 47		2
3	Subtract line 2 from line 1. If zero, **stop**; you cannot take this credit		3
4a	Earned income (see instructions on back). If your main home was in a Midwestern disaster area when the disaster occurred, and you are electing to use your 2007 earned income, check here ▶ ☐	4a	
b	Nontaxable combat pay (see instructions on back)	4b	
5	Is the amount on line 4a more than $8,500? ☐ **No.** Leave line 5 blank and enter -0- on line 6. ☐ **Yes.** Subtract $8,500 from the amount on line 4a. Enter the result	5	
6	Multiply the amount on line 5 by 15% (.15) and enter the result **Next.** Do you have three or more qualifying children? ☐ **No.** If line 6 is zero, stop; you cannot take this credit. Otherwise, skip Part II and enter the **smaller** of line 3 or line 6 on line 13. ☐ **Yes.** If line 6 is equal to or more than line 3, skip Part II and enter the amount from line 3 on line 13. Otherwise, go to line 7.		6

Part II Certain Filers Who Have Three or More Qualifying Children

7	Withheld social security and Medicare taxes from Form(s) W-2, boxes 4 and 6. If married filing jointly, include your spouse's amounts with yours. If you worked for a railroad, see instructions on back	7	
8	**1040 filers:** Enter the total of the amounts from Form 1040, lines 27 and 58, plus any taxes that you identified using code "UT" and entered on the dotted line next to line 61. **1040A filers:** Enter -0-. **1040NR filers:** Enter the total of the amounts from Form 1040NR, line 53, plus any taxes that you identified using code "UT" and entered on the dotted line next to line 57.	8	
9	Add lines 7 and 8	9	
10	**1040 filers:** Enter the total of the amounts from Form 1040, lines 64a and 65. **1040A filers:** Enter the total of the amount from Form 1040A, line 40a, plus any excess social security and tier 1 RRTA taxes withheld that you entered to the left of line 43 (see instructions on back). **1040NR filers:** Enter the amount from Form 1040NR, line 60.	10	
11	Subtract line 10 from line 9. If zero or less, enter -0-		11
12	Enter the **larger** of line 6 or line 11 **Next,** enter the **smaller** of line 3 or line 12 on line 13.		12

Part III Additional Child Tax Credit

13	**This is your additional child tax credit**		13

1040 1040A 1040NR

Enter this amount on Form 1040, line 66, Form 1040A, line 41, or Form 1040NR, line 61.

For Paperwork Reduction Act Notice, see back of form. Cat. No. 10644E Form **8812** (2008)

Appendix L: 2008 Schedule A

SCHEDULES A&B (Form 1040)
Department of the Treasury Internal Revenue Service (99)

Schedule A—Itemized Deductions

(Schedule B is on back)

► Attach to Form 1040. ► See Instructions for Schedules A&B (Form 1040).

OMB No. 1545-0074
2008
Attachment Sequence No. 07

Name(s) shown on Form 1040 | Your social security number

Medical and Dental Expenses		Caution. Do not include expenses reimbursed or paid by others.			
	1	Medical and dental expenses (see page A-1)	1		
	2	Enter amount from Form 1040, line 38	2		
	3	Multiply line 2 by 7.5% (.075)	3		
	4	Subtract line 3 from line 1. If line 3 is more than line 1, enter -0-			4
Taxes You Paid (See page A-2.)	5	State and local (check only one box): a ☐ Income taxes, or b ☐ General sales taxes	5		
	6	Real estate taxes (see page A-5)	6		
	7	Personal property taxes	7		
	8	Other taxes. List type and amount ►	8		
	9	Add lines 5 through 8			9
Interest You Paid (See page A-5.)	10	Home mortgage interest and points reported to you on Form 1098	10		
	11	Home mortgage interest not reported to you on Form 1098. If paid to the person from whom you bought the home, see page A-6 and show that person's name, identifying no., and address ►	11		
Note. Personal interest is not deductible.	12	Points not reported to you on Form 1098. See page A-6 for special rules	12		
	13	Qualified mortgage insurance premiums (see page A-6)	13		
	14	Investment interest. Attach Form 4952 if required. (See page A-6.)	14		
	15	Add lines 10 through 14			15
Gifts to Charity If you made a gift and got a benefit for it, see page A-7.	16	Gifts by cash or check. If you made any gift of $250 or more, see page A-7	16		
	17	Other than by cash or check. If any gift of $250 or more, see page A-8. You must attach Form 8283 if over $500	17		
	18	Carryover from prior year	18		
	19	Add lines 16 through 18			19
Casualty and Theft Losses	20	Casualty or theft loss(es). Attach Form 4684. (See page A-8.)			20
Job Expenses and Certain Miscellaneous Deductions (See page A-9.)	21	Unreimbursed employee expenses—job travel, union dues, job education, etc. Attach Form 2106 or 2106-EZ if required. (See page A-9.) ►	21		
	22	Tax preparation fees	22		
	23	Other expenses—investment, safe deposit box, etc. List type and amount ►	23		
	24	Add lines 21 through 23	24		
	25	Enter amount from Form 1040, line 38	25		
	26	Multiply line 25 by 2% (.02)	26		
	27	Subtract line 26 from line 24. If line 26 is more than line 24, enter -0-			27
Other Miscellaneous Deductions	28	Other—from list on page A-10. List type and amount ►			28
Total Itemized Deductions	29	Is Form 1040, line 38, over $159,950 (over $79,975 if married filing separately)? ☐ No. Your deduction is not limited. Add the amounts in the far right column for lines 4 through 28. Also, enter this amount on Form 1040, line 40. ☐ Yes. Your deduction may be limited. See page A-10 for the amount to enter.			29 ►
	30	If you elect to itemize deductions even though they are less than your standard deduction, check here ► ☐			

For Paperwork Reduction Act Notice, see Form 1040 Instructions. Cat. No. 11330X Schedule A (Form 1040) 2008

Appendix M: 2008 Standard Deduction Worksheet

Form 1040—Line 40

Itemized Deductions

To figure your itemized deductions, fill in Schedule A.

Standard Deduction

If you checked the box on line 39b, your standard deduction is zero, even if you were born before January 2, 1944, were blind, paid real estate taxes, or had a net disaster loss.

Most people can find their standard deduction by looking at the amounts listed under "All others" to the left of Form 1040, line 40. But use the worksheet below to figure your standard deduction if:

- You, or your spouse if filing jointly, can be claimed as a dependent on someone's 2008 return,
- You checked any box on line 39a,
- You paid state or local real estate taxes in 2008, or
- You have a net disaster loss on Form 4684, line 18a.

Standard Deduction Worksheet—Line 40

Keep for Your Records

Do not complete this worksheet if you checked the box on line 39b; your standard deduction is zero.

1. Enter the amount shown below for your filing status.
   - Single or married filing separately—$5,450
   - Married filing jointly or Qualifying widow(er)—$10,900
   - Head of household—$8,000

   ..... 1. _____
2. Can you (or your spouse if filing jointly) be claimed as a dependent?
   - ☐ **No.** Skip line 3; enter the amount from line 1 on line 4.
   - ☐ **Yes.** Go to line 3.
3. Is your **earned income*** more than $600?
   - ☐ **Yes.** Add $300 to your earned income. Enter the total
   - ☐ **No.** Enter $900

   ..... 3. _____
4. Enter the **smaller** of line 1 or line 3. ..... 4. _____
5. If born before January 2, 1944, or blind, multiply the number on Form 1040, line 39a, by $1,050 ($1,350 if single or head of household). Otherwise, enter -0- ..... 5. _____
6. Enter any net disaster loss from Form 4684, line 18a. If more than zero, check the box on Form 1040, line 39c ..... 6. _____
7. Enter the state and local real estate taxes you paid that would be deductible on Schedule A, line 6, if you were itemizing your deductions. See the instructions for Schedule A, line 6. Do not include foreign real estate taxes ..... 7. _____
8. Enter $500 ($1,000 if married filing jointly) ..... 8. _____
9. Enter the **smaller** of line 7 or line 8. If more than zero, check the box on Form 1040, line 39c ..... 9. _____
10. Add lines 4, 5, 6, and 9. Enter the total here and on Form 1040, line 40. ..... 10. _____

* ***Earned income*** *includes wages, salaries, tips, professional fees, and other compensation received for personal services you performed. It also includes any amount received as a scholarship that you must include in your income. Generally, your earned income is the total of the amount(s) you reported on Form 1040, lines 7, 12, and 18, minus the amount, if any, on line 27.*

Need more information or forms? See page 87.

Predicting Aggregate Taxpayer Compliance Behavior

Alan Plumley, Internal Revenue Service; Brian Erard, Brian Erard & Associates; and Derek Snaidauf, IBM Global Business Services[1]

The IRS seeks to be able to estimate the impact that its service and enforcement interventions have on the voluntary compliance of taxpayers. A key method for doing that is to conduct field experiments among taxpayers as they respond to their real tax obligations. A typical experimental design includes tracking the behavior of both a test group and a control group, which are sufficiently identical in all relevant respects except that the intervention is applied only to the test group. We believe, however, given the nature of most IRS interventions and the diversity and geographically non-uniform distribution of the taxpayer population, that it will often be very difficult to construct a control group that is sufficiently identical to a test group in all relevant respects, except for the intervention (or in preventing the control group from being affected by the intervention). The goal of this research, therefore, was to develop the capability to control statistically for factors that influence taxpayer behavior, supplementing the role of control groups in future field experiments. To do this, we develop econometric models to predict aggregate reporting behavior among individuals. That is, we seek to estimate what taxpayers would have done in the absence of an intervention. The difference between what they would have done and what they actually did in any given time period is a measure of the change in voluntary compliance. If that measure of behavioral change improves as a result of an intervention, then the intervention is considered to have improved compliance.

Data

We extend in this paper the foundational studies by Plumley (1996) and Dubin (2007), which were focused on estimating the independent effects of many of the determinants of voluntary compliance. Our methodology is similar, but our focus is on developing accurate predictions of the dependent variable: taxpayer reporting behavior. To do this, we compiled a robust database containing detailed state-level longitudinal data on taxpayer characteristics and behavior, as well as IRS activities, from Tax Years 1982 through 2009. This includes data for 750 variables from over 20 different sources. Approximately 200 of these variables were updated from both Dubin's and Plumley's studies and about 550 additional variables, not included in these earlier databases, were incorporated to provide a richer set of potential determinants of taxpayer behavior. Beyond our current empirical work, the analysis database will serve as a valuable tool for researchers to employ when analyzing various issues relating to taxpayer filing and reporting behavior across states and over time. To ensure uniformity between all our variables, we combined the District of Columbia with Maryland, since data from the District of Columbia was not available for all variables and it was already combined with Maryland in previous studies.

In the case of the non-IRS source data, a major task was the construction of variables from the Annual Social and Economic Supplement of the Current Population Survey conducted by the Census Bureau. This data source includes nationally representative micro-level cross-sectional data on residents of housing units (homes, apartments, group living arrangements, etc.) based on survey responses. To construct the desired variables from this data source, it was necessary to construct filing (or potential filing) units out of the household residents for each year of the survey, assess whether a tax return was required to be filed, and construct measures of income, filing status, and other relevant factors at the level of the filing unit before aggregating to the state level.

Once the data were collected from the various sources, the data were processed, standardized, and validated to help ensure their accuracy. Ultimately, we created a comprehensive analysis database that included detailed state-level longitudinal data on taxpayer characteristics, income tax reporting behavior, IRS service and

enforcement activities, and other factors. Where feasible, annual state-level observations were collected for the entire period from 1982 through 2009. However, in some cases variables could be measured only at a national level or only over a shorter time span. For example, variables available only at a national level include several of the political science variables, such as the ratio of Democrats to Republicans in the Senate and House and the average political ideology score of members in Congress, economic variables, including the GDP deflator and CPI, and finally certain IRS variables including web data and complexity measures. Variables that are available only over a shorter time span, either due to lack of data availability or a change in the data collection methodology, include burden, service activity measures, variables from the Bureau of Labor Statistics (BLS), income and offsets data, and criminal investigation measures.

Model Development

Using our robust database, we conducted a preliminary econometric analysis to explore its suitability as a source for predicting taxpayer reporting behavior. We followed a logical analytic progression beginning with specifications similar to previous studies and gradually introducing a number of methodological refinements to incorporate new variables, explore alternative functional forms, and test predictive performance. The focus in this paper is on the reporting of overall income and offsets on the tax return. Thus, our current study not only extends previous efforts for a longer time-frame, but it also integrates a richer set of variables and incorporates innovations to the econometric methodology. Ultimately, the results of our analysis provide a preliminary assessment of the feasibility of using state-level panel data to predict taxpayer reporting behavior.

Our econometric methodology builds on the prior work of Dubin et al. (1990), Dubin (2007), and especially Plumley (1996). As in those studies, we employ panel data regression techniques to explain the aggregate reporting behavior of taxpayers across different states and time periods as a function of various IRS activities and other relevant behavioral determinants.

In general terms our econometric specification is as follows:

$$Y_{it} = \alpha_i + \gamma_t + \beta_A^{'} A_t + \beta_O^{'} O_t + \varepsilon_{it} ,$$

where *Y* represents a measure of reporting behavior (such as total income reported, total offsets reported, net income reported, or income reported for a specified line item), *A* represents a set of IRS activities (including both enforcement and service activities), and *O* represents a set of other relevant measured determinants of reporting behavior. The subscripts "*i*" and "*t*" represent individual states and years, respectively, reflecting our objective of explaining the variation in reporting behavior across both states and time. In the above specification, the parameters β_A and β_O represent coefficients to be estimated. The term ε_{it} is an error term that is meant to capture the net impact of unobserved factors across states and over time on state-level reporting behavior. Finally, the terms α_i and γ_t represent possible sources of state-specific and year-specific heterogeneity. More specifically, α_i represents unobserved time-invariant differences across states and that drive inter-state differences in reporting behavior, while γ_t represents unobserved state-invariant differences across years that drive inter-temporal differences in reporting behavior.

Following Plumley (1996), we specified two alternative definitions of total income: (1) an "A" version that excluded income items that were subject to substantial changes in reporting requirements over the estimation period; and (2) a "B" version that included all taxable income sources, regardless of changes in reporting requirements. A comparable pair of measures for total offsets was also developed. We found that the levels of income and offsets were relatively steady over time under the more restrictive "A" definition, but tended to be somewhat more variable under the "B" definition. Our analysis for this paper focused on the "A" definition.

Fixed vs. Random Effects

The two most common approaches to modeling heterogeneity in panel data are fixed effects and random effects. In the context of the state-specific heterogeneity term α_i in our above specification, a fixed effects specification treats this term as a state-specific constant term in the analysis. In contrast, a random effects

specification treats the value of α_i for each state as a random draw from a probability distribution. An advantage of the fixed effects specification is that it produces consistent estimates of the parameters of the model even when the α_i terms are correlated with one or more of the explanatory variables in the model. However, if these terms are not correlated with any of the explanatory variables, the random effects specification produces more efficient (precise) estimates; intuitively, the random effects specification exhausts fewer degrees of freedom, because it is not necessary to estimate the value of α_i ("nuisance parameter") for each state as one does with the fixed effects specification. The fixed effects specification also yields only conditional predictions, in the sense that it is limited to predicting observations that come from units for which a fixed effect has been estimated. However, as Plumley (1996) points out, since the units in our case are states and essentially all states are included in our analysis,[2] this is not a meaningful limitation for our application. Like Plumley (1996), we tend to favor the fixed effects approach for this study as it produces consistent estimates under a wider range of circumstances than the random effects approach. However, we perform some comparisons with the random effects approach to see how sensitive the findings are to the choice of method.

One can also apply a fixed or random effects specification for the time-specific heterogeneity term, in which case one has what is known as a "two-way" fixed or random effects specification. An alternative approach we employ in much of our analysis is to model the term γ_t using one or more time trend terms; for instance:

$$\gamma_t = \gamma_1 t + \gamma_2 t^2 .$$

In this example, time-specific heterogeneity would be modeled using a quadratic trend.

Endogeneity

Both Plumley (1996) and Dubin (2007) recognize that the audit rate is likely to be an endogenous explanatory variable. To account for this, they employ an instrumental variables approach. We follow Plumley in using measures involving state level measures of direct examination time as instruments; specifically, our instruments are the direct examination time percentage (the share of examiners' time directly devoted to examination activities) and the lagged value of the average direct examination time. For our fixed effects specifications, we employ the standard two-stage least squares approach to estimation. In our random effects specifications, we employ the instrumental variables approach proposed by Balestra and Varadharajan-Krishnakumar (1987).

Another explanatory variable that is likely to be endogenous in our model is the combined state and federal marginal tax rate. Owing to the graduated federal (and in some cases, state) tax structure, the state level marginal tax rate will tend to be lower when state level income reporting is low. In our analysis, we experiment with using the combined state and federal marginal tax rate based on a fixed national measure of the income distribution as an instrument. We find that we get extremely similar results when we directly substitute this instrument for the endogenous measure in our analysis. Since the latter approach simplifies prediction, we use it in our prediction exercises.

Other Statistical Issues

Our work goes beyond the previous studies to address a host of statistical issues, including the use of: specifications with ratio dependent and explanatory variables versus alternative functional forms; short versus long panels; and year dummies versus trend terms. For the most part, many of our results are reasonably robust against these alternatives. For instance, we generally obtained qualitatively similar parameter estimates (in terms of coefficient signs and statistical significance) when we substituted alternative functional forms for the base-case ratio specifications. Details of these analyses are presented later in this paper.

Predictive Accuracy

Since an important objective of this study is to evaluate the potential of our alternative specifications to forecast future reporting behavior, we have developed two alternative methodologies for measuring forecasting

performance. The first is based on a "leave-one-out" prediction methodology in which one year of data at a time is left out of the estimation sample and the resulting parameter estimates are then used to predict reporting behavior within each state for the excluded year. Successively leaving out each year from the estimation sample produces a set of out-of-sample predictions of reporting behavior for each state and year, which can then be compared against actual reporting behavior. Under our second methodology, we exclude the last four years of the data sample from estimation and then use the resulting parameter estimates to forecast reporting behavior in each of these four years. A comparison against the actual reporting behavior provides an assessment of forecasting performance one, two, three, and four years into the future.

For both our leave-one-out and step-ahead forecasting approaches, we focus on two alternative measures of out-of-sample predictive performance. The first is the mean absolute deviation of the out-of-sample prediction of reported income in each state and time period from the true value of reported income. The second is the root mean-squared error (i.e., the square root of the average squared deviation of the out-of-sample prediction from the actual value). Both of these measures are normalized by dividing them by the average value of reported income over all states and time periods. We refer to the first measure as the "absolute deviation as a percentage of income". The second measure is known in the statistics literature as the "coefficient of variation of root mean-squared error".

A limitation of modeling the time-specific heterogeneity term using fixed effects is that the value of the fixed effect would not be known for years outside of the sample period, which makes forecasting difficult. We therefore employ trend terms rather than yearly fixed effects in much of our analysis. However, a comparison of our results based on our longer panel analyses indicates that certain parameter estimates (notably, the audit rate coefficient) are sensitive to whether yearly fixed effects or trend terms are employed. To investigate the impact of this choice on predictive performance, we have developed an econometric approach to forecasting with yearly fixed effects. Under this approach, we predict the value of the fixed effects for years outside of the sample period based on the estimated sample period fixed effects. We use a Box-Jenkins time series approach (autoregressive integrated moving average, or ARIMA, analysis) to model the fixed effects. Results of our analysis indicate that an autoregressive process of order 2 provides a reasonable fit to the data in the specification we have investigated.

Another complication of our analysis for prediction purposes is the presence of endogenous explanatory variables. Consider a fixed effects specification of the following form:

$$Y_{it} = \alpha_i + \gamma_t + \beta_A A_t + \beta_O' O_t + \varepsilon_t ,$$

where the variable *A* represents the audit rate – an endogenous explanatory variable. We can consistently estimate the parameters of this model using an instrumental variables approach. Suppose that we then substitute the predicted values of the coefficients in for the actual values and attempt to predict *Y* as:

$$\hat{Y}_{it} = \hat{\alpha}_i + \hat{\gamma}_t + \hat{\beta}_A A_t + \hat{\beta}_O' O_t .$$

In general, this will not be a consistent predictor of *Y*, because the error term ε will be correlated with *A*. Consequently, the conditional expectation of $(Y - \hat{Y})$ given *O* and *A* will (asymptotically) converge to the value $E(\varepsilon \mid A)$ the value of the conditional expectation of the error term given the audit rate *A*. Since ε and *A* are correlated, this expectation will not be equal to zero. To address this issue, we employ a two-stage approach to prediction motivated by the Durbin-Wu-Hausman specification test for endogeneity. In the first stage, we regress the audit rate against all of the explanatory variables of the model as well as the instruments (just as in the first stage of two-stage least squares estimation). We obtain the residual (*u*) from this regression. In the second stage, we estimate the following regression specification:

$$Y_{it} = \alpha_i + \gamma_t + \beta_A A_t + \beta_O' O_t + \lambda u_t + \varepsilon_t .$$

Under the Durbin-Wu-Hausman test, one performs a t-test of whether the coefficient λ is equal to zero. The intuition for this test is that this extra term involving the residual *u* accounts for the correlation between ε and *A,* so that if λ =0, there is no correlation and, hence, *A* is exogenous. Although we have not seen this specification used in the econometric literature for purposes of prediction, it can also serve this function. In particular, this extra term involving the residual *u* directly accounts for the conditional expectation of ε given *A* that was left out of the above prediction formula and was the source of inconsistent estimation.

Estimation Results

We have employed a systematic approach to estimation to explore the sensitivity of our methodology to the choice of time period, the selection of explanatory variables, the specification of functional forms, and the use of fixed vs. random effects. A comparison of the results provides evidence of the degree to which the methodology is robust to different modeling assumptions and yields some insights about productive areas for further data collection and modeling refinements.

Our preliminary econometric models explore the reporting of a broad measure of the overall total amount of income reported on tax returns before any statutory adjustments or deductions. We first present the estimation results for our base specification and extensions for our model of total income reporting. We then discuss the predictive performance of selected specifications.

Base Specification and Extensions for Total Income Reporting

We begin by specifying the model of income reporting presented by Plumley (based on his "A" definition of total income) using the same time period (1982-1991). Consistent with his approach, we have employed a limited definition of income that controls for some of the changes in federal income reporting requirements over time. In addition, we have included certain control variables in our analysis to account for various changes in federal tax laws, such as the forms of income that are excluded from taxation, the amount allowed for dependent exemptions, various features of the Tax Reform Act of 1986 (captured through a dummy variable), and certain other tax changes (captured through either trend terms or yearly fixed effects).[3]

In his model, Plumley employed a two-way fixed effects model (state and year effects) to explain state level income reporting on required returns (returns that were legally required to be filed). We examine how the results are impacted by substituting trend terms for the year effects. We then explore the sensitivity of the results to using updated measures for some variables, excluding certain variables that we were unable to update for future years, and including some new or substitute explanatory variables. Next, we extend the analysis to different time periods and examine the role of some additional explanatory variables.

We have observed that both Dubin (2007) and Plumley (1996) have relied extensively on ratio variables in their analyses. As summarized by Wiseman (2009), the use of ratio measures in regression analysis is controversial, and there is a growing literature demonstrating that such measures can sometimes lead to spurious and inconsistent findings. We have therefore estimated some alternative specifications that do not rely as heavily on ratios. For instance, we have investigated specifications in which the natural log of reported income is regressed against the natural log of personal income and other explanatory variables rather than using the ratio of reported income to personal income as the dependent variable as is done in Plumley's study. We have also estimated specifications in which many of the ratio explanatory variables have been replaced by variables that separately account for their numerators and denominators.

Table 1 below provides a preliminary sensitivity analysis of the model of income reporting presented by Plumley.[4] The first column includes the results for his original specification. The dependent variable in this specification is the ratio of income reported on returns that were required to be filed to total personal income. In the second column, various modifications have been made, including dropping his information returns matching (lnirp), criminal investigations (lncid), taxpayer service calls (tps_callspc), and IRS return preparation variables (tps_retpreppc) for which we do not have updated measures for subsequent years. Also in the second column, trend terms have replaced the yearly fixed effects; the marginal tax rate variables (mtr15k and mtr57k) have been replaced by a combined state-federal measure of the marginal tax rate (which has been

instrumented); a broader measures of the value of dependent exemptions has replaced the measure of the value of child exemptions (childexemptspct); and updated versions of certain other variables (such as soleprops—the percentage of sole proprietors) have been introduced. Overall, the results are not very sensitive to these changes, although the estimates of the coefficients of soleprops and lnaud (the natural log of one plus the audit rate) have become less precise.

In the original Plumley specification, the audit start rate was employed as an explanatory variable. In column 3, the audit close rate was substituted.[5] This has only a very modest impact on the results. In column 4, the specification in column 3 is estimated using random effects rather than state-level fixed effects. The results are quite comparable.

The first column in Table 2 repeats the information in column 3 of the previous table for the case in which the audit close rate is employed. The second column extends the original time period (1982-1991) to a longer time span (1982-2007). While many of the coefficients have the same signs and similar levels of precision in the longer panel, there are some noteworthy exceptions. In particular, the coefficients of the audit rate and the marginal tax rate change signs and become significant. In the case of the marginal tax rate, the new negative coefficient is intuitive, suggesting that high marginal tax rates lead to less compliance. However, the new negative coefficient on the audit rate is counter-intuitive. One would expect, all else equal, that a higher audit rate would yield relatively greater (not less) compliance. In the third column, we have included some additional trend terms in our specification. This does not substantively alter the results. In column 4, we apply random effects estimation to the specification from column 3. This also has only a modest impact on the results. Finally, in column 5, we employ a two-way fixed effects specification that includes year dummies rather than trend terms. This specification change does have an important impact on the results. In particular, the coefficient of the audit rate now becomes positive and significant. Apparently, the year dummies are able to capture certain state-invariant changes in taxpayer circumstances that the trend terms cannot. We have performed a Wald test of the joint significance of the year dummies in our specification, and the evidence strongly supports the alternative hypothesis that the year dummies are jointly significant explanatory variables. We later examine whether the inclusion of year dummies translates into an improved forecasting performance over the use of trend terms.

In Table 3, we experiment with some additional variables not included in the original Plumley specification. In column 1, the specification includes explanatory variables describing the percentages of potential returns for which the primary taxpayer has some college education, is male, and is a homeowner. In addition, population density and the Gini coefficient based on CPS family income (a measure of income inequality within the state) are included as explanatory variables. Only the population density is found to be significant over the 1982-1991 estimation period. In the second column, the estimation period is extended to 2004. With these additional data points, all of the new explanatory variables are found to be statistically significant. However, as with Table 2, the coefficient of the audit rate becomes negative and significant when the time period is extended. In column 3 of the table, the rate of criminal sentences for tax evasion and money laundering is included as an explanatory variable for the 1988-2004 period. This variable is not found to have a significant impact on reporting behavior. In the fourth column, year dummies are employed rather than trend terms. The criminal sentence rate remains insignificant in this specification. However, consistent with previous findings, the audit rate coefficient becomes positive when year dummies are employed (however, the estimate is statistically insignificant). We have again employed a Wald test of the joint significance of the year dummies and the evidence again strongly supports the alternative hypothesis that these variables are jointly significant.

TABLE 1. Model of Income Reporting Presented by Plumley (1996) and Some Variations

	(1)	(2)	(3)	(4)
	Original Specification	Various Changes	Audit Close Rate	Random Effects
Lnaud	11.259	9.358	9.298	7.075
	(1.52)	(0.82)	(1.29)	(0.96)
filingrate	0.302	0.300	0.308	0.314
	(5.35)**	(4.75)**	(5.65)**	(5.62)**
fthresholdpct	0.935	1.286	1.111	0.986
	(2.39)*	(5.21)**	(4.28)**	(3.89)**
mtr15k	1.292			
	(1.18)			
mtr57k	-1.421			
	(0.68)			
childexemptspct	1.582			
	(1.90)			
Lnburden	8.489	7.924	7.039	6.710
	(1.18)	(3.27)**	(2.72)**	(2.62)**
Soleprops	2.635	0.702	0.666	0.595
	(2.78)**	(1.41)	(1.50)	(1.40)
soleproptfs	-0.056	-0.020	-0.018	-0.018
	(2.96)**	(1.97)*	(2.10)*	(2.13)*
Paidprep	-0.124	-0.106	-0.116	-0.133
	(2.89)**	(2.25)*	(3.34)**	(3.99)**
Lnirp	-9.160			
	(1.63)			
Lncid	1.122			
	(3.17)**			
tps_callspc	-0.006			
	(1.78)			
tps_retpreppc	0.055			
	(0.69)			
Singles	0.114	0.240	0.251	0.273
	(0.58)	(2.55)*	(3.09)**	(3.35)**
under30	-0.099	-0.020	0.038	0.002
	(1.06)	(0.22)	(0.42)	(0.03)
over64	-0.060	0.021	0.084	0.033
	(0.56)	(0.19)	(0.69)	(0.28)
Pcbirths	0.725	0.659	0.809	0.609
	(2.99)**	(3.60)**	(4.32)**	(3.65)**
exclincomepct	-0.502	-0.815	-0.583	-0.673
	(1.21)	(1.53)	(1.50)	(1.75)
unemplrate	-0.473	-0.485	-0.448	-0.400
	(3.03)**	(2.44)*	(3.29)**	(2.96)**
Trend		0.739	0.552	0.268
		(1.15)	(1.51)	(0.73)
tra86dum		-6.142	-9.384	-8.381
		(3.45)**	(4.50)**	(4.33)**
Tratrend		1.211	1.288	1.382
		(4.80)**	(5.20)**	(5.56)**
depamountpct		0.418	1.204	0.524
		(0.59)	(1.60)	(0.76)
c_marg		71.440	15.070	11.798
		(4.44)**	(0.62)	(0.52)
Constant	10.459	-67.978	-51.611	-39.961
	(0.09)	(3.32)**	(2.22)*	(1.72)
Observations	490	490	490	490

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 2. Results of Estimation of Model Using a Longer Panel

	(1)	(2)	(3)	(4)	(5)
	Audit Close Rate	Longer Sample	Additional Trend Terms	Random Effects	Year Dummies
lnaudnw	9.298	-8.134	-7.995	-6.982	8.789
	(1.29)	(4.40)**	(4.51)**	(3.73)**	(2.36)*
filingrate	0.308	0.351	0.358	0.323	0.310
	(5.65)**	(10.62)**	(11.05)**	(9.73)**	(10.21)**
fthresholdpct	1.111	0.249	0.595	0.283	0.718
	(4.28)**	(1.38)	(3.36)**	(1.74)	(3.94)**
c_marg	15.070	-67.941	-3.899	-21.186	-12.561
	(0.62)	(6.52)**	(0.31)	(2.10)*	(0.65)
depamountpct	1.204	1.957	1.285	0.677	1.488
	(1.60)	(4.22)**	(2.95)**	(1.60)	(3.81)**
lnburden	7.039	4.887	5.298	4.860	5.153
	(2.72)**	(5.51)**	(6.40)**	(5.80)**	(6.03)**
soleprops	0.666	-0.315	-0.629	-0.500	0.525
	(1.50)	(0.97)	(2.03)*	(1.78)	(1.78)
soleproptfs	-0.018	0.008	0.012	0.007	-0.013
	(2.10)*	(1.33)	(2.21)*	(1.48)	(2.33)*
paidprep	-0.116	-0.070	-0.081	-0.093	-0.002
	(3.34)**	(2.69)**	(3.37)**	(4.10)**	(0.11)
singles	0.251	0.224	0.234	0.224	0.163
	(3.09)**	(4.04)**	(4.48)**	(4.24)**	(3.47)**
under30	0.038	-0.012	-0.105	-0.183	0.029
	(0.42)	(0.22)	(1.93)	(3.41)**	(0.56)
over64	0.084	0.007	-0.071	-0.142	0.058
	(0.69)	(0.10)	(1.11)	(2.22)*	(0.96)
pcbirths	0.809	0.847	0.613	0.441	0.337
	(4.32)**	(7.09)**	(5.40)**	(4.29)**	(3.08)**
exclincomepct	-0.583	-0.828	-0.539	-0.767	-0.726
	(1.50)	(3.00)**	(2.09)*	(3.00)**	(3.08)**
unemplrate	-0.448	-0.338	-0.420	-0.357	-0.241
	(3.29)**	(3.82)**	(4.91)**	(4.08)**	(2.70)**
trend	0.552	-0.953	-1.036	-1.271	
	(1.51)	(4.68)**	(5.47)**	(6.93)**	
tra86dum	-9.384	-6.291	3.112	3.555	
	(4.50)**	(5.58)**	(2.05)*	(2.57)*	
tratrend	1.288	0.731	-0.333	-0.125	
	(5.20)**	(4.27)**	(1.62)	(0.63)	
dum91			1.795	1.827	
			(3.87)**	(3.75)**	
Constant	-51.611	10.515	-7.172	17.365	-22.213
	(2.22)*	(0.88)	(0.63)	(1.65)	(1.84)
dum92			1.369	1.323	
			(3.06)**	(2.84)**	
trendsq			0.035	0.033	
			(8.21)**	(8.51)**	
Observations	490	1274	1274	1274	1274

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 3. Variables Added to the Model of Income Reporting Presented by Plumley (1996)

	(1)	(2)	(3)	(4)
	New Variables 1982-1991	New Variables 1982-2004	Cidsentrate 1988-2004	Year Dummies
lnaudnw	9.716	-13.751	-9.996	3.871
	(1.29)	(5.89)**	(4.32)**	(1.14)
filingrate	0.292	0.389	0.414	0.409
	(4.71)**	(11.26)**	(10.46)**	(11.64)**
fthresholdpct	1.052	0.967	1.026	1.647
	(3.64)**	(4.65)**	(3.79)**	(6.91)**
c_marg	-25.591	-20.222	47.021	37.828
	(0.72)	(1.47)	(2.30)*	(1.28)
depamountpct	1.297	0.695	1.356	1.762
	(1.68)	(1.40)	(2.68)**	(4.15)**
lnburden	8.724	2.896	2.715	2.730
	(2.99)**	(3.01)**	(3.02)**	(3.47)**
soleprops	0.481	-0.311	0.097	0.617
	(1.04)	(0.94)	(0.21)	(1.59)
soleproptfs	-0.013	0.005	-0.002	-0.013
	(1.53)	(0.84)	(0.23)	(1.81)
paidprep	-0.150	-0.074	0.029	0.036
	(4.09)**	(2.74)**	(0.82)	(1.20)
singles	0.196	0.155	0.202	0.130
	(1.91)	(2.17)*	(2.78)**	(2.11)*
under30	0.024	-0.219	-0.048	0.070
	(0.24)	(3.46)**	(0.74)	(1.21)
over64	0.082	-0.083	0.011	0.061
	(0.65)	(1.12)	(0.14)	(0.93)
pcbirths	0.431	0.413	0.457	0.050
	(1.97)*	(3.22)**	(2.24)*	(0.29)
exclincomepct	-0.419	-1.000	-0.950	-0.236
	(1.05)	(3.51)**	(3.17)**	(0.91)
unemplrate	-0.791	-0.332	-0.559	-0.516
	(3.83)**	(3.44)**	(4.58)**	(4.23)**
trend	-1.733	-1.386	-1.961	
	(1.64)	(6.39)**	(6.99)**	
tra86dum	3.001	-2.061		
	(0.40)	(1.03)		
tratrend	-1.276	0.388		
	(0.88)	(1.50)		
trendsq	0.295	0.007	0.043	
	(1.85)	(0.97)	(4.37)**	
collegepct	-0.031	0.131	0.105	0.057
	(0.41)	(2.68)**	(2.16)*	(1.32)
malepct	0.055	0.206	0.124	0.105
	(0.51)	(2.69)**	(1.56)	(1.58)
homeownerpct	-0.081	-0.158	-0.092	-0.052
	(0.90)	(2.73)**	(1.58)	(1.04)
popdensity	0.105	0.044	0.041	0.020
	(3.42)**	(5.22)**	(4.23)**	(2.25)*
gini_faminc	-1.773	29.430	29.943	1.342
	(0.10)	(3.17)**	(3.13)**	(0.15)
dum91		0.913	1.560	
		(1.77)	(3.27)**	
dum92		0.621	1.516	
		(1.26)	(3.36)**	
cidsentrate			-0.041	-0.322
			(0.13)	(1.26)
Constant	-33.145	7.290	-28.571	-46.565
	(1.31)	(0.54)	(1.85)	(3.11)**
Observations	490	1127	833	833

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

In Table 4, we experiment with including some more new explanatory variables. These include an annual national measure of hours by IRS personnel in taxpayer-facing service occupations and annual national measure of tax return complexity. The state level measure of the rate of criminal sentences for tax evasion and money laundering that was introduced in Table 4 is also included. The sample period extends from 1996-2005 as this is the period for which we have measures of these three new variables. The first column of Table 4 presents a base level specification for that sample period that excludes the new variables, while column 2 includes them. The results for taxpayer-facing service hours and complexity are somewhat unexpected, suggesting that more service hours leads to lower taxpayer reporting and that greater tax system complexity leads to higher reporting. The estimated impact of criminal sentences on reporting behavior is positive for this time period, although it is not very precisely estimated. In column 3, the criminal sentence rate variable is dropped, allowing us to include an additional two years in the sample period. The signs of the estimated coefficients on service hours and complexity are unchanged, although their magnitude has been reduced. In column 4, the trend term has been replaced by year dummies in the base specification without any of the new variables. As observed in previous specifications, the use of year dummies results in a change in the signs of the estimated audit rate and marginal tax rate coefficients. In column 5, the new criminal sentence has been included in the two-way fixed effects specification. It was not possible to include the service hours and complexity variables in this specification, because these national level estimates are perfectly collinear with the year dummies. With the two-way fixed effects specification, the sign of the criminal sentence variable has reversed, although the estimate is statistically insignificant.

In Table 5, we experiment with a state level measure of attempted calls to the IRS help line. This is similar to the measure of calls handled by taxpayer service (tps_callspc) that was used by Plumley for the period from 1982-1991. Our measure is available for the 2002-2007 time period. The first column of the Table provides a base specification for this period that excludes the attempted calls variable, while the second column includes this variable. The results indicate that telephone assistance is positively associated with income reporting, although the coefficient estimate is not very precise. In columns 3 and 4, we repeat this exercise, this time using a two-way random effects specification rather than including a time trend. In this specification, the coefficient of the calls attempted variable becomes negative, but insignificant. Also observe that the audit rate and marginal tax rate coefficients have increased substantially compared to the earlier specification involving the time trend. It appears that the results for this time period (2002-2007) are rather fragile.

We have also experimented with functional forms. For instance, we have estimated variants of our specifications in Tables 4 and 5 in which the dependent variable is the log of income reported rather than the ratio of income reported to total personal income. In one variant, the natural log of total personal income is included as an additional explanatory variable and the other variables are the same as in the previous specification. In the other variant, many of the ratio variables are eliminated. In their place are separate measures of the numerators and denominators of these ratios. We have experimented with year dummies and trend terms as alternatives in these specifications. As with the results in this section, a Wald test supports the joint significance of the year dummies. The estimation results are qualitatively similar to the results presented in Tables 4 and 5.

Predictive Performance of Models of Total Income Reporting

An important objective of this project is to develop a preliminary assessment of the predictive capability of the panel data modeling approach. We begin by evaluating how well alternative econometric specifications of total income reporting forecast out of sample when they are based on the same 1982-1991 period employed in the Plumley study. We then explore how the forecasting performance changes when the models are estimated over a longer time span.

The first two columns of Table 6 below respectively present results from a two-way fixed effects version and a one way fixed effects with trends version of the parsimonious specification of income reporting behavior provided earlier in the first column of Table 2. The results for the two specifications are quite similar. In these specifications, both the natural log of the audit close rate and the combined state-federal marginal tax rate are treated using instrumental variables. Since having two instrumented variables complicates the prediction process to some extent, we experiment in column 3 with directly using the instrument for the state-federal marginal tax rate in the specification rather than as an instrumental variable. This variable represents

the computed combined state-federal marginal tax rate based on a fixed national distribution of income in 1995. The results indicate that this approach yields very similar estimates to those shown for the instrumental variables specification in column 2. In column 4, we extend the specification in column 3 to include some additional explanatory variables that were not included in Plumley's original specification. Of these additional variables, only population density proves to be statistically significant.

TABLE 4. Experimentation with Some Additional New Explanatory Variables

	(1)	(2)	(3)	(4)	(5)
	Base Model 1996-2005	New Variables 1996-2005	Longer Period (no cidsentrate) 1996-2007	Base Year Dummies 1996-2005	Cidsentrate Year Dummies 1996-2005
lnaudnw	-9.938	-12.306	-2.209	18.132	18.223
	(3.08)**	(3.11)**	(0.70)	(2.43)*	(2.44)*
filingrate	0.407	0.408	0.310	0.364	0.369
	(7.84)**	(7.59)**	(6.39)**	(7.21)**	(7.29)**
fthresholdpct	2.240	2.163	0.945	2.084	2.116
	(7.07)**	(6.74)**	(3.82)**	(6.86)**	(6.94)**
c_marg_95_fixed	89.795	76.055	60.857	-81.169	-84.842
	(5.05)**	(3.88)**	(3.42)**	(2.88)**	(2.99)**
depamountpct	0.970	1.167	0.841	-0.268	-0.338
	(1.51)	(1.76)	(1.42)	(0.42)	(0.53)
lnburden	1.356	0.603	1.671	2.321	2.314
	(1.54)	(0.65)	(1.90)	(2.67)**	(2.67)**
soleprops	-0.427	-0.639	0.163	1.009	0.932
	(0.47)	(0.68)	(0.21)	(1.07)	(0.99)
soleproptfs	0.009	0.012	-0.004	-0.020	-0.019
	(0.57)	(0.75)	(0.33)	(1.21)	(1.14)
paidprep	0.114	0.144	0.038	0.135	0.134
	(2.51)*	(3.08)**	(0.92)	(3.04)**	(3.03)**
singles	0.234	0.249	0.181	0.187	0.187
	(2.77)**	(2.92)**	(2.38)*	(2.40)*	(2.40)*
under30	0.135	0.147	0.119	0.167	0.166
	(1.82)	(1.97)*	(1.77)	(2.44)*	(2.43)*
over64	0.000	0.018	-0.001	-0.020	-0.025
	(0.00)	(0.21)	(0.02)	(0.27)	(0.33)
pcbirths	0.281	0.548	0.153	0.433	0.398
	(0.86)	(1.60)	(0.57)	(1.41)	(1.29)
exclincomepct	0.002	-0.085	-0.681	-0.556	-0.588
	(0.00)	(0.21)	(2.10)*	(1.63)	(1.71)
unemplrate	-1.183	-1.109	-1.009	-0.782	-0.765
	(8.13)**	(7.39)**	(7.79)**	(5.09)**	(4.96)**
collegepct	0.058	0.068	0.116	0.018	0.014
	(1.05)	(1.23)	(2.37)*	(0.36)	(0.29)
malepct	0.014	0.013	0.049	0.036	0.036
	(0.15)	(0.14)	(0.55)	(0.42)	(0.43)
homeownerpct	-0.026	-0.000	0.006	0.065	0.068
	(0.36)	(0.01)	(0.09)	(0.99)	(1.03)
popdensity	0.004	0.009	-0.007	-0.008	-0.008
	(0.36)	(0.78)	(0.72)	(0.68)	(0.72)
trend	0.139	-0.559	-0.013		
	(0.86)	(1.70)	(0.04)		
hoursrvrate		-19.377	-8.777		
		(3.02)**	(1.75)		
complexity		13.540	7.056		
		(3.15)**	(2.28)*		
cidsentrate		0.723			-0.338
		(1.77)			(0.93)
Constant	-70.995	-69.674	-31.680	-31.592	-30.355
	(3.80)**	(3.68)**	(1.88)	(1.83)	(1.75)
Observations	490	490	588	490	490

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 5. Inclusion of State-level Measure of Attempted Calls to the IRS Help Line

	(1)	(2)	(3)	(4)
	Base 2002-2007	callattemptpct	Base 2002-2007, Year Dummies	Year Dummies, callattemptpct
lnaudnw	18.554	16.233	67.785	71.291
	(1.95)	(1.71)	(1.77)	(1.77)
filingrate	0.218	0.233	0.191	0.189
	(2.77)**	(2.97)**	(1.57)	(1.54)
fthresholdpct	0.138	0.245	0.090	0.083
	(0.40)	(0.71)	(0.11)	(0.10)
c_marg_95_fixed	19.546	-12.034	-252.328	-257.324
	(0.92)	(0.42)	(2.70)**	(2.66)**
depamountpct	1.238	1.301	0.541	0.455
	(1.30)	(1.38)	(0.43)	(0.35)
lnburden	2.751	2.908	4.192	4.220
	(2.91)**	(3.11)**	(2.94)**	(2.91)**
soleprops	2.180	1.992	-0.788	-0.832
	(1.67)	(1.54)	(0.36)	(0.37)
soleproptfs	-0.040	-0.037	0.012	0.013
	(1.82)	(1.71)	(0.33)	(0.35)
paidprep	-0.202	-0.166	-0.087	-0.095
	(3.17)**	(2.49)*	(0.96)	(1.02)
singles	-0.008	-0.014	-0.033	-0.029
	(0.08)	(0.14)	(0.24)	(0.21)
under30	-0.071	-0.079	0.027	0.029
	(0.73)	(0.81)	(0.21)	(0.22)
over64	-0.041	-0.038	-0.107	-0.115
	(0.41)	(0.38)	(0.79)	(0.81)
pcbirths	-0.159	-0.128	1.078	1.125
	(0.39)	(0.32)	(1.50)	(1.51)
exclincomepct	0.097	0.083	-0.660	-0.677
	(0.23)	(0.21)	(0.89)	(0.90)
unemplrate	-0.431	-0.420	-0.271	-0.266
	(1.90)	(1.88)	(0.74)	(0.72)
collegepct	0.078	0.057	0.056	0.067
	(1.17)	(0.85)	(0.58)	(0.64)
malepct	0.503	0.522	0.464	0.454
	(3.48)**	(3.65)**	(2.41)*	(2.29)*
homeownerpct	0.008	-0.007	0.071	0.075
	(0.10)	(0.08)	(0.65)	(0.66)
popdensity	-0.144	-0.141	-0.149	-0.151
	(4.33)**	(4.29)**	(3.37)**	(3.30)**
trend	0.893	1.011		
	(4.29)**	(4.61)**		
callattemptpct		0.131		-0.088
		(1.58)		(0.46)
Constant	11.707	10.871	75.980	78.444
	(0.51)	(0.48)	(2.02)*	(1.99)*
Observations	294	294	294	294

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 6. Two-way Fixed Effects and a One-way Fixed Effects with Trends

	(1)	(2)	(3)	(4)
	Base Year Dummies	Base Trends	Base Trends c_marg_95_fixed	Base Trends c_marg_95_fixed Additional Variables
lnaudnw	9.728	9.298	8.494	6.938
	(1.36)	(1.29)	(1.23)	(1.04)
filingrate	0.281	0.308	0.315	0.332
	(4.58)**	(5.65)**	(5.99)**	(6.45)**
fthresholdpct	0.888	1.111	1.131	1.247
	(2.14)*	(4.28)**	(4.27)**	(4.85)**
c_marg	-22.506	15.070		
	(0.67)	(0.62)		
depamountpct	1.168	1.204	1.220	1.399
	(1.48)	(1.60)	(1.62)	(1.88)
lnburden	10.473	7.039	6.827	6.252
	(3.40)**	(2.72)**	(2.71)**	(2.55)*
soleprops	0.403	0.666	0.643	0.705
	(0.84)	(1.50)	(1.44)	(1.62)
soleproptfs	-0.013	-0.018	-0.017	-0.018
	(1.42)	(2.10)*	(2.03)*	(2.13)*
paidprep	-0.147	-0.116	-0.117	-0.126
	(3.99)**	(3.34)**	(3.36)**	(3.75)**
singles	0.270	0.251	0.251	0.206
	(3.25)**	(3.09)**	(3.09)**	(2.11)*
under30	0.056	0.038	0.040	0.003
	(0.59)	(0.42)	(0.44)	(0.03)
over64	0.060	0.084	0.073	0.094
	(0.49)	(0.69)	(0.61)	(0.77)
pcbirths	0.782	0.809	0.846	0.484
	(3.63)**	(4.32)**	(4.58)**	(2.38)*
exclincomepct	-0.546	-0.583	-0.548	-0.483
	(1.27)	(1.50)	(1.41)	(1.27)
unemplrate	-0.603	-0.448	-0.450	-0.603
	(3.47)**	(3.29)**	(3.26)**	(3.99)**
trend		0.552	0.511	0.248
		(1.51)	(1.45)	(0.73)
tra86dum		-9.384	-9.695	-11.228
		(4.50)**	(5.08)**	(5.75)**
tratrend		1.288	1.305	1.463
		(5.20)**	(5.21)**	(5.59)**
c_marg_95_fixed			10.674	16.170
			(0.51)	(0.79)
collegepct				-0.019
				(0.26)
malepct				0.064
				(0.61)
homeownerpct				-0.082
				(0.96)
popdensity				0.102
				(3.48)**
gini_faminc				-3.727
				(0.22)
Constant	-49.943	-51.611	-50.076	-43.105
	(2.09)*	(2.22)*	(2.17)*	(1.77)
Observations	490	490	490	490

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

Table 7 presents measures of the predictive performance of the more and less parsimonious specifications presented in columns 3 and 4 of Table 6. These measures are based on the leave-one-out prediction approach described earlier. The results indicate a rather similar forecasting performance across the two specifications. In both cases, the average absolute deviation of the forecast from the true level of income represents a little more than 2 percent of income, and the coefficient of variation of the root mean-squared error (CV of RMSE) is approximately 3.6 percent. As discussed above, the use of ratio specifications such as those presented in Table 6 is rather controversial. We have therefore estimated alternative versions of these models in which the dependent variable has been specified in natural log rather than ratio form and, in one variant, where many of the ratio explanatory variables have been replaced with separate variables representing the numerators and denominators of the ratios. The parameter estimates from these specifications have been fairly comparable in terms of signs and statistical significance. Further, the predictive performance of these specifications have turned out to be quite similar to that of the original specifications based on ratios.

TABLE 7. Leave-One-Out Predictive Performance of Models in Table 6

Specification from Table 6 Column #	Absolute Deviation as a % of Income Reported	CV of RMSE
3	2.12	3.57
4	2.18	3.62

We now explore the forecasting performance of specifications estimated from a longer panel. Table 8 summarizes the estimation results of some selected specifications that have been estimated from data spanning the period from 1982 to 2007 (or 2004 for the specifications that include the Gini coefficient as an explanatory variable). As with Table 6, the first two columns of Table 8 respectively provide a two-way fixed effects specification and a one way fixed effects with trend terms specification of a parsimonious model of income reporting behavior. While many of the parameter estimates are comparable in sign and significance across these two specifications, they do produce conflicting estimates of the coefficients of the natural log of the audit close rate and the combined state-federal marginal tax rate. This discrepancy between the results of the two alternative specifications is consistent with similar findings presented above. In column 3 we verify that directly substituting the instrument for the combined marginal tax rate as an explanatory variable yields comparable results to those presented in column 2. That specification is extended to include some additional explanatory variables in column 4. Finally, column 5 presents the results of estimating the specification in column 4 using a two-way fixed effects specification rather than using trend terms. Once again, the use of two-way fixed effects yields more intuitive coefficient estimates for the audit and marginal tax rate variables.

To investigate whether this translates into improved predictive performance, we have employed the leave-one-out prediction methodology described in Section 4.3 for each of the specifications in columns 3, 4, and 5. In the case of the year dummy specification in column 5, we have used an autoregressive process of order 2 to forecast the value of each year dummy when the corresponding year is left out of the estimation sample. The leave-one-out forecasting results are summarized in Table 9. All of the specifications predict reasonably well out of sample, with an average absolute forecast deviation of less than 3% of income reported and a CV of RMSE of 4.3 to 5.1 percent. Overall, the predictive performance is slightly weaker for the longer panel specifications summarized in Table 9 than for the comparable shorter panel specifications summarized earlier in Table 7. Interestingly, the specification based on year dummies performs slightly less well than those based on trend terms.

We have also used the results of the last two specifications presented in Table 8 to develop one, two, three, and four step-ahead forecasts. These specifications were estimated using a sample period of tax year 1982 through tax year 2000, and the results were then employed to develop state level forecasts of income reported for tax years 2001 through 2004. These forecasts were compared against actual levels of income reported to

produce measures of the average absolute forecast deviation as a percentage of income reported and the CV of RMSE. The results are summarized in Table 10 below.

TABLE 8. Results of Estimation Using a Longer Panel

	(1)	(2)	(3)	(4)	(5)
	Base Year Dummies	Base Trends	Base Trends c_marg_95_fixed	Base Trends c_ marg_95_fixed Extra Variables	Year Dummies
lnaudnw	8.789	-3.791	-3.890	-9.124	4.881
	(2.36)*	(2.15)*	(2.20)*	(3.89)**	(1.42)
filingrate	0.310	0.350	0.350	0.393	0.366
	(10.21)**	(12.13)**	(12.04)**	(12.92)**	(12.22)**
fthresholdpct	0.718	0.576	0.552	1.063	0.994
	(3.94)**	(3.57)**	(3.59)**	(5.99)**	(5.51)**
c_marg	-12.561	8.787			
	(0.65)	(0.75)			
depamountpct	1.488	1.630	1.643	0.736	1.427
	(3.81)**	(3.99)**	(4.00)**	(1.65)	(3.46)**
lnburden	5.153	5.039	5.030	3.236	3.594
	(6.03)**	(6.50)**	(6.45)**	(3.72)**	(4.15)**
soleprops	0.525	-0.148	-0.188	-0.160	0.589
	(1.78)	(0.52)	(0.67)	(0.54)	(2.07)*
soleproptfs	-0.013	0.003	0.003	0.001	-0.014
	(2.33)*	(0.53)	(0.71)	(0.25)	(2.71)**
paidprep	-0.002	-0.064	-0.066	-0.069	-0.018
	(0.11)	(2.80)**	(2.92)**	(2.87)**	(0.77)
singles	0.163	0.210	0.209	0.135	0.114
	(3.47)**	(4.32)**	(4.28)**	(2.06)*	(1.94)
under30	0.029	-0.042	-0.043	-0.214	-0.077
	(0.56)	(0.85)	(0.86)	(3.80)**	(1.36)
over64	0.058	-0.007	-0.009	-0.056	0.073
	(0.96)	(0.11)	(0.16)	(0.81)	(1.15)
pcbirths	0.337	0.575	0.582	0.349	0.269
	(3.08)**	(5.34)**	(5.28)**	(2.91)**	(2.34)*
exclincomepct	-0.726	-0.513	-0.512	-0.855	-1.012
	(3.08)**	(2.11)*	(2.09)*	(3.25)**	(4.19)**
unemplrate	-0.241	-0.392	-0.396	-0.361	-0.186
	(2.70)**	(5.05)**	(5.11)**	(4.33)**	(2.14)*
trend		-0.453	-0.488	-0.969	
		(2.40)*	(2.80)**	(4.70)**	
tra86dum		-3.506	-3.625	-2.569	
		(3.35)**	(3.51)**	(2.46)*	
tratrend		0.626	0.653	0.603	
		(4.17)**	(4.59)**	(4.05)**	
dum92on		-4.875	-4.811	-3.757	
		(10.34)**	(10.47)**	(6.47)**	
c_marg_95_fixed			5.826	13.295	-2.461
			(0.58)	(1.26)	(0.16)
collegepct				0.140	0.055
				(3.17)**	(1.24)
malepct				0.189	0.182
				(2.69)**	(2.88)**
homeownerpct				-0.195	-0.127
				(3.70)**	(2.62)**
popdensity				0.045	0.037
				(5.79)**	(5.20)**
gini_faminc				21.313	4.573
				(2.46)*	(0.54)
Constant	-22.213	-17.122	-15.592	-6.228	-17.692
	(1.84)	(1.58)	(1.50)	(0.51)	(1.51)
Observations	1274	1274	1274	1127	1127

Absolute value of z statistics in parentheses * significant at 5%; ** significant at 1%

TABLE 9. Leave-One-Out Predictive Performance of Selected Models in Table 8

Specification from Table 8 Column #	Absolute Deviation as a % of Income Reported	CV of RMSE
3	2.47	4.64
4	2.42	4.31
5	2.78	5.10

TABLE 10. Step-Ahead Predictive Performance of Selected Models in Table 8

Specification from Table 8 Column #	2001		2002		2003		2004	
	Absolute Dev'n %	CV of RMSE	Absolute Dev'n %	CV of RMSE	Absolute Dev'n %	CV of RMSE	Absolute Dev'n %	CV of RMSE
4	2.11	3.09	2.78	3.84	4.22	6.06	2.96	4.49
5	2.45	3.90	2.95	4.23	3.28	4.96	2.22	3.92

Generally, the forecasting performance is reasonably strong. As expected, the performance tends to decline to some extent as one predicts further out, although the four-year ahead forecast performance for 2004 is comparable to the one-year ahead performance for 2005 in the specification involving year dummies.

We have also estimated variants of the specifications summarized in Table 8 that rely less on ratio variables. The results for these variants were qualitatively similar to those based on the ratio variables.

Conclusions

We have found that the forecasting performance of our preliminary models of overall income reporting is reasonably strong. This performance is slightly stronger for our shorter panel (1982-1991) than our longer panel (1982-2007), although the performance is reasonably good in both cases.

Overall, the results of our analysis indicate that it is possible to develop reasonably good forecasts of what overall state level income reporting behavior would be in the absence of a major innovation, such as a significant change in service level or quality. However, the lack of a reasonably lengthy time series of high quality state-level measures of IRS service activities limits the potential for our current models to predict how such activities influence reporting behaviors. Fortunately, compiling such data is a high priority for the IRS Service-Compliance Initiative going forward.

References

Balestra, P., and J. Varadharajan-Krishnakumar, "Full information estimations of a system of simultaneous equations with error component structure," *Econometric Theory*, 1987, 3: 223–246.

Dubin, Jeffrey A., "Criminal investigation enforcement activities and taxpayer noncompliance," *Public Finance Review*, 2007. 35: p. 500-529.

_____, M. Graetz, and L. Wilde (1990). "The Effect of Audit Rates on the Federal Individual Income Tax, 1977-1986," *National Tax Journal* 43:395-409.

Plumley, Alan H., "The Determinants of Individual Income Tax Compliance: Estimating the Impact of Tax Policy, Enforcement and IRS Responsiveness," *Internal Revenue Service, Publication 1916 (Rev. 11-96),* Washington, DC, 1996.

Wiseman, R.M. "On the Use and Misuse of Ratios in Strategic Management Research" in *Research Methodology in Strategy and Management,* Volume 5, ed. by D.D. Berg and D.J. Ketchen, Jr., pp. 75-110, Bingley, U.K.: Emerald Group Publishing, Ltd., 2009.

Endnotes

[1] RAS Office of Research (IRS), Brian Erard & Associates, and IBM Global Business Services, respectively. The views expressed in this paper are those of the authors, and do not necessarily reflect the views of the Internal Revenue Service.

[2] One exception is Alaska, which is excluded because of compatibility issues resulting from the need for all recipients of Alaska Permanent Fund Dividends—including children—to file federal tax returns. Also, Maryland and DC have been combined.

[3] See the listing of variable definitions provided at the end of the paper.

[4] Our results for this specification differ from Plumley (1996), as we use a more standard approach to instrumental variables estimation. The variable definitions are provided at the end of the paper.

[5] The audit start rate is defined as the number of audits started in a given year, expressed as a percentage of the total number of returns filed in the calendar year before the beginning of the audit. The audit close rate, which is the more standard measure, is defined as the number of audits completed in a given year, expressed as a percentage of the total number of returns filed in the calendar year before the closure. Plumley (1996) theorized that the information about audits that gets "rippled" into the general population at the start of an audit affects people's perception of the *probability* of an audit, while the information communicated when the audit is closed has more to do with the *consequence* of an audit. As a practical matter, since the two audit rates are so highly correlated, they appear to be fairly interchangeable in an analysis such as this, so it makes sense to use the audit closure rate, which is more readily available.

Appendix

Definitions of Included Variables

Variable	Definition
lnaud	Natural log of one plus the audit start rate
filingrate	Returns Filed/ Returns Required
fthresholdpct	Income below filing threshold among all potential returns as a % of PI
mtr15k	Marg. tax rate @ $15K taxable income (weighted by Singles & Marrieds)
mtr57k	Marg. tax rate @ $57K taxable income (weighted by Singles & Marrieds)
childexemptspct	Value of allowed dependent child exemptions/Personal Income
lnburden	Natural log of average burden (in dollars) based on the IMF Population
soleprops	% of Potential Returns having non-farm proprietorship income
soleproptfs	SoleProps x percentage of non-farm employment in Trade, Finance & Service sectors
paidprep	% of Returns Filed prepared by paid practitioner
lnirp	Natural log of information returns matching
lncid	Natural log of criminal investigations
tps_callspc	Taxpayer service calls handled per thousand of population
tps_retpreppc	Returns prepared by taxpayer service calls per thousand of population
singles	% of Potential Returns likely to qualify for Single filing status
under30	% among Potential Returns under age 30
over64	% among Potential Returns over age 64
pcbirths	Number born per thousand of population
exclincomepct	Excluded Income/Personal Income
unemplrate	Unemployment Rate
trend	Trend
tra86dum	TRA86 dummy variable equal to one for years subsequent to 1986
tratrend	Interaction of trend and tra86dum (trend times tra86dum)
depamountpct	Total Value of the Dependent exemption as a percent of personal income
c_marg	Combined Marginal Tax Rate Based on the Actual distribution of Reported Income / IMF Population
lnaudnw	Natural log of the audit close rate
dum91	Dummy variable for 1991
dum92	Dummy variable for 1992
trendsq	Trend squared
collegepct	% among Potential Returns having at least some college
malepct	% of Potential Single & HeadHhld Returns associated with males
homeownerpct	% of Potential Returns associated with homeowners
popdensity	Population density
gini_faminc	Smoothed state-level gini coefficient based on family income
cidsentrate	Total sentenced violations as a percentage of population
c_marg_95_fixed	Combined marginal tax rate based on 1995 national distribution of reported income / IMF population
hoursrvrate	National measure of number of hours worked by IRS employees in taxpayer-facing service occupations
complexity	National measure of the complexity of individual returns based on word counts of IRS individual income tax code
callattemptpct	Total call attempts as a percent age of the overall state population

2

Individual Compliance Behavior

Alm ◆ Bloomquist ◆ McKee
Phillips

Comparing Student and Non-Student Reporting Behavior in Tax Compliance Experiments

James Alm, Tulane University; Kim M. Bloomquist, Internal Revenue Service; and Michael McKee, Appalachian State University

Laboratory experiments have been increasingly used to examine various issues in tax compliance. The use of laboratory experiments in economics began in the early 1960s with the establishment of a well-defined framework for experimental work by Smith (1976, 1982), and laboratory methods are now widely accepted as a methodological approach in the analysis of theory and policy, especially of behavioral economics. For comprehensive surveys of experimental methods, see Davis and Holt (1993) and Kagel and Roth (1995).

Tax compliance is an area that seems especially amenable to laboratory experiments, given limitations in empirical approaches based on field data. As discussed in more detail later, theoretical models are not able to incorporate fully, appropriately, or tractably many factors deemed relevant to the individual compliance decision; also, these modeling efforts can benefit from the behavioral insights obtained through laboratory investigations that illuminate many of the factors relevant to the individual compliance decision. Empirical studies of tax compliance using field data are plagued by the absence of reliable information on individual compliance decisions: it is difficult to measure—and measure accurately—something that by its very nature people want to conceal, and it is difficult to control in econometric work for the many unobservable factors that may affect the compliance decision. In contrast, laboratory methods allow many factors suggested by theory to be introduced in experimental settings. Also, experiments generate precise data on individual compliance decisions, which allow econometric estimation of individual responses in ways that are simply not possible with field data. Indeed, laboratory methods have examined a wide range of factors in the compliance decision, factors that have not proven amenable to either theoretical analyses or empirical analyses with field data.

However, laboratory studies of compliance are also sometimes viewed with some skepticism. The most common criticism of experimental investigations of tax compliance behavior is that the student subjects typically used in experiments may not be representative of taxpayers. Undergraduates may have little experience with filing tax returns, and their economic and demographic backgrounds may differ from that of taxpayers. It is this issue that we examine here.

We present evidence that relates directly to the use of student subjects. In particular, we compare the experimental responses of *student subjects* to the experimental responses of *non-student subjects*, in *identical experiments*; that is, do students behave differently than non-students in identical experiments? These data are generated from a series of experiments conducted as part of an Internal Revenue Service (IRS) funded research program, in which student subjects and non-student subjects participated. We examine the mean levels of compliance, the frequency distribution of compliance rates, and the econometrically estimated behavioral responses. While we find that the mean *levels* of compliance of students are not always the same as non-students, the behavioral *responses* of students in laboratory experiments to policy innovations are largely the same as non-students in identical experiments. Also, the frequency distributions of individual compliance rates are virtually identical for students and non-students. (Note that in other work we compare experimental data versus non-experimental data from the National Research Program of the IRS, in an attempt to answer whether students in experiments behave differently than non-students in naturally occurring settings. These other results address explicitly the issue of generalizing from experiments to the naturally occurring world (e.g., the "external validity" of tax compliance experiments). These results are not presented here, but these comparisons also indicate largely similar patterns. See Alm, Bloomquist, and McKee (2011) for a detailed

analysis of all student versus non-student and experimental versus non-experimental comparisons.) Our results largely confirm that the observed behaviors of student and non-student subjects are qualitatively and quantitatively similar.

The Uses and Misuses of Experimental Economics

Experimental economics involves the creation of a real microeconomic system in the laboratory, one that parallels the naturally occurring world that is the subject of investigation and one in which subjects (usually students) make decisions that yield individual financial payoffs whose magnitude depends on their decisions. The essence of such a system is control over the environment, the institutions, the incentives, and the preferences that subjects face. Of these, control over preferences is particularly crucial. As emphasized by Smith (1976), "[s]uch control can be achieved by using a reward structure to induce prescribed monetary value on actions."

Why use experimental methods? On some level, the use of experimental methods derives from a fundamental problem with economics. Like other sciences, economics is based on the development of theory and on the ability of that theory to explain observed activities. However, unlike some other sciences, especially the natural sciences, economics faces substantial difficulties in empirically testing the predictive power of its theories using data from the naturally occurring world. Given the dizzying array and complexity of forces that operate in market (and non-market) systems, economists can never be quite certain that they are "holding constant" the many factors that may be driving individual choices, so that they can focus on the "true" driving factors that are the object of empirical testing. Methods for achieving such identification have become increasingly sophisticated over time, especially with the use of so-called "natural experiments" and "controlled field experiments". Even so, there are few instances in which such identification is uncontroversial.

There are of course numerous avenues for testing the predictions of economics, aside from experimental methods, including the use of naturally occurring field data, hypothetical choices, natural experiments, and controlled field experiments. Even so, experimental methods have often been a common approach, and economists have increasingly begun to emulate the methods of natural scientists by conducting carefully controlled laboratory experiments.

Economics generally and public economics specifically have profited from the use of laboratory experiments, for several reasons. Econometric data on research questions obtained from the naturally occurring world can be unreliable, can fail to show the variation or distinctions of interest, or can fail to provide sufficient identification to discern "cause and effect". Indeed, in some cases data simply cannot be assembled outside the laboratory because the real world setting of relevance does not exist. For its part, theoretical analyses often cannot incorporate fully, appropriately, or tractably many relevant factors.

What can experiments do? Unlike standard theoretical work, experiments are not as constrained by the same degree of simplification required in analytical studies, which allows the impact of numerous factors not amenable to theoretical work to be examined precisely and unambiguously in a controlled environment. Unlike traditional empirical work based on naturally occurring data, experiments generate data under settings in which there is control over extraneous influences. Laboratory experiments also provide a controlled environment that allows one to examine the mechanisms of interest, as well as the changes in these environments and institutions, in isolation from each other.

Given the limitations of theoretical and econometric work, there are, we believe, compelling reasons for the use of experiments, as an *additional* (and not as the *only*) methodological tool, in large part because laboratory experiments give a researcher the twin advantages of *control* (including *data generation* and *replicability* from this control) and *flexibility*.

However, despite the demonstrated usefulness of experimental methods, there are sound reasons for caution in interpreting and generalizing experimental results. Perhaps the most common criticism of experimental economics is that the student subjects typically used may not be representative of taxpayers (Levitt and List, 2007). Although there is now much evidence that the experimental responses of students are seldom different than the responses of other subject pools (Plott, 1987), there are no comparisons of student versus non-student behavior in the specific context of tax compliance. It is this issue that we examine here.

Some Simple Comparisons of Students and Non-students in Identical Experiments

We consider the responses of student subjects versus non-student subjects in identical experiments. These comparisons are based upon results in Alm, Cherry, Jones, and McKee (2010, 2011) who conducted experiments designed to investigate the compliance behavior of individuals under various policy initiatives. In both studies the subject pool consisted of both students and non-students (e.g., university staff and faculty), and we report here a comparison of students and non-students in these different experimental settings. The first setting investigates the effects on tax reporting of the provision of information services by the tax agency in an environment in which subjects may not know with certainty their true tax liability. The second setting introduces positive inducements via social programs as an incentive to truthfully report tax liabilities; these social programs include an income tax credit (in which receipt requires that the subject must file a tax return) and unemployment benefits (in which benefits are a positive function of past reported taxes). We first discuss the experimental design of these experiments, and then we present the comparison of student versus non-student responses in order to answer the question of whether students behave differently than non-students in identical experiments.

The experimental setting implements the fundamental elements of the voluntary reporting system of the individual income tax in most countries. Participants earn income by performing a task, and they self report tax liability to the tax authority. At the time of reporting only the individual knows his or her true level of tax liability, and the subject can choose to report any amount from zero on up. An audit occurs with an announced probability, and any unreported taxes are discovered. If the participant has underreported the tax liability, then both the unpaid taxes and a penalty are collected. This process is repeated over a number of rounds each representing a tax period. Participants are informed that they will be paid their after-tax earnings at the end of the experiment, converted from lab dollars to U.S. dollars at a fixed and announced conversion rate.

Participants are informed, with certainty, of the audit probability, the penalty rate, and the tax rate. The tax rate is set at 35 percent for all sessions; the penalty rate is also fixed for all sessions at 150 percent (i.e., the participants must pay unpaid taxes plus a penalty of 50 percent of unpaid taxes if audited). The audit probability is varied once within the session, and the participants are told that there is zero probability of audit if no tax form is filed. There is no public good financed by the tax payments in order to focus subject attention on the tax setting.

Into this setting, various policy innovations are introduced in the different studies of Alm, Cherry, Jones, and McKee (2010, 2011). A first set of experiments investigates the effects of taxpayer information services on compliance decisions (Alm, Cherry, Jones, and McKee, 2010). Here the tax reporting decision is "complicated" through the introduction of uncertainty regarding the true tax liability, and then information services are provided by the "tax administration" that partially or fully resolve the uncertainty, thereby allowing subjects to compute more easily their tax liabilities. Complicating factors include both a tax deduction (a reduction in taxable income) and a tax credit (a reduction in tax owed, comparable to the U.S. Earned Income Tax Credit), each of which is conditional upon filing. The tax deduction is set at 15 percent of income, and tax credit is introduced in which the credit starts at a given level and declines at a stated rate as income increases. These factors complicate the tax reporting decision. Also, the exact levels of the deduction and the credit are, in some settings, uncertain to the taxpayer at the time of filing, and this uncertainty adds an additional level of complications. These uncertainties on the credit and the deduction are implemented via mean-preserving spreads (with a uniform distribution) in each, where the participants are informed of the means of the allowed credit and deduction and the ranges for each. When information services are provided, information is complete, accurate, and costless to the participant.

A more direct set of positive inducements is investigated in Alm, Cherry, Jones, and McKee (2011). Here positive inducements for filing are introduced in several alternative treatments. In one treatment tax credits are introduced that are available to participants but only to those who file a tax return. In a second treatment a "social safety net" (e.g., unemployment replacement income) is present in which individuals face some probability of unemployment but also in which unemployment replacement income may be provided, with any benefits again conditional upon past filing behavior. There is a known probability of unemployment,

and, if the individual becomes unemployed and earns no income, then they are unemployed for two periods. Unemployment replacement income is received only if the individual has filed a tax return in each of the two previous periods. Other features of the experimental design are identical to Alm, Cherry, Jones, and McKee (2010).

Table 1 summarizes the experimental design of Alm, Cherry, Jones, and McKee (2010, 2011), with the top panel showing the information services design and the bottom panel showing the positive inducements design. Treatment T1 provides a baseline setting that entails no uncertainty and no tax authority information. The second treatment (T2) introduces tax liability uncertainty, in which participants face uncertainty regarding their allowed deduction and tax credit. The third treatment (T3) entails the same uncertainty as in the second treatment, but introduces the option of resolving the uncertainty by receiving information from the tax authority; that is, participants in this treatment are able to click on a button to reveal the true levels of the deduction and the tax credit. In Table 2, treatment T4 establishes a baseline with no positive inducements. In T5 a tax credit is introduced, and in T6 an individual may be able to collect income benefits while unemployed. The unemployment benefits are based on the average filed earnings for the previous two periods. Thus, an individual who chooses to not file can earn no benefits. See Alm, Cherry, Jones, and McKee (2010, 2011) for a detailed discussion of the experimental designs.

TABLE 1. Experimental Treatments

Information Services

	Information Service Provided?	
Tax Liability Uncertain?	**Yes**	**No**
No	T1	—
Yes	T2	T3

Positive Inducement via Social Programs

Positive Inducements Provided?		
No	**Yes, via Tax Credit**	**Yes, via Unemployment Benefits**
T4	T5	T6

TABLE 2. Aggregate Results by Experimental Treatment by Subject Pool

	Mean Reporting Compliance Rate		
Treatment	**All**	**Staff**	**Students**
Information Services			
No Uncertainty (T1)	0.673	0.795	0.618
Uncertainty—No Information (T2)	0.621	0.571	0.689
Uncertainty—Information (T3)	0.704	0.657	0.768
Positive Inducements via Social Programs			
No Positive Inducements (T4)	0.483	0.444	0.504
Tax Credit (T5)	0.599	0.678	0.487
Unemployment Benefits (T6)	0.681	0.709	0.655

The dedicated experimental lab consists of 25 networked computers, a server, and software designed for this series of experiments. Sessions were conducted at a major state university, using both students and staff as participants. Recruiting was conducted using the Online Recruiting System for Experimental Economics (ORSEE). The participant database was built using announcements sent via email to all students and staff. Participants were invited to a session via email, and were permitted to participate in only one tax experiment, although other experimental projects are ongoing at the time and participants may have participated in other types of experiments. Only participants recruited specifically for a session were allowed to participate, and no participant had prior experience in this experimental setting. Methods adhere to all guidelines concerning the ethical treatment of human subjects.

Of most importance, participants included both students and non-students, thereby allowing one aspect of the external validity of experiments to be examined: do students behave differently than non-students in identical experiments? Students were recruited from the pool of undergraduate students at a U.S. public university. Non-students were recruited from faculty and staff at this same university. The student portion of the subject pool covered a very broad range of year in studies and major, and no single major exceeded 8 percent of the pool. The staff pool was similarly diverse, covering all levels of support staff, non-academic professional staff, and faculty. Also, the compensation varied for students versus non-students. For students the rate was 80 lab dollars to 1 U.S. dollar; faculty and staff participants received a higher exchange rate to reflect their higher outside earnings, with a conversion rate of 50 lab dollars to 1 U.S. dollar. Earnings averaged $18 for student subjects. The average payoffs for faculty and staff were $28.

Four hundred forty nine individuals participated in a session in one of the two sets of experiments. In the sessions designed to investigate the role of tax information services on reporting (T1 through T3), there were 131 subjects, 54 percent of whom were students. In the sessions designed to investigate the effects of positive inducements (T4 through T6), there were 318 subjects (68 percent were students). Table 2 reports the aggregate figures for reporting behavior only by treatment and by subject group.

Treatments T1 – T3 concern the taxpayer information services design. These aggregate numbers indicate that uncertainty concerning tax liability results in lower reporting compliance rates but that providing information that resolves the uncertainty increases reporting. With tax liability uncertainty, the overall reporting compliance rate is 0.621 (T2), which is statistically lower than the 0.673 rate without uncertainty (T1). Further, reporting compliance significantly increases when information services are provided in the uncertain setting (T3), or 0.704 versus 0.621.

When these aggregate levels of compliance are broken down by subject type, we see some differences by subject type in reporting compliance rates. In T1, the mean reporting compliance rate of staff is higher (0.795) than the reporting compliance rate of students (0.618), a difference that is statistically different. Similarly, the mean compliance rates of staff versus students are also different for the other two treatments, T2 and T3, although in these latter two treatments students report at higher rates than staff.

However, despite the somewhat different *levels* of compliance of the two subject pools, the *changes* in compliance rates in response to the treatment effects are similar in both pools of subjects. Comparing T2 and T3, we observe that the changes in compliance rates in response to the information services treatments are similar in magnitude and sign. For both subject pools, the provision of information that resolves tax liability uncertainty leads to an increase in the mean reporting rate, by 8.6 percentage points for staff and by 7.9 percentage points for students (or T3 versus T2). The introduction of uncertainty has different effects on these responses by subject pool (or T2 versus T1), reducing reported income for staff and increasing reported income for students; even so, the change in mean compliance rates for students is not statistically significant in these two treatments.

The positive inducements treatments demonstrate a similar pattern. Treatment T4 provides the baseline setting for the investigation of the tax reporting effects of positive inducements. When looking at all subjects combined, the provision of the tax credit (T5) and unemployment benefits (T6) leads in both cases to higher tax reporting relative to T4. Also, both the student subjects and the staff subjects respond positively to the positive inducements. From the results in Table 2, it appears the staff responses to these inducements are greater than for the student subjects, but the qualitative treatment effects results are similar.

Also, the frequency distributions of individual compliance rates are virtually identical for students and non-students in the uncertainty sessions and in the positive inducements sessions.

It should be noted that there are several significant differences between the subject pools, and a simple comparison of means does not hold these differences constant. For example, the different pools are compensated at different rates, with staff compensated at roughly twice the rate of student subjects. Also, the average ages of the two groups are different, with the average age of staff exceeding 30 years and the average age of students slightly exceeding 20 years. There are other differences between the pools as well.

To hold constant these differences, we use a conditional analysis at the individual level to re-examine the initial impressions from the aggregate data. For the entire sample, staff only, and students only, we estimate the effects of various subject features and design parameters on reporting behavior, while holding other factors constant. Our basic specifications estimate individual reporting of each subject by round as a function of subject demographic characteristics (e.g., subject age, subject sex, subject own preparation of tax returns, subject claimed as a dependent on parental tax returns), subject variables that change by round (e.g., income, accumulated earnings, audit probability), and session characteristics (e.g., indicator variables that signify the presence of uncertainty about tax features, of agency-provided information, of a tax credit that the subject can claim on filing a tax report, or of a safety net that (partially) makes up for income lost due to unemployment). We estimate these specifications using Tobit estimation procedures. These detailed econometric results are not presented here, but they confirm even more strongly our initial impressions from the aggregate data. In particular, we see that the *changes* in compliance behavior of the two groups are quite similar, as measured by the estimated coefficients on the policy innovations, even if the average *levels* of compliance differ as in Table 2.

Conclusions

Our results indicate that the experimental behavior of students is often—although not always—similar to the experimental behavior of other subject pools. Although the *levels* of compliance may differ between student and non-student subjects in identical experiments, the *changes* in compliance behavior of these pools largely parallel each other, especially when potentially confounding influences are controlled in a regression framework. Further, the frequency distributions of individual compliance rates for students versus staff are very much the same. Other results that compare experimental data versus non-experimental data from the National Research Program of the IRS also indicate largely similar patterns. Again, see Alm, Bloomquist, and McKee (2011) for complete discussion and analysis of all results.

In sum, our results are consistent with other experimental studies that demonstrate that student and non-student subjects behave and, especially, respond similarly. Indeed, our evidence is consistent with the broader notion that there is also no reason to believe that the cognitive processes of students are different from those of "real" people, at least in the context of tax compliance and in the comparison of changes in behavior.

References

Alm, James, Kim M. Bloomquist, and Michael McKee (2011). "On the External Validity of Tax Compliance Experiments." Tulane University Working Paper. New Orleans, LA.

Alm, James, Todd Cherry, Michael Jones, Michael McKee (2010). "Taxpayer Information Assistance Services and Tax Reporting Behavior." *Journal of Economic Psychology* 31 (4): 577-586.

Alm, James, Todd Cherry, Michael Jones, Michael McKee (2011). "Social Programs as Positive Inducements for Tax Participation." Tulane University Working Paper. New Orleans, LA

Davis, Douglas D. and Charles A. Holt (1993). *Experimental Economics*. Princeton, NJ: Princeton University Press.

Kagel, John H. and Alvin E. Roth, eds. (1995). *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.

Levitt, Steven D. and John A. List (2007). "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *Journal of Economic Perspectives* 21 (2): 153-174.

Plott, Charles R. (1987). "Dimensions of Parallelism: Some Policy Applications of Experimental Methods." In *Laboratory Experimentation in Economics: Six Points of View*, edited by Alvin E. Roth, 193-229. New York, NY: Cambridge University Press.

Smith, Vernon L. (1976). "Experimental Economics: Induced Value Theory." *The American Economic Review Papers and Proceedings* 66 (2): 274-279.

Smith, Vernon L. (1982). "Microeconomic Systems as an Experimental Science." *The American Economic Review* 72 (5): 923–955.

Reconsidering the Deterrence Paradigm of Tax Compliance[1]

Mark D. Phillips, Department of Economics, University of Chicago, Internal Revenue Service

An extensive literature on the determinants of tax compliance began nearly forty years ago with the theoretical treatments of Allingham and Sandmo (1972) and Yitzhaki (1974). These theories represent applications of the Becker (1968) economic theory of crime in which a rational, expected-utility-maximizing agent chooses how much income to self-report to the government by comparing his consumption when noncompliance is or is not detected. The theory has been dubbed the "deterrence" paradigm as it assumes that taxpayers inherently wish to pay no tax liability and are "deterred" from doing so solely by the risk of audit, detection, and penalty. The model's stylized taxpayer is essentially identical to a gambler who chooses how much to wager based on the odds and payouts of the noncompliance bet.

While the IRS estimated a sizeable $345 billion tax gap for tax year 2001 ($197 billion for individual income tax, see IRS (2007)), one might predict a much greater amount under the classical deterrence theory. Therefore, while the deterrence paradigm represented economists' initial attempt towards understanding compliance, it has fallen out of favor in recent decades. According to Slemrod (2007), "the dismissive argument goes as follows: given the average probability of audit…, the penalties typically assessed for noncompliance…, and what we know about the degree of risk aversion from other contexts, noncompliance should be much, much higher than it apparently is."[2] In other words, the odds and payoffs of the gamble appear so favorable that optimal risk-taking motivations do not appear sufficient to explain observed noncompliance. More assertively, Kirchler *et al*. (2010) states "though the [deterrence theory] provides useful tools for tax policy … empirical evidence for its validity is rather weak."

Given these observations, Alm *et al.* (2010) summarizes that "the puzzle of tax compliance behavior may be why people pay taxes, not why they evade them." To solve this puzzle, economists have offered up many alternative explanations. These alternatives have been broadly categorized as belonging to the "behavioral" paradigm, a catch-all categorization of all other factors that involve "more than amoral cost-benefit calculation." (Slemrod (2007)) Such factors include (but are not limited to) guilt, adherence to social norms, or overestimation of actual audit and penalty rates.

In the paper, I offer an alternative explanation for the relatively low observed levels of noncompliance. However, the current explanation falls within the context of the deterrence paradigm, expanding the classical theory to more realistically model the audit and detection regime that taxpayers face. In particular, I account for the fact that taxpayers do not face a constant likelihood of audit and noncompliance detection; instead, the probability of audit and detection depends on the taxpayer's noncompliance itself.

First, different types of income have different degrees of inherent noncompliance detectability. In particular, I distinguish between "matched" and "unmatched" income, "matched" being income that has been reported to the IRS by a third-party and "unmatched" being that which has not. Any misreporting of matched income leads to a significantly higher probability of detection relative to a correct report of matched income. Misreporting of unmatched income may instead carry a relatively small probability of detection. Second, even if a taxpayer misreports only unmatched income, greater amounts of underreporting increase the likelihood of audit and detection. This arises from the fact that audits are targeted towards those taxpayers who are most likely to have been noncompliant and have the greatest expected amount of noncompliance. Furthermore, increases in underreporting require that taxpayers transition from underreporting income that is relatively undetectable during an audit towards income that is more readily detectable. In Section 2, I discuss in more detail the mechanisms under which audit and detection rates depend on the taxpayer's noncompliance itself.

These realistic features of the examination environment imply that the probability of audit and detection is increasing with respect to the amount that a taxpayer underreports his true tax liability. This in turn implies that taxpayers have some ability to affect the likelihood of detection via their noncompliance behavior; therefore, the stylized taxpayer of the deterrence paradigm has an additional incentive for compliance, even in the presence of low audit and penalty rates. Even though the "gamble" has generous payoffs and odds, the taxpayer may strategically gamble less in order to lessen the suspiciousness of his reported return and in turn tilt the odds in his favor. In Section 3, I discuss how the predictions of the deterrence paradigm change when properly accounting for an environment of targeted audits and endogenous detection rates. In net, the empirical evidence is much stronger in support of this expanded version of the classical deterrence theory, in which case the lessons, motivations, and policy prescriptions of the deterrence paradigm warrant greater authority.

Targeted Audits and Endogenous Detection

Implicit in the criticism of the classical model is an assumption that audit probabilities are constant and exogenous. However, this assumption corresponds to a tax agency audit strategy that consists merely of drawing taxpayers' names out of a hat, with each taxpayer facing the same likelihood of audit independent of his self-reported tax filings or any other characteristics that the tax agency observes. This assumption is unrealistic as the IRS (and other tax agencies around the world) devotes significant efforts towards developing algorithms that successfully target audits towards those taxpayers who are most likely to have underreported liability and to have underreported by the greatest expected amounts.

One straightforward way in which audit probabilities depend on the taxpayer's noncompliance itself relies on the distinction between matched and unmatched income. While misreporting of unmatched income may yield a relatively low probability of audit, the same cannot be said for misreporting of matched income. In the U.S., the IRS's Document Matching Program (or Automated Underreporter Program) uses computer automation to find discrepancies between an individual's reported return and information it has received from third-parties. All returns undergo this document matching, and even when a mismatch is detected and pursued, it is usually cheaper than an in-person audit. Therefore taxpayers face a discrete, large increase in detection probability when matched income underreporting goes from \$0 (i.e. the taxpayer correctly reports matched income) to some strictly positive amount.

Even if the taxpayer correctly reports matched income, greater amounts of unmatched income underreporting also increase the likelihood of audit. This results from the fact that the tax agency observes several variables upon which it can base its targeted audit strategy. The IRS employs both the Discriminant Function (DIF) score that "rates the potential for change, based on past IRS experience with similar returns" as well as the Unreported Income DIF (UIDIF) score that "rates the return for the potential of unreported income." (See IRS press release FS-2006-10.) The IRS is unsurprisingly secretive regarding the inputs and internal mechanism of the DIF and UIDIF scoring, but one can imagine reasonable and intuitive examples of how these procedures might work. For instance, consider a taxpayer who earns \$5,000 in third-party-reported interest income and \$95,000 in unmatched sole proprietor income, for a total income of \$100,000. The taxpayer in turn debates whether to report the full \$95,000 of sole proprietor income, or perhaps only \$45,000 of it. If prior audits have shown the tax agency that interest income typically accounts for 5% of a taxpayer's true total income, the report of \$45,000 may appear more suspicious and face a higher probability of audit than the report of \$95,000.

In this example, the probability of audit depends directly on the taxpayer's self-reported income, as this amount is *ex ante* observable to the IRS. The expectation is that the probability of audit is decreasing in a taxpayer's self-reported income, but this in turn means that the probability of audit is increasing in the amount of underreporting (holding constant the taxpayer's true income).[3] In addition to the taxpayer's self-reported income, the tax agency has access to many other observable characteristics upon which it can base its targeting. It also seems likely that more egregious noncompliance is more likely to lead to mistakes and inconsistencies across the tax return, further increasing suspicion. In sum then, there are several reasons to expect that the probability of audit is generally increasing in the amount of underreporting.

Furthermore, the probability of detection during an audit is also likely to be increasing in the amount of the underreporting. For instance, a taxpayer might start underreporting income paid in cash, with a minimal

corresponding "paper trail," and eventually transition to underreporting income paid by credit card, for which a well-documented "paper trail" exists. Furthermore, if higher levels of underreporting lead to increased suspicion on the part of the tax agency, it may in turn conduct its audit with increased intensity and ultimately detect a greater portion of the noncompliance.

Some evidence of the IRS's ability to effectively target noncompliant returns is given by the fact that 63% of a weighted random sample of taxpayers were found to have correctly reported Form 1040 Total Income (see Phillips (2011)[4]), while the no-change rate among the IRS's targeted operational audits was only 19% (see IRS (2002)). However, it is certainly not the case that any tax agency can perfectly identify, target, and detect even the most egregious instances of underreporting. Instead, the claim is simply that more egregious noncompliance is relatively more likely to result in audit and detection compared to less egregious noncompliance. Even if significant underreporting results in a relatively low audit and detection probability *level*, a strategic incentive for partial compliance remains intact so long as additional compliance results in a *relatively* lower probability.

What Does the Expanded Deterrence Theory Predict?

Properly accounting for targeted audits and endogenous detection rates, the predictions of the deterrence theory differ significantly from when a constant, exogenous probability is assumed. First, the high probability of detection associated with matched income misreporting results in a prediction that taxpayers correctly report matched income. This prediction is consistent even with aggregate data. For instance, IRS (2007) estimates only a 1.2% Net Misreporting Percentage (NMP) for wages, salaries, and tips, 4.5% NMP for income line items subject to substantial information reporting (interest income, dividend income, state income tax refunds, pensions and annuities, unemployment compensation, and Social Security benefits), and 8.6% NMP for income line items subject to some information reporting (partnership income, S-Corp income, estate and trust income, alimony income, capital gains, deductions, and exemptions).[5,6]

Since the deterrence paradigm's predictions regarding matched income are well aligned with observed compliance, we must instead look to unmatched income for evidence that contradicts the theory. IRS (2007) estimates a 53.9% NMP of income items subject to little or no information reporting (Form 4797 income, other income, nonfarm proprietor income, farm income, rents and royalties, and total statutory adjustments). While 53.9% represents a considerable amount of underreporting, aggregate measures of the gamble's payoffs and odds appear so generous that one would predict an even greater NMP.[7]

For example, consider a taxpayer who is choosing how much unmatched income to underreport under the following parameterization of the noncompliance gamble: the probability of audit and detection is exogenously fixed at 11.6%, the tax rate is 35%, and the penalty rate is 75% on unpaid tax liability.[8] Each marginal dollar of underreporting results in the taxpayer retaining 35 cents in unpaid tax liability. With 88.4% probability, the taxpayer is never contacted by the tax agency and he keeps the 35 cents. With 11.6% probability however, the taxpayer is audited, pays back the 35 cents in unpaid tax liability, and pays an additional 35 * .75 = 26.25 cents in penalty. In net, the dollar of underreporting has a large, positive expected payoff of -26.25 * .116 + 35 * .884 = 27.895 cents. In the case of the risk-neutral taxpayer, the deterrence theory (with exogenous audit and detection probability) predicts 100% unmatched income underreporting so long as this net expected payoff is positive, as is the case under most any reasonable parameterization.[9]

However, let us now consider the case where more egregious noncompliance is more likely to result in audit due to the tax agency having a targeted audit strategy. For instance, assume that the taxpayer faces the average 11.6% six-year probability of audit (2.0% annual probability) only if he underreports 57% of his unmatched income (57% being the IRS (2007) estimated NMP for Schedule C sole proprietor income). If he instead underreports 100% of his unmatched income, he faces a hypothetical 35% six-year probability of audit (6.9% annual probability). While a 6.9% annual audit probability is significantly larger than the aggregate annual estimated probability of 2.0%, bear in mind that it applies only to a taxpayer who underreports 100% of his unmatched income. It is not obvious to the author that this is unreasonable for, let's say, a taxpayer who earns $100,000 per year in unmatched sole proprietor income and reports none of it.

If the risk-neutral taxpayer simply faced a constant 35% audit rate, the terms of the gamble are still generous enough that 100% underreporting is predicted: the net expected payoff of each dollar of underreporting

is -26.25 * .35 + 35 * .65 = 13.5625 cents. However, even though the last marginal dollar of underreporting has a positive expected payoff, the taxpayer actually has the opportunity to lessen his underreporting and in turn increase the odds that his remaining underreporting goes undetected. Specifically, if the taxpayer underreports 100% of his unmatched income of *U*, he consumes *U* with 65% probability and *U* – 1.75 * .35 * *U* = .3875 * *U* with 35% probability, for an expected consumption of .35 * .3875 * *U* + .65 * *U* = .785625 * *U*. If he instead underreports only 57% of *U* (and correspondingly reports 43% of *U*), he pays .43 * .35 * *U* = .1505 * *U* in self-reported tax liability, and therefore consumes only *U* - .1505 * *U* = .8495 * *U* if he goes unaudited. If he is audited, he now pays the penalty only on the 57% of *U* which he failed to self-report, for a consumption of *U* - .1505 * *U* – 1.75 * .35 * .57 * *U* = .500375 * *U*. The taxpayer's partial compliance increases the likelihood of the "preferred" state of non-audit and the taxpayer's expected consumption is .116 * .500375 * *U* + .884 * .8495 * *U* = .8090015 * *U*, which is greater than the .785625 * *U* consumption he could have expected under 100% underreporting.

This simplistic example is meant to demonstrate that a taxpayer who is motivated solely by "amoral cost-benefit calculation" may find it beneficial to strategically sacrifice some underreporting in order to appear less suspicious and in turn increase the likelihood that his remaining underreporting goes undetected. Furthermore, the previous example should not be interpreted as saying that *all* taxpayers with unmatched income are predicted to underreport less than 100%. Instead, the deterrence theory with targeted audits and endogenous detection predicts the existence of two types of taxpayers: those who underreport 100% of unmatched income and those who underreport only a portion of unmatched income.

There are two primary factors that determine whether the taxpayer is willing to strategically deviate from 100% underreporting. The first factor is the degree to which marginal increases in compliance decrease the joint audit and detection probability. For instance, if the audit rates are 35% for 100% underreporting and 11.6% for 57% (as above), the reduction in audit probabilities is sufficient to incentivize the partial compliance. On the other hand, if the taxpayer reduces his audit rate from only 12% to 11.6% by switching from 100% to 57% noncompliance, it is unlikely that the 0.4% reduction in probability is worth the 43% decrease in underreporting.

The second factor is the amount of unmatched income that the taxpayer possesses. A taxpayer's willingness to deviate from 100% underreporting will crucially depend on the *difference* in his consumption when he is or is not audited. This stems from the fact that the benefit of a marginal increase in compliance is the increase in likelihood that the "preferred" no-audit state occurs and off-setting decrease in likelihood that the audit state occurs. When the taxpayer considers underreporting 100% of unmatched income, the difference in consumption between these two states is larger when unmatched income is larger.

The two factors each suggest that low-unmatched-income taxpayers are more likely to underreport 100%, while high-unmatched-income taxpayers are more likely to be partially compliant. First, consider a taxpayer with only $1,000 in unmatched income. Whether this taxpayer reports this income correctly, reports only $500, or reports none of it, his probability of audit is unlikely to change very much. Even 100% underreporting represents a relatively small absolute amount of misreporting that is perhaps only slightly more likely to be selected for audit relative to a correct report. On the other hand, a taxpayer with $100,000 in unmatched income is engaging in increasingly egregious noncompliance as he goes from $0 to $50,000 to $100,000 of underreporting; this in turn provides the high-unmatched-income taxpayer with a greater opportunity to influence his likelihood of audit and detection by deviating from 100% underreporting. Second, the $1,000 taxpayer faces a relatively small difference in consumption in the audit vs. no-audit states. For instance, if the taxpayer faces a 35% tax rate and underreports 100% of $1,000 instead of 0% of $1,000, he stands to gain only $350 if unaudited and lose only $262.50 if audited (assuming a 75% penalty). On the other hand, if a taxpayer faces a 35% tax rate and underreports 100% of $100,000 instead of 0% of $100,000, he stands to gain $35,000 if unaudited and lose $26,250 if audited (assuming a 75% penalty). Since 100% underreporting represents such a larger-stakes gamble for the latter taxpayer, he will be more willing to sacrifice some profitable underreporting in order to better ensure that the "preferred" no-audit state occurs.

These predictions imply that aggregate measures of unmatched income misreporting are not sufficient to gauge the empirical validity of the deterrence paradigm. In particular, aggregate measures will entangle the

two types of noncompliant taxpayers, those who are predicted to underreport 100% and those who deviate from 100%. Phillips (2011) finds that the majority of taxpayers with little unmatched income do in fact underreport 100% of unmatched income, with many other low-unmatched-income taxpayers actually underreporting in excess of 100%.[10] Taxpayers with larger amounts of unmatched income are instead found to be more likely to be partially compliant, entirely consistent with the expanded deterrence theory.

The previous discussion focused on the prevalence of 100% underreporting and the incentive for deviations below 100%. However, it is also worth discussing the expanded model's predictions regarding 0% underreporting, i.e. the incidence of any noncompliance. Accounting for the high detection rate associated with matched income misreporting explains a significant portion of the low aggregate incidence of observed noncompliance. For instance, Phillips (2011) estimates that 32.5% of the total population underreports Form 1040 Total Income; however, only 41.0% of the population receives unmatched income. Disaggregating the population into those with and without unmatched income, the rates of underreporting are instead 49.6% for taxpayers with unmatched income and only 20.5% for taxpayers without unmatched income.[11] Furthermore, the underreporting incidences are even higher when one limits the sample to taxpayers with strictly nonnegative unmatched income line items (54%) or taxpayers with positive unmatched Schedule C income (61%).

However, classical deterrence theory predicts that all taxpayers (at least those with unmatched income) will be noncompliant.[12] Even the expanded deterrence theory described in this paper makes a similar prediction. As previously discussed, the benefit associated with marginal increases in compliance is proportional to the difference in consumption between the audit and no-audit states. When the taxpayer considers whether to engage in the very first dollar of underreporting, this difference in consumption is $0; therefore, the taxpayer has no incentive for compliance and is predicted to underreport at least some of his income.[13] While the deterrence theory explains the low incidence of underreporting for the majority of taxpayers with only matched income, explaining the less-than-universal underreporting among the remaining taxpayers still appears to require the addition of alternative determinants.

Conclusion

In this paper, I have discussed an extension of the classical deterrence theory of income tax noncompliance that more accurately accounts for the targeted audit and detection regime that taxpayers face. Of primary importance is the fact that taxpayers have the opportunity to exert some control over the terms of the underreporting gamble, with increases in compliance making taxpayers appear less suspicious and thus reducing the likelihood of audit and detection. Therefore, even a taxpayer who is motivated solely by amoral cost-benefit calculations has some incentive to deviate away from full noncompliance.

With this in mind, I contend that referring to the classical model as a paradigm of "deterrence" may be misleading. The noncompliance gamble clearly has a profitable risk-reward profile (so long as the taxpayer has some income that is not third-party-reported), such that the risk of audit, detection, and penalty is insufficient to "deter" the stylized taxpayer from significant amounts of underreporting. However, this prediction of "lack of deterrence" (i.e. a prediction of 100% unmatched income underreporting) is in fact borne out in the taxpayer-level analysis of Phillips (2011), at least among taxpayers with relatively small amounts of unmatched income. Furthermore, the model predicts only partial noncompliance for high-unmatched-income taxpayers, but it is not the case that these taxpayers are "deterred" from greater amounts of underreporting. Instead, the theory should be considered a paradigm of "strategy" as these taxpayers sacrifice some profitable underreporting in order to tilt the odds of the noncompliance gamble in their favor.

Additionally, I would like to point out that the argument made in this paper should not be read as a refutation of the "behavioral" paradigm of noncompliance. A valuable body of research has confirmed the existence of these alternative margins, beyond amoral cost-benefit calculation, that affect noncompliance. In fact, there is no reason that the "deterrence" and "behavioral" paradigms need be mutually exclusive. Instead, the paper should be read as a defense of the classical deterrence theory's ability to characterize the primary motivations that underlie real-world noncompliance. When one more carefully considers the theory's predictions in a realistic environment of targeted audits and endogenous detection, and furthermore relies on taxpayer-level rather than aggregate measures of underreporting, the empirical evidence in support of the theory is quite strong.

From a policy perspective, this finding implies that the lessons and implications of the deterrence paradigm should not be ignored, as one might be inclined to do when "behavioral" alternatives are thought to primarily explain observed behavior. Namely, policies that directly affect the taxpayer's basic cost-benefit analysis can significantly affect gross noncompliance. Among such policy parameters, the average audit and detection rates are clearly important. However, the paper also demonstrates the equal (or perhaps greater) importance of another policy parameter, the degree to which the tax agency effectively targets and detects the most egregious instances of noncompliance. A successful targeting strategy not only increases the recovery of a given amount of noncompliance, but can also generate an endogenous response of increased self-reported compliance.

References

Allingham, Michael and Agnar Sandmo (1972). "Income Tax Evasion: A Theoretical Analysis," *Journal of Public Economics*, 1: 323-338.

Alm, James, Jorge Martinez-Vazquez, and Benno Torgler (2010). "Developing Alternative Frameworks for Explaining Tax Compliance." In *Developing Alternative Frameworks for Explaining Tax Compliance*, ed. James Alm, Jorge Martinez-Vazquez, and Benno Torgler. Florence, Kentucky: Routledge.

Andreoni, James, Brian Erard, and Jonathan Feinstein (1998). "Tax Compliance." *Journal of Economic Literature*, 36(2): 818-860.

Becker, Gary (1968). "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(2): 169-217.

Bloomquist, Kim (2003). "Tax Evasion, Income Inequality, and Opportunity Costs of Compliance." Presented at 96th Annual Conference of the National Tax Association. <http://www.irs.gov/pub/irs-soi/bloomq.pdf>

Kirchler, Erich, Stephan Muehlbacher, Barbara Kastlunger, and Ingrid Wall (2010). "Why Pay Taxes? A Review of Tax Compliance Decisions." In *Developing Alternative Frameworks for Explaining Tax Compliance*, ed. James Alm, Jorge Martinez-Vazquez, and Benno Torgler. Florence, Kentucky: Routledge.

Phillips, Mark (2011). *Essays on Individual Income Tax Compliance*. Ph.D. Dissertation, Department of Economics, University of Chicago.

Slemrod, Joel (2007). "Cheating Ourselves: The Economics of Tax Evasion." *Journal of Economic Perspectives*, 21(1): 25-48.

U.S. Department of the Treasury, Internal Revenue Service (2002). *IRS Data Book: 2001*. <http://www.irs.gov/pub/irs-soi/01databk.pdf>

U.S. Department of the Treasury, Internal Revenue Service (2003). *Report to Congress: IRS Tax Compliance Activities*. <http://www.irs.gov/pub/irs-soi/03congressrpt.pdf>

U.S. Department of the Treasury, Internal Revenue Service (2007). *Tax Year 2001 Federal Tax Gap (Extended Edition)*. <http://www.irs.gov/pub/irs-utl/tax_gap_update_070212.pdf>

Yitzhaki, Shlomo (1974). "A Note on `Income Tax Evasion: A Theoretical Analysis.'" *Journal of Public Economics*, 3(2), 201-202.

Endnotes

[1] This paper summarizes portions of the content of my 2011 Ph.D. dissertation (Phillips (2011)) for the Department of Economics at the University of Chicago. Part of the dissertation was researched and written while interning with the Office of Research at the Internal Revenue Service. The content of the dissertation and paper is the opinion of the writer and does not necessarily represent the position of the Internal Revenue Service.

[2] It is worth noting that Professor Slemrod is not necessarily endorsing but simply stating this prevailing argument.

[3] Allowing for the endogeneity of audit probability with respect to self-reported income (and therefore

indirectly with respect to the amount of underreported income) has long been a staple of theoretical research on noncompliance, and in fact was included in the original Allingham and Sandmo (1972) theoretical treatment. Such endogeneity has also provided the foundation for an extensive game theoretic literature on noncompliance. The game typically consists of taxpayers who have private information regarding their true income and choose how much to report. The tax agency in turn observes taxpayers' reported incomes and chooses whom to audit based on this signal.

[4] This and all future statistics from Phillips (2011) are based on raw data as detected by auditors in the 2001 National Research Program (NRP) study of individual income tax compliance.

[5] Net Misreporting Percentage is defined as "the net amount of income or offset misreported divided by the amount that should have been reported." The IRS (2007) estimates are based on data from the 2001 NRP study, controlling for noncompliance that was undetectable to NRP auditors. Phillips (2011) instead uses raw NRP data.

[6] The importance of third-party-reporting in the compliance decision has been recognized by many other researchers. See for example Andreoni, Erard, and Feinstein (1998), Slemrod (2007), and Alm et al. (2010).

[7] Andreoni, Erard, and Feinstein (1998) states that "taking information reporting into account, taxpayers still appear to be more honest than might be expected..." Slemrod (2007) instead states that "whether the… noncompliance rate of [unmatched] nonfarm sole proprietors is lower than deterrence theory predicts is less clear."

[8] 11.6% represents the probability that a noncompliant Schedule C (nonfarm sole proprietor) taxpayer avoids audit for six consecutive years, at an annual audit rate of 2.0%. Per the 2001 IRS Data Book, 1.6% of Schedule C returns were examined in tax year 2000. Of these examinations, 77.0% were deemed to have additional liability, reflecting the IRS's ability to target its examinations towards those returns most likely to be noncompliant. Using the Phillips (2011) estimate that 61% of taxpayers with positive Schedule C income are estimated to have underreported income, the aggregate annual rate of audit conditional on having underreported is therefore estimated to be (77%) * (1.6%) / (61%) = 2.0%. The cumulative audit rate over six years is then estimated to be 1 - (1 - 2.0%)6 = 11.6%. 75% is the penalty for "civil fraud" whereas the statutory penalty for "substantial understatement" is 20%. The higher 75% penalty rate is used in the current discussion to demonstrate that aggregate probabilities are so low that the deterrence theory (with exogenous audit and detection probability) nonetheless predicts large amounts of noncompliance.

[9] The prediction of 100% underreporting occurs even when allowing for risk-aversion. For instance, under the current parameterization and assuming that the taxpayer exhibits constant relative risk aversion and has the entirety of his income unmatched, 100% underreporting is predicted so long as the coefficient of relative risk aversion is less than 2.45. If the penalty rate is lower than 75% or only a portion of the taxpayer's total income is unmatched, the critical value of risk aversion required for anything less than 100% unmatched income underreporting is even larger. For instance, a 20% penalty rate increases the critical value of risk aversion to 6.68.

[10] Underreporting in excess of 100% means that taxpayers earned a positive amount of unmatched income and instead reported negative earnings. 100% instead implies that taxpayers earned a positive amount of unmatched income and reported $0 in earnings.

[11] A few notes are in order regarding these estimates. First, the analysis in Phillips (2011) categorizes 2001 Form 1040 line items as either "matchable" or "unmatchable." Matchable income is defined as those income line items that IRS (2007) categorizes as "subject to substantial information reporting and withholding" or "subject to substantial information reporting." Unmatchable income corresponds to line items "subject to some information reporting" or "subject to little or no information reporting," with the exception of capital gain distributions which are considered matchable per Bloomquist (2003). Therefore the statistics presented here are based on imprecise measures of actually "matched" and "unmatched" income. For instance, some of the taxpayers with only "matchable" income may have actually possessed some "unmatched" income that simply appears on "matchable" lines. Furthermore, the statistics in Phillips (2003) are based on raw NRP data and reflect only noncompliance that was detectable to NRP auditors. Therefore estimates of the frequency of underreporting experience downward bias relative to the true frequencies since some of these taxpayers may have underreported but were not detected to have done so.

[12] As previously discussed, the net expected return to a marginal dollar of income underreporting is significantly positive. Even if the stylized taxpayer of the model is highly risk averse, the first marginal dollar of underreporting (i.e. going from $0 to $1 of underreporting) leads to essentially no difference in consumption between the audit and no-audit states and the taxpayer should not be deterred from at least some strictly positive underreporting.

[13] If the audit process itself is costly (in ways beyond penalties that are incurred), the taxpayer does experience some difference in utility between the audit and no-audit states, in which case the expanded model would predict that some taxpayers are fully compliant. However, this explanation seems valid only for those taxpayers who have very little unmatched income and therefore don't consider the small expected gains of underreporting to be worth the hassle of a potential examination. This explanation has little intuitive appeal for explaining the non-0% correct reporting incidence of taxpayers with any significant amounts of unmatchable income.

3

Estimating the Tax Gap

Rubin

Erard ◆ Feinstein

Lawrence ◆ Udell ◆ Young

The Practicality of the Top-Down Approach To Estimating the Direct Tax Gap

Marcus Rubin,
Her Majesty's Revenue and Customs, United Kingdom

I. Scope and Key Concepts

Scope of Paper

The paper evaluates the top-down approach to direct tax gaps. It tackles two distinct issues:

- The practicality of estimating entire direct tax gaps top down; and
- The contribution of top-down methods to elements of the direct tax gap.

The focus of the paper is on UK corporate and personal taxes, corporation tax, income tax and capital gains tax, and national insurance contributions. The paper covers relevant experience in other countries, and draws some conclusions that apply beyond the UK.

Key Concepts

The tax gap is the difference between the tax collected and the tax that would be collected if all individuals and companies complied with both the letter and the spirit of the law. The letter of the law is the literal interpretation of tax legislation. The spirit of the law is the intention of Parliament in legislating on tax. In contrast to the letter of the law, which implies a literal legal interpretation of tax liability, the spirit of the law implies a purposive interpretation.

The *top-down approach* is based on a single estimate of the tax base from which a theoretical tax liability is calculated. *Theoretical liability* is the tax that would be collected if all individuals and companies complied with both the letter and the spirit of the law. The tax gap estimated top down is the difference between theoretical liability and total tax collected.

The alternative to the top-down approach is the bottom-up approach. The bottom-up approach builds a total tax gap from estimates of components of the tax gap. The components for the UK are estimated using tax information, random enquiries, risk registers and data matching.

For the direct taxes covered by this paper the tax base is taxable income. Top-down methods, however, do not estimate taxable income. They estimate a total for income that also includes underground production or the shadow economy.

Underground production would be legal, but the income from underground production is not declared to public authorities and, as a result, it is not included in the primary sources used to compile the national accounts.[1]

The shadow economy includes, in addition to underground production, illegal production, such as smuggling.

Under-declared income is the income that should be declared to the tax administration, but is not. It includes both income under-declared on a tax return and income not declared because a tax return that should be submitted is not submitted.[2] In the UK, and other countries where the primary sources used for the income

in the national accounts largely reflect tax administration data, the income from underground production is under-declared income.

II. Requirements, Advantages and Disadvantages of the Top-Down Approach

Introduction

The top-down approach has two basic requirements:

(i) suitable information on the tax base; and

(ii) a calculation of theoretical liability.

This section discusses these basic requirements and the suitability of UK data for top-down direct tax estimates. It also discusses the pros and cons of the top-down approach relative to the alternative bottom-up approach.

Suitable Information

Information on the tax base is only suitable if it is both independent of the tax administration and sufficiently reliable. Information that is not independent of the tax administration, such as national income based on aggregating administration data, does not give the administration new knowledge. A top-down estimate based on aggregated administration data would do more than approximate a bottom-up estimate that the administration could make from the disaggregated data.

The income information in the UK national accounts generally meets the requirement of sufficient reliability. The required reliability does depend on the size of the tax gap and could be an issue where the tax gap is small. The main issue, however, with the UK national accounts is their dependence on HMRC data, a much more fundamental problem than reliability.

The personal, mixed and corporate income statistics in the UK national accounts are largely based on an aggregation of HMRC data.[3] In contrast, the information in the national accounts on consumption, the tax base for indirect taxes, which is derived from consumer surveys, is independent of HMRC and suitable for tax gap estimates and as such is used by HMRC to estimate the VAT gap.

Any estimate of national income that includes the shadow economy, or rather underground production, can be used to calculate theoretical liability and the tax gap. In principle all methods that estimate the shadow economy can be used for tax gap estimates. The issue with methods of estimating the shadow economy is reliability. In addition national income does not provide a complete measure of the direct tax base. Also, macro model methods provide only aggregate information on total income without any indication of how much income is subject to corporate and now much to income tax.

Calculation of Theoretical Liability

Theoretical liability is more difficult to calculate for direct taxes than for indirect taxes. Individual circumstances are crucial for direct tax liability, which depends on various reliefs and allowances. In addition tax rates vary with income. For indirect taxes, in contrast, liability generally depends on the value or volume of sales. The tax liability on sales does not generally depend on the circumstances of the individual businesses making the sales.

The key difference between direct and indirect tax theoretical liability is that direct tax liability requires much more information than indirect tax liability, which can generally be calculated from total sales. A calculation of direct tax liability requires, in addition to total income, sufficient information on individual circumstances to calculate the total allowances and reliefs available and the overall distribution of income.

In addition direct tax theoretical liability depends on the reliefs and allowances that would be claimed without avoidance. The available information on reliefs and allowances, which comes from tax records, is on

claims against declared income and does not indicate the reliefs and allowances that could be claimed without avoidance against total income, including under-declared income.

The difficulty of calculating theoretical liability is particularly acute for corporate tax. Reliefs and allowances are more important than for personal tax. Also, HMRC estimates suggest that avoidance is a larger share of the UK corporate tax gap than it is of the personal tax gap.[4] For these reasons information on total income produced by the top-down approach gives even less indication of theoretical tax liability for corporate tax than for personal tax.

Pros and Cons of the Top-Down and Bottom-Up Approaches

An assessment of the top-down approach must consider the relative merits of the alternative bottom-up approach. The main advantages of the top-down approach are:

- it gives a single estimate;
- it is based on sources independent of the tax administration;
- it produces timely estimates; and
- it requires few resources.

The main disadvantage is the uncertainty of top-down estimates. Uncertainty is also a problem of bottom-up estimates. However, uncertainty reduces the advantage of the single estimate given by the top-down approach. Indeed, uncertainty is such a serious problem that it can render top-down estimates meaningless.

Uncertainty is inherent to the top-down methodology, which estimates a relatively small number by subtracting one large number, the tax actually paid, from another large number, the tax theoretically due. The smaller the tax gap the greater, relative to the size of the tax gap, is the uncertainty inherent to top-down estimates.

For the bottom-up approach uncertainty is less of a problem. The bottom-up approach does not depend on the value of a single estimate. It gives estimates, which are often broken down into considerable detail, of each component of the tax gap. Uncertainty is much less of an issue for some components than for others.

With the exception of most payroll taxes, such as UK national insurance contributions, the tax base for direct taxes is broader than the income estimated by the top-down approach.[5] The tax base includes interest, capital gains and, for corporate tax, significant other non-trading taxable income. The single top-down estimate only covers the tax gap related to employees' earned income and corporate trading income. The top-down approach does not cover the entire direct tax gap. To give an entire tax gap income estimated top down has to be adjusted to taxable income. The adjustment, which requires considerable extra information, is problematic.[6] Indeed, the difficulty of making the adjustment with any degree of reliability may be such that the top-down approach is only capable of giving the part of the direct tax gap that relates to employee compensation and trading profit.

The top-down approach is better suited to providing a broad measure of the main elements of a large tax gap than to estimating an entire small tax gap. A small tax gap requires more precision and, also, the adjustment to taxable income is likely to cover a larger part of the total tax gap.

While the bottom-up approach, like the top-down approach, provides only an uncertain estimate of the total tax gap, it has the important advantage of providing operationally useful information. The detailed tax gap breakdown required for a bottom-up approach enables better prioritisation of compliance resources to maximise tax collection.

The useful information provided by a bottom-up approach does require significant resources, for example, random enquiries. Yet, the limited resources required by the top-down approach do not constitute a real advantage over the bottom-up approach. The useful information provided by the bottom-up approach can justify much of the required resources even without a tax gap estimate.

For HMRC, which is already obtaining the information necessary for a bottom-up estimate of the direct tax gap, the issue is the value that the top-down approach can add to the established bottom-up approach.

Any added value depends on the reliability and comprehensiveness of estimates produced by the top-down approach. The reliability of the top-down approach, particularly macro model methods, is highly questionable.

III. The UK Direct Tax Gap and Under-Declared Income

HMRC Tax Gap Estimates

HMRC estimates a tax gap for different components of the tax gap and adds these together to give a total tax gap for direct taxes, indirect taxes and all taxes (Measuring **Tax Gaps** 2010). The estimated total tax gap in 2008/9 is £42bn, almost 9 percent of the tax due, and the direct tax gap (which includes taxes such as stamp duties not covered in this paper) is £22.5bn. Table 1 below gives the estimates for the direct taxes covered by this paper.

TABLE 1. HMRC Estimates of the Direct Tax Gap, 2008/9

Taxes	Tax Gap	Gap as % of Tax Due
Income tax, National insurance contributions, Capital gains tax	£14.5bn	5.4%
Corporation tax	£6.9bn	13.9%

The bottom-up methodology used for the direct tax gap does not allow more than a rough approximation of under-declared income. Under-declared income cuts across most components of the tax gap, including evasion, the failure, in order fraudulently to pay less tax, to declare income that should be declared or to provide accurately other required information. Fraudulent non-declaration of income, of course, forms part of under-declared income, but fraudulent provision of inaccurate information, for example, to claim greater reliefs or allowances, does not. The tax gap arises from overstated expenses and improper claiming of allowances as well as from understated income.

For income tax, national insurance contributions and capital gains tax (hereafter IT, NIC and CGT) £6.6bn, close to half the total direct tax gap is due to inaccurate income tax self-assessment returns. The next largest component is estimated to be the 'hidden economy', the tax loss due to ghosts, moonlighters and individuals with unearned income who are not in self assessment, where the gap is £3.4bn. Inaccurate returns from employers, who are responsible for collection of personal tax from employees as well as payment of employer national insurance contributions, add a further £3.2bn to the gap. Together, the three components total £13.1bn.[7]

Much of the tax gap for IT, NIC and CGT, which is £14.5bn, is due to under-declared income. For the corporation tax gap, which is £6.9m, under-declared income is much more important for small and medium businesses than for large businesses.[8] For the largest businesses the main tax gap issue is avoidance, for which under-declared income is relatively unimportant. Avoidance accounts for £3.6bn of the estimated £4.3bn corporation tax gap for large businesses.

ONS Adjustments and Under-Declared Income

The Office for National Statistics (ONS) makes adjustments to include under-declared income in the national accounts. These adjustments aim to make estimates of national income, and other estimates, such as data in the sector accounts, comprehensive.

In terms of the components of the tax gap, the ONS adjustments correspond to most evasion, the hidden economy and the net tax gap due to errors and failure to take reasonable care. The adjustments affect each of the three types of factor income in the national accounts, employee compensation, mixed income and operating surplus. Mixed income roughly corresponds to self-employed trading profit and operating surplus to company trading profit.[9]

The adjustments provide only an indication of the under-declared income. The uncertainty around the adjustments is small relative to national income, but it is very large relative to under-declared income. The ONS does not estimate under-declared income as such, but its adjustments provide an implicit estimate.

Table 2 shows the three types of factor income, and total factor income, in the 2009 national accounts, the adjustments and the implicit under-declared income.

TABLE 2. 2009 National Accounts Income Adjustments and Under-Declared Income

Factor Income	Total	Adjustment (%)	Implicit Under-declared Income
Employee compensation	£770bn	0.25 percent	£1.9bn
Mixed income	£84bn	23.3 percent	£19.6bn
Gross operating surplus	£386bn	0.18* percent	£0.7bn
Total income	£1,240bn	1.9 percent	£22.2bn

* The ONS adjustment is to just one component of gross operating surplus, the surplus of private non-financial corporations. This adjustment is 0.3 percent of net trading profit.

The under-declared income implicit in the national accounts at first appears low relative to HMRC's tax gap estimates. The tax due on the £21.5 billion implicit undeclared employee compensation and mixed income is significantly less than the £13.1 billion IT, NIC and CGT tax gap excluding avoidance. However, the tax gap after exclusion of avoidance is not all due to under-declared income.

The value of the implicit under-declared income figures is not the absolute levels, but the relative levels. Under-declared income is a much greater direct tax gap issue for the self-employed than for employees, whose tax is mainly deducted at source, or companies. Also, it is more of an issue for small rather than large businesses.

IV. Top-Down Methods

Introduction

Top-down methods of estimating direct tax gaps fall into three groups, according to their source of information on the tax base. The groups are:

- The national accounts, which give information on the operating surplus of corporations, employee compensation and the 'mixed income' of the self-employed;
- Macro models, which give a single estimate of national income; and
- Micro methods, which use surveys to estimate household income or the labour force.

Table 3 lists the three groups of methods and the eight top-down methods of estimating the direct tax gap, or major elements of the gap.

TABLE 3. Top-Down Methods

A) Methods based on the national accounts
1) Calculation of theoretical liability from income
2) Discrepancy using the income measure of GDP
B) Macro model methods
3) Monetary methods
4) Other single indicator methods
5) Latent variable method
C) Micro methods
6) Discrepancy in labour force measures
7) Direct surveys of households
8) Discrepancy between reported income and income inferred from spending

The rest of this section outlines and assesses the various methods, and also looks at the use of the methods other than by tax administrations. Section V looks at how tax administrations use the methods and Section VI discusses the practicality of using the methods in the UK.

National Accounts Methods

National accounts methods are based on factor incomes, the compensation of employees, mixed income and gross operating surplus, which are the main components of GDP. There are two methods. Method 1 calculates total tax theoretical liability from factor income. Method 2 uses the discrepancy between factor income and some other measure of income.

The value of the methods depends on the comprehensiveness of the national accounts. The information national statistical institutions (NSIs) need to ensure comprehensiveness is similar to that needed by tax administrations to estimate a tax gap. Some NSIs undertake considerable work to estimate under-declared income in sectors where non-declaration of income that should be declared is a particular problem.[10]

In Italy the NSI uses a labour input method, considered here as a micro method, Method 6. Since the national accounts are based on surveys, the distinction between national accounts and micro methods may seem artificial. The difference between the methods is that the national accounts methods use aggregate factor incomes whereas the micro methods use results from surveys, which do not generally contribute to the national accounts.

Conversion of National Accounts Data to Taxable Income

A critical problem for both Methods One and Two is converting national accounts income measures to taxable income. The USA illustrates the difficulty of converting national accounts employee compensation to taxable income.[11] Until recently the Bureau of Economic Affairs (BEA), the US NSI, adjusted its personal income statistics to make them comparable to the IRS's adjusted gross income (AGI), a measure of taxable income, by, for example, excluding non-taxable income and including benefits, capital gains and pensions.

In 2005, the last year for which the adjustment was done, an unexplained gap of $778bn, or 8.9 percent of the adjusted BEA measure remained after "all known and measurable statistical and definitional differences", including a "misreporting adjustment" of $508bn for income not declared to the IRS (under-declared income plus income that did not need to be declared). The unexplained gap is attributed to "known definitional differences that cannot be estimated, statistical discrepancies, data sampling and nonsampling errors, use of different source data for AGI and for personal income, incomplete source data, timing anomalies, and other unknown factors." These reasons, along with "immeasurable sources of tax-exempt income" mean the unexplained gap "is not a proper measure of non-compliance" (Ledbetter, 2007, in a BEA journal).

The "misreporting adjustment" for under-declared income is based on IRS tax gap estimates. The unexplained gap could include some under-declared income missing from IRS estimates. However, the BEA's explanation of the unexplained gap means that national accounts methods are not applicable to income tax in the USA. An attempt to apply Method 1 would simply add the tax theoretically due on the unexplained gap, which is not a measure of non-compliance, to IRS tax gap estimates. Method 2 would include the unexplained gap in the discrepancy used to calculate theoretical liability.

The conversion of employee compensation in the national accounts to income that should be declared to the tax administration may not be as difficult in other countries as in the USA. However, in all countries there must be an issue of whether the "unexplained gap" after all possible adjustments is a measure of non-compliance.

For corporate tax the adjustments required to convert operating surplus in the national accounts to taxable income are more difficult than the adjustments for personal income.[12] Gross operating surplus and taxable income are quite distinct concepts.

For a payroll tax gap, such as the national insurance gap, national accounts methods may be more applicable. The base for payroll taxes is usually closer to the employee compensation in the national accounts than the income

tax base. The OECD in its 2004 Employment Outlook used Method 1 to calculate "shortfalls in receipts" of compulsory social security contributions for 27 member countries, including the UK. The NIC tax gap of 20 percent, or more, calculated for the UK was discounted, because of a "relatively simplistic" calculation that did take not account of the lower contribution rates for "contracted out" employees and some married women.

Employment Outlook did not discuss the dependence of national accounts data on tax administration data.[13] However, the OECD Secretariat has agreed in private correspondence that where, as in the UK, the national accounts are based on tax administration data and include an estimate of under-declared income, the estimate is the best measure that can be derived from the national accounts.

Use of Method 2 by National Statistical Institutions

Method 2 uses the discrepancy between income in the national accounts and the income given by some other source. In the UK the ONS did, over thirty years ago, use the initial residual discrepancy (IRD) between the initial expenditure and income measures of GDP as a measure of the shadow economy. MacAfee (1980) found a "glimpse" of the UK shadow economy in the IRD, which was 3.3 cent of GDP in 1978.

The idea here is that the surveys of consumer expenditure used for the expenditure measure include under-declared income when it is spent. The initial income measure does not include under-declared income.

Unfortunately, the IRD only permits a glimpse of the shadow economy through a very dark glass. The UK IRD fell sharply from 1978, becoming negative in the mid-1980s. A study by the Rockwool Foundation of Denmark concluded, "IRD can probably only be interpreted as a reflection of random fluctuations in the underlying statistics used to estimate GDP." While subsequent revisions to the national statistics mean that the UK IRD was not actually negative, the apparent absence since the 1980s of any use of the IRD as a measure of the UK shadow economy indicates widespread acceptance of this conclusion.

The IRD does not provide a reliable measure of the UK shadow economy. It is of no value for work on the tax gap.

Where the national accounts are independent of the tax administration, Method 2 can use the discrepancy between the final measure of income in the national accounts and the income declared to the tax administration. Statistics Netherlands has used this version of Method 2 to estimate under-declared income.[14] The NSI's work highlights the limitations of the Method.

The net adjustments to convert operating surplus in the national accounts of the Netherlands to taxable income were three times as large as the estimate of under-declared income. Presumably, there is a balance of relatively large positive and negative adjustments. The uncertainty in adjustments will result in much greater uncertainty in an estimate of under-declared income that is a small percentage of total income.

The NSI's analysis at sector level showed a further limitation of Method 2. In some sectors the taxable income declared to the tax administration was more than the theoretical taxable income estimated from the operating surplus. Sectoral analysis of Method 2 results, particularly for sectors with little or no under-declared income, can give an indication of bias and uncertainty in the results. Information on sectors with high under-declaration could be useful in prioritising resources and as a check on bottom-up estimates. However, the value of an estimate of an entire tax gap that includes negative values for some sectors can be questioned.

Macro Model Methods

Macro models, which are used to estimate the shadow economy, provide a potential means of calculating tax gaps. These models use a macroeconomic model to estimate national income independently of the national accounts. The difference between the independent estimates of national income and the GDP in the national accounts provides a measure of the shadow economy.[15]

The macro models estimate the shadow economy, which combines under-declared income from underground production, which is a cause of the tax gap, with income from crime, which is not. Where income in the national accounts includes income from crime, the income from crime is separately estimated and can

be easily excluded from a tax gap calculation. The macro models, unlike the national accounts, provide just a single estimate of national income from which income from crime cannot readily be excluded.

In general macro models are used simply to estimate the shadow economy. However, estimates of the shadow economy are sometimes also interpreted as a measure of the tax gap. Such an interpretation is simplistic. Rather than assuming that the shadow economy and the tax gap are the same, a proper tax gap estimate would combine macro models with both a conversion of national accounts information to taxable income and a theoretical liability calculation. Yet, the research for this paper has not found any example of such a proper estimate. The apparent absence of any proper estimate of the tax gap from the shadow economy may reflect a widespread view of the unreliability of macro model estimates of the shadow economy. The difficulty of excluding crime may also be a factor.

Indicator Methods

The three types of macro model, monetary methods, other indicator methods and the latent variable method, are all indicator methods. They rely on an indicator or indicators of the shadow economy. The basic idea behind indicator methods is that the shadow economy leaves some trace in particular statistics. For example, the monetary methods assume that monetary statistics provide an indicator of the shadow economy. The change in a monetary statistic that cannot be explained by factors unrelated to the shadow economy indicates the change in the shadow economy.

Various indicators have been used. The indicators used for Method 3, monetary methods, include the share of large denomination notes in cash in circulation, the cash-deposit ratio and the share of cash in a broad measure of money that includes savings deposits as well as cash deposits. Other single indicator methods, Method 4, have used various measures of electricity consumption, labour force measures, such as multiple job holding and the number of self-employed, and the number of very small enterprises. The latent variable method, Method 5, uses a number of indicators.

Indicator methods have two fundamental failings. The first failing is that the methods only estimate change in the shadow economy. They cannot estimate the size of the shadow economy. In general the methods assume the size of the shadow economy in one year, or in several years, to provide a starting point for the changes that they estimate and, so, give new estimates of the size of the shadow economy.[16] Unfortunately, the size of the shadow economy is not known in any year. If it were, the method used to establish the share in that year would be used to establish the share in other years and indicator methods would be largely redundant.

The second fundamental failing of indicator methods is ignorance of the relationship between the change in the indicator and the change in the shadow economy. Again, an assumption is needed. It is generally assumed that the indicator shows the same effect from a given change whether the change is in the shadow economy or in the economy observed in the national accounts. For example, the electricity consumption method assumes that a change in the total economy indicated by a variation in electricity consumption is the same whether the change is in the shadow economy or the observed economy. This assumption, however, is unjustified and seems to be adopted simply as a matter of convenience.

The monetary methods, and also the latent variable method, assume the same increase in the demand for cash when the shadow economy grows as when the observed economy grows. This assumption, in more technical terms that the velocity of circulation of cash is the same in the shadow economy as in the observed economy, is questionable.

There are good reasons to expect a lower velocity of circulation of cash in the shadow economy. Cash from the shadow economy is less likely to be placed in a savings account because of the risk of questions about its source. Controls to prevent money laundering must contribute to the hoarding of cash by criminals, which can be extensive. Cash hoarding by tax evaders is also common. If shadow economy cash is hoarded, its velocity of circulation will be lower.

A significantly lower velocity of circulation in the shadow economy would mean that the assumption of equal velocity causes serious overestimation of changes in the shadow economy. For example, if an extra

pound of income from the shadow economy increases the demand for cash by half as much as an extra pound of income in the observed economy, the assumption of equal velocity means that the estimated change in the shadow economy is twice the actual change (assuming no other reason for the estimate to differ from the actual change).

Method 5, Latent Variable Method

Recent estimates of the shadow economy tend to use Method 6, the latent variable method, sometimes in combination with a cash demand equation, Method 4. The latent variable method uses a 'multiple indicator multiple causes' or MIMIC model or a dynamic variation, a DYMIMIC model. The MIMIC model was developed for factor analysis in psychometrics to estimate intelligence, which is an unobservable latent variable.

The shadow economy is not a latent or hypothetical quantity like intelligence. It exists, and has the same units of measurement as the observed economy. A fundamental criticism of the latent variable method is that, as the shadow economy is not a latent variable, it cannot be estimated by a MIMIC model.[17]

An obvious advantage of Method 6 over the single indicator methods is the use of more than one indicator. However, this advantage is very much reduced by controversy over whether the indicator variables used are valid indicators.

The main issues with Method 6 are the fundamental failings identified for all indicator methods, the need to assume an anchor for estimated changes in the shadow economy and ignorance of the relationship between the change in the indicator and the change in the shadow economy. These failings are acknowledged in a recent latent variable study, Schneider, Buehn and Montenegro (2010).

"This [MIMIC] analysis provides only relative estimates, not absolute, of the size of the shadow economy. Therefore an additional procedure, benchmarking or calibration procedure, is required in order to calculate absolute values of the size of the shadow economy" (ibid, page 13). "The base values necessary for … the calibration procedure are from the year 2000 and taken from Schneider (2007), who presents estimates of the shadow economies in 145 countries around the world using the MIMIC and the currency demand approach" (ibid, page 19).

The estimates of the base values, or anchors, rely on "the assumption of no shadow economy in the base year."[18] The assumption is justified as "Relaxing this assumption would … imply an upward adjustment of the shadow economy" (ibid, page 39).

The justification of the assumption implies that the shadow economy estimates after an upward adjustment would be implausible. The real issue may be whether macro model estimates are implausibly high even before any upward adjustment due to relaxing the assumption of no shadow economy in the base year.

It is, of course, true that the shadow economy cannot be negative. It does not, however, follow that an anchor based on no shadow economy in a particular year produces under-estimates. If the uncertainty in the estimates of change is larger than the shadow economy, there can be negative estimates in some years. The assumption, which is actually that the shadow economy is never estimated as negative, rather than that the actual shadow economy is never negative, can cause an upward bias.

The second fundamental failing, ignorance of the relationship between the change in the indicator and the change in the shadow economy, is also acknowledged in the recent study. "Without knowledge about the velocity of circulation in the shadow economy, one has to accept the assumption of an 'equal' money velocity" (ibid, page 38).

Ignorance does not justify a single assumption. Where results of a study depend on assumptions, the study normally examines the sensitivity of results to alternative assumptions. Sensitivity analysis, however, is not necessary to indicate the wide uncertainty around latent variable estimates. The authors of latent variable studies generally acknowledge their limitations.[19] Latent variable, and other macro model estimates, of the shadow economy are unsuitable for tax gap estimates.

Micro Methods

Method 6, Discrepancies in Labour Force Measures

Discrepancies in labour force measures are widely used by NSIs to estimate the number of undeclared workers and to ensure that the national accounts are comprehensive. The idea is that household surveys give all workers while business surveys only give declared workers. The difference in survey results represents the number of undeclared workers who do not pay tax.

In Italy the NSI uses discrepancies in labour force measures to calculate the labour input of undeclared workers in each sector and build up a detailed picture of income concealed by employers through not declaring workers. Undeclared workers are more of an issue for the tax gap in Italy, where the NSI has a particular need to estimate the shadow economy, than in Denmark, the UK and probably most other OECD countries.[20]

Method 7, Direct Surveys

Surveys of households and individuals are an obvious way to obtain information on the shadow economy. The problem, however, is the willingness of tax evaders to take part and to reveal earnings deliberately concealed from the tax administration. The design of survey questionnaires, and of any advance letters used, is crucial in addressing this problem.

In 2006 HMRC commissioned a feasibility study from Bristol University. After a pilot survey, HMRC decided not to go ahead with a full survey. The reason was doubt over the ability of the survey to obtain useful information from ghosts and moonlighters with a high level of under-declared income.

Unfortunately, the study did not give detailed consideration to the design of the questionnaire in the light of earlier work. The Netherlands national statistical institution, the European Commission and a Danish research institute, the Rockwool Foundation, have done considerable work on how best to design and use surveys to obtain information on the shadow economy and the tax gap.

Research on undeclared work for the European Commission, Directorate General for Employment and Social Affairs, has included surveys. In 2004 an expert seminar sponsored by the Commission agreed that direct methods (observation, interviews and surveys on the supply side) provided the best means of studying undeclared work. The method of questioning is important for results. The Rockwool Foundation of Denmark had achieved very promising recent results in internationally comparable direct surveys (Rooney et al, 2004).

The Rockwool surveys approach the subject of earnings from undeclared work gradually, first asking about aspects related to the subject, but not particularly sensitive, such as general opportunities to earn extra money in the interviewee's line of work. Such 'warming up' questions help establish interest in the topic and create confidence between interviewer and interviewee. More sensitive 'core' questions about the interviewee's undeclared work are then supposed to be answered more frankly. Indeed, the Rockwool Foundation has shown that the gradual approach finds a much higher level of undeclared work than the direct approach used in the Bristol University pilot

The Netherlands NSI has also adopted a gradual approach. The NSI carried out surveys in 1983 to compare the results of different research designs, as the Rockwool Foundation did later. With the gradual approach 12 percent of interviewees admitted undeclared work, twice the proportion found by the direct approach (OECD, 2002).

Method 8, Discrepancy Between Inferred and Reported Income

Pissarides and Weber (1989) pioneered a method of estimating income concealed by the self-employed through not declaring some of their earnings.

The P-W methodology first estimates the marginal propensity to consume food using cross-section data for a sample of the employed and self-employed. It then estimates under-declared income by comparing the marginal propensities.

The original P-W study in 1989 estimated that under-declared income was 55 percent of the income declared by the self-employed in the UK, but the proportion was lower in later studies. The methodology makes a number of assumptions:

- All respondents report food expenditure accurately.
- Employed respondents report income accurately.
- Employment status does not affect the marginal propensity to consume food.
- Households with tax evaders are not significantly under-represented in a survey that primarily covers spending.

HMRC has shown considerable interest in the P-W methodology, sponsoring work in 1993 and carrying out its own analyses in 2001 and 2005. The 2005 analysis was a useful cross check on the results of the random enquiry programme. It also provided potentially useful information on employment with the greatest risks of under-declaration.

More recently, the PW methodology has been used to estimate income under-declared by the self-employed in Canada, Finland and Sweden. It has also been used to estimate total under-declared income in Turkey, relying on an assumption that the income of employees of large companies is reported accurately.

V. Tax Administration Experience

Introduction

A partial survey of the use of top-down methods by tax administrations to estimate under-declared income or an entire direct tax gap has been conducted.[21] The survey concentrates on developed countries, few of which have used top-down methods.[22] Table 4 summarises the findings of the survey, which includes how the results are or were used.

TABLE 4. Use of Top-Down Methods by Tax Administrations

Country/ Region	Method Used and What is or Was Estimated	Application of Results
Denmark	Method 2 Under-declared personal income	Used to calculate a performance objective, but administration is seeking to measure objective bottom up
Latin America	Method 1 Corporate tax gap Method 7 Personal tax gap	Broad assessment of relative levels of tax gaps for different taxes
New Zealand	Method 5 Total tax gap	None - the administration has considerable reservations about the methodology
Sweden	Method 2 Personal tax gap	Reconciliation with bottom up under-declared income estimates No plan for further top-down work

Denmark

The Danish tax administration has, since 1947, calculated a personal tax gap using the difference between personal income in the national accounts and declared income. The tax gap is "before tax", that is under-declared income, rather than the tax due on under-declared income.

The method used in Denmark avoids the difficulty, and uncertainty, of calculating theoretical liability. Yet, uncertainty is still a serious problem. This problem is reduced by publishing the tax gap as a five year moving average rather than as a potentially misleading annual figure. While the moving average does not reduce systematic errors, it does lessen the fluctuations due to random variation in the data and so provides a better indication of trends.

The Danish "before tax" personal tax gap, measured as a share of GDP, varied between 15 and 20 percent from 1947 to 1955, and then declined, falling below 5 percent in the early 1970s. The latest figure, which is for

2005 to 2009, is 2.3 percent. The decline since 1955 is attributed to structural changes in the economy reducing the share of hard to tax employment, such as self-employment, day labourers paid in cash and domestic service (OECD, 2004).

In Denmark the "before tax" tax gap is a performance objective for the administration. However, the administration is seeking to measure its objective bottom up using random audits. Even with a long history of top-down measurement and without a theoretical liability calculation, the administration views a bottom-up method as better fitted for a performance measure.

While there have no doubt been some methodological improvements since the personal tax gap was first measured in Denmark, the estimated tax gap is now less than one fifth of its 1947 level. The uncertainty in the top-down estimate may mean that a method valuable in 1947 for a relatively large tax gap has ceased to be useful. In contrast the value of bottom-up methods is much greater than in 1947 because of tax administration developments, such as random audits.

Latin America

In Latin America work on the tax gap has tended to focus on VAT rather than direct taxes. Still, most countries have estimated direct tax gaps. The Colombian tax administration, whose first estimate is for 1987 (Shome, 1995), is perhaps the second after Denmark to estimate top-down direct tax gaps.

Columbia has used Method 1, adjusting the operating surplus from the national accounts to give an estimate of taxable income and then to split the taxable income according to whether it was corporate income, subject to corporation tax, or self-employed income, subject to income tax. Seven other Latin American countries covered in a recent study, including Chile and Mexico, have also made Method 1 estimates of the corporation tax gap (Jiminez et al, 2010). The seven countries have also estimated personal income tax gaps through Method 7 surveys.

In the seven countries corporate tax gaps ranged from around 50 to as high as 65 percent of tax liability while personal tax gaps ranged from 30 to 50 percent. The study, in commenting on why the personal tax gap is lower than the corporate tax gap refers to possible over-estimation of the corporate tax gap because of issues with the national accounts data and accounting for revenue loss through deductions. The possible overestimation of the corporate tax gap indicates the major difficulty in estimating corporate tax top down.

The relatively high tax gap levels in Latin America mean that the uncertainty of top-down estimates is a less serious problem than in the UK. Estimates used to assess relative levels of gaps for different taxes require less precision than performance measures. Accordingly, national accounts methods are of more value for Latin American countries than for the UK and Denmark, where tax gaps are smaller. Also, a top-down estimate is much more useful if information required for bottom-up measures is lacking.

New Zealand

In 1999 the New Zealand Inland Revenue Department commissioned a Method 5 latent variable macro study to estimate the tax gap, but shortly later decided not to estimate a tax gap at all. The reason for not using top-down methods is that they lack reliability and do not indicate where the tax gap exists. A method that fails to indicate where the tax gap exists is of no use for decisions on targeting compliance activity.

The study for the Department estimated the total tax gap. It assumes that the discrepancy represents income from undeclared work, under-declared income, and that the income evades indirect taxes when spent as well as evading direct taxes when earned.

Sweden

In 2006 the Swedish tax administration made experimental use of Method 2, estimating under-declared personal income from the national accounts discrepancy between expenditure and income measures.[23] The discrepancy was estimated as about 5 percent of GDP.

As with the New Zealand study, the Swedish administration estimated a total tax gap and assumed that where income from an activity is under-declared, the sales of the goods of services produced by that activity are also under-declared so that both direct and indirect tax are not paid.

The tax gap estimate has a wide range, plus or minus 10 percent, reflecting the uncertainty introduced by the calculation of theoretical liability. This range is in addition to the uncertainty over the level of under-declared income used to calculate the tax gap. For reconciliation with bottom-up estimates the administration used under-declared income, which does not suffer the additional uncertainty caused by the calculation of theoretical liability. The reconciliation used top-down information from direct surveys of households, Method 7, and estimates based on income inferred from food consumption, Method 8.

Other Countries

Comprehensive information on countries that do not appear in Table 2 is lacking. Yet, tax administrations more closely comparable to HMRC, such as the IRS in the USA, have never used top-down methods to estimate entire direct tax gaps. Perhaps, the main reason top-down methods are not used is a widespread view that under-declared income cannot be reliably estimated.[24] Most administrations focus on identifying and assessing risk factors and prioritising compliance resources to areas of highest risk. Top-down estimates are of no value for decisions on compliance priorities.

A number of tax administrations have decided against the top-down approach to the direct tax gap, as has HMRC. Indeed, the tax administrations in Australia, Canada and New Zealand have decided against tax gap measures altogether. The Australian Tax Office "has concluded that accurate and defensible measures of the absolute size of the tax gap are impossible to achieve in a practical sense" (OECD Forum on Tax Administration, 2008). In the United States the IRS has concentrated on bottom-up measures and does not see that the top-down approach would add anything of value.

Conclusions on Tax Administration Experience

Experience in Denmark and Sweden shows how serious the problem of the uncertainty of theoretical liability is for direct tax gaps. Denmark avoids the uncertainty by estimating only under-declared income. When Sweden calculated theoretical liability, the margin of error, 10 percent, was so large as to cast doubt on the value of the calculation.

Tax administrations similar to HMRC do not generally estimate entire tax gaps by top-down methods. Even in Denmark, which is an exception, the administration would prefer a performance measure calculated bottom up.

NSIs and tax administrations have a common interest in under-declared income. Estimates of under-declared income, which may be required to ensure the comprehensiveness of the national accounts, are valuable for tax gap purposes. This creates the possibility of joint work. In the UK, where the ONS has not reviewed the methodology of their under-declared income estimates for some considerable time, HMRC and the ONS are exploring the possibility of working together.

VI. Evaluation of the Methods and Conclusions

This Chapter evaluates the eight methods, including their applicability to the UK, before giving overall conclusions.

National Accounts Methods, Methods 1 and 2

The dependence of the UK national accounts on HMRC means that Method 1, Calculation of theoretical liability from national accounts income, is not applicable in the UK. Method 1 can only be used if the income information in the national accounts is independent of the tax administration.

Method 2, which is based on a discrepancy using the income measure of GDP, can be applied where Method 1 is inapplicable. Indeed, the version of Method 2 that uses the Initial Residual Difference between

income and expenditure measures of GDP exploits the dependence of the income measure on the tax administration. However, for the UK the IRD reflects fluctuations in the underlying statistics used to estimate GDP rather than the shadow economy.

Method 2 has been applied in Denmark and, experimentally, in Sweden to calculate under-declared income or a personal tax gap. The discrepancy used in Denmark, the difference between income declared to the tax administration and income in the national accounts, is not available in the UK. Method 2 in Sweden used the IRD where, in contrast to the UK, the IRD has been remarkably stable over 40 years.

The personal tax gap in Sweden, according to a Method 7 survey, is twice that in the UK.[25] Sweden, unlike the UK, does not limit tax liability to activities with a business purpose. The IRD in Sweden in 2006 was 5 percent, which is considerably more than the last available figures for the UK. Given the relatively low level of the personal tax gap and the large fluctuations in IRD in the UK, the experimental use of IRD in Sweden provides no reason to revise the conclusion based on UK experience that the IRD cannot provide useful tax gap information for the UK.

Macro Model Methods, Methods 3 to 5

Unlike the national accounts methods, macro models are readily applicable to the UK. The basic requirement for a macro model is simply an equation relating national income to an economic variables or variables. Studies can cover as many as 162 countries and their scope is only limited by data availability, which is not an issue for the UK.

The main issue with the macro model methods is their reliability. The methods require information that does not exist in the UK or in other countries and have to make assumptions.[26] For example, monetary methods require information on the velocity of circulation in the shadow economy. The general assumption of equal velocity in the shadow and observed economies is unjustified and probably results in exaggerated estimates of change in the shadow economy.

The macro model methods only identify changes in indicators, such as cash demand, unexplained by factors not related to the shadow economy. They rely on assumptions to estimate the shadow economy. First, there is an assumption, such as on the velocity of circulation, to estimate the change in the shadow economy from changes in the indicator or indicators not explained by other factors in the estimating model. Then, there is a further assumption to estimate the size of the shadow economy from the change in the shadow economy.[27]

Macro models sometimes estimate the size of the shadow economy from the change in the shadow economy from a year in which there is assumed to be no shadow economy. The basis of this assumption is that the shadow economy cannot be negative. However, macro models will produce negative estimates of the shadow economy if the uncertainty in the estimates of change is sufficiently large relative to the actual shadow economy. As a result there may be a strong upward bias in estimates of the shadow economy, particularly in countries like the UK with relatively small shadow economies.

Macro model methods have not produced reliable estimates of changes in the shadow economy or of the size of the shadow economy. They are of no value for work on the tax gap. This conclusion on macro model methods is very much in line with a declaration by the world's statistical institutions.[28]

Declaration by the world's statistical institutions

"Unofficial estimates [of the shadow economy] are often based on macro economic models. … The OECD-ILO-IMF-CIS manual on measuring the non-observed economy rejects such 'macro-model' methods because these methods suffer from serious problems that cast doubt on their utility for any purpose in which accuracy is important. In particular, they are completely unsuitable for use in compiling the national accounts."

Micro Methods, Methods 6 to 8

Method 6 uses discrepancies in employment statistics to estimate undeclared work just as Method 2 uses discrepancies in income statistics to estimate the shadow economy. Method 6 also requires information on average income from undeclared work. It is not particularly suitable for UK tax gap estimates. Illegal workers, such as non-EU nationals working in the UK without work permits, are part of the discrepancy in employment statistics, but may be paying taxes and so making no contribution to the tax gap.

Employment discrepancies identify ghosts, whose work is entirely underground production, but not moonlighters, only part of whose work is underground production. Their value for tax gap estimates depends on the relative importance of ghosts and moonlighters. In the UK the limited available evidence suggests that ghosts are not more important than moonlighters.[29] As in Denmark, Method 6 might not even find employment discrepancies in the UK. The Method appears to have no value for the UK.

Method 7, the direct survey method, has considerable potential for use in UK tax gap estimates, provided that a gradual approach is used. Indeed, Method 7 already contributes to HMRC's tax gap estimates for moonlighters through the use of the results of a 2003 study by the Rockwool Foundation of Denmark.

Method 8, reported and inferred income, like Method 7, has been used by HMRC for work on the tax gap. Methods Seven and Eight share the difficulty of all top-down methods in calculating a tax gap from under-declared income. Data matching with HMRC records, which is now open to external researchers through the HMRC Data Lab, subject to safeguards on taxpayer confidentiality, overcomes this difficulty. National insurance numbers of individuals of interest for under-declared income could potentially be used in a booster to the ONS Living Costs and Food survey with the main sample a control group.

Overall Conclusions

Top-down methods do not provide a practical method of estimating entire UK direct tax gaps. Yet, direct surveys (Method 7) and inferred and reported income (Method 8) can potentially contribute to elements of the direct tax gap by providing estimates of under-declared personal income, particularly for the self-employed. Indeed, elements of these methods are already used in the UK for direct tax gaps. Internationally, top-down methods give more meaningful information for developing countries than for developed countries, where tax gaps are generally smaller.

Appendix

GLOSSARY

Bottom-up approach	The bottom-up approach builds a total tax gap from estimates of components of the tax gap, generally based on tax information.
Evasion	Evasion is the failure, in order fraudulently to pay less tax, to declare income that should be declared or to provide accurately other information required by a tax administration.
Ghosts	Ghosts are individuals with taxable income (from employment or self-employment) but who fail to make any tax return.
Hidden economy	The hidden economy consists of productive activities on which tax is due, but which are hidden from the tax administration by not making any tax return (as opposed to concealment through non- declaration when a tax return is made)
Illegal production	Illegal production consists of productive activities, including the distribution and sale of goods and services, forbidden by law.
Letter of the law	The letter of the law is the literal interpretation of tax legislation.
Money laundering	Money laundering is the processing of income from illegal or underground production to disguise the origin of the income.
Moonlighters	Moonlighters are individuals who pay tax on income from one employment (generally through employer deductions) but not on income from other employment or self-employment.
Non-observed economy	The non-observed economy consists of the shadow economy plus informal and other activities, such as household production for own use.
Shadow economy	The shadow economy consists of both underground and illegal production.
Spirit of the law	The spirit of the law is the tax administration's interpretation of the legislature's intention in legislating on tax.
Tax gap	The tax gap is the difference between tax collected and the tax that should be collected (the theoretical liability). The tax gap does not include theoretical liability due to illegal production.
Theoretical liability	Theoretical liability is the tax that would be collected if all individuals and companies complied with both the letter and the spirit of the law.
Top-down approach	The top-down approach estimates the tax gap as the difference between theoretical liability and total tax collected.
Under-declared income	Under-declared income is the income that should be declared to the tax administration, but is not. In the UK and other countries where the primary sources used to compile the national accounts are not independent of the tax administration, it is also the income from underground production.
Underground production	Underground production consists of productive activities that are in themselves permitted by law but are illegal because income is not declared to public authorities and, as a result, not included in the primary sources used to compile the national accounts

References

Breusch, Trevor (2005), "Estimating the Underground Economy using MIMIC Models."

The Danish Tax and Customs Administration (2009), "Compliance with the tax rules by businesses in Denmark 2006".

The Danish Tax and Customs Administration (2009), "Compliance with the tax rules by private individuals in Denmark 2006".

Dell'anno, Roberto, and Schneider, Friedrich, (2006), Estimating the underground economy by using MIMIC models: A response to T. Breusch´s critique.

Declaration of the ISWGNA (2006), Estimates of the unrecorded economy and national accounts.

European Commission (2006), "Feasibility Study of Direct Survey on Undeclared Work".

European Commission (2007), Special Eurobarometer 284 "Undeclared Work in The European Union".

Franklin, Jon (2010), "Can Money Demand be used to Measure the Informal Economy in the UK?", MSc Thesis, Birkbeck College, London.

HMRC (2010), Measuring Tax Gaps.

Jiménez, Juan Pablo, Sabaini, Juan Pablo Gómez and Podestá, Andrea (2010), "Tax gap and equity in Latin America and the Caribbean", GTZ and United Nations Economic Commission for Latin America.

Ledbetter, Mark A. (2007), "Comparison of BEA Estimates of Personal Income and IRS estimates of Adjusted Gross Income," Survey Of Current Business, 87 (November): pp. 35–41.

OECD Centre for Tax Policy and Administration (2008), "Monitoring Taxpayers' Compliance: A Practical Guide Based on Revenue Body Experience".

OECD Centre for Tax Policy and Administration (2009), "Tax Administration in OECD and Selected Non-OECD Countries: Comparative Information Series (2008)".

OECD (2002), "Measuring the Non-Observed Economy: A Handbook".

OECD (2004). Employment Outlook, Chapter 5, Informal Employment and Promoting the Transition to a Salaried Economy

Pissarides, Christopher and Weber, Guglielmo (1989), "An Expenditure-based Estimate of Britain's Black Economy," Journal of Public Economics, 39: pp. 17–32.

Pedersen, Søren (1998), "The Shadow Economy Results", The Rockwool Foundation, Copenhagen.

Pedersen, Søren (2003), "The Shadow Economy in Germany, Great Britain and Scandinavia", The Rockwool Foundation, Copenhagen.

Rand Europe & National Audit Office (2008), "Comparing How Certain Tax Authorities Tackle the Hidden Economy."

Schneider, Friedrich, Buehn, Andreas and Montenegro, Claudio E. (2010), "Shadow Economies All over the World New Estimates for 162 Countries from 1999 to 2007", The World Bank Development Research Group Poverty and Inequality Team & Europe and Central Asia Region Human Development Economics Unit.

Schneider, Friedrich and Dominik Enste (2000), "Shadow Economies: Size, Causes, and Consequences," The Journal of Economic Literature, 38/1, pp. 77-114

Shome, Parthasrathi (1995), "Comprehensive Tax Reform: The Colombian Experience," IMF Occasional Paper No. 123.

Swedish Tax Agency (2008), "Tax Gap Map for Sweden".

Endnotes

[1] The informal sector, for example individuals below the tax threshold, is legal but not included in the primary sources and is separately estimated in the national accounts.

[2] HMRC does not estimate under-declared income.

[3] The text is a simplification for corporate income. The UK national accounts use company accounts for the surplus of large companies. However, companies use the same information for their accounts and tax returns. Accordingly, the accounts of large companies are not independent of tax returns.

[4] For personal tax the 2008/09 avoidance estimate is £1.4m, which is 9 percent of the total personal tax gap estimate of £14.5bn. While an avoidance estimate is not published for all corporate tax, the estimate for avoidance by large businesses is £3.6bn, over half the total corporate tax gap of £6.9bn.

[5] Even for national insurance contributions the top-down approach does not cover the entire tax base. National insurance offshore avoidance, for example, the supply of staff as part of a composite service by an overseas company, does not affect income estimated top down.

[6] Details of the adjustments for corporate income in the UK are given in endnote 12. The value of the personal income gap remaining after adjustments as a measure of the tax gap is discussed for the USA and the Netherlands.

[7] The figure of £13.1bn is illustrative rather than definitive because it includes the hidden economy for which the estimates are illustrative.

[8] A large business has either 250 or more employees or both a turnover over €50 million and a balance sheet total over €43 million.

[9] For details on mixed income see paragraphs 5.29 to 5.31 of the ONS publication "National accounts concepts sources and methods."

[10] UNECE (2008), Non-observed economy in the national accounts, provides a survey of NSI practice.

[11] The difficulty is even more acute for 'mixed income', the surplus of unincorporated enterprises owned by households before interest or rent. The income is mixed because, in addition to the operating surplus of unincorporated enterprises, it includes the employee compensation of household members from an unincorporated enterprise owned by a member of the same household. A calculation of income tax and national insurance liability requires information on the distribution of income between household members. To calculate theoretical liability from under-declared 'mixed income' would require a series of heroic guesses that cannot produce reliable estimates.

[12] In the UK the main adjustments concern interest and finance lease rentals, which are not part of taxable income, and financial charges such as commissions, stamp duties and other expenses connected with mortgages, capital issues and transfers of financial assets, which are not deducted in calculating taxable income. Chapter 4, The Income Approach, in the UK Gross National Income (ESA95) Inventory gives details of the adjustments, which also concern expenses associated with take-over activity, launch aid subsidies, entertainment expenses, royalties on patents and profits earned abroad.

[13] A Vienna Institute for Economic Studies working paper, which calculates both income tax and NIC tax gaps, similarly does not mention the dependence (Christie and Holzner, 2006). The working paper finds a UK NIC gap of 35 percent and a UK income tax gap of 22 percent in 2002/3. On the basis of a reduction in the UK income tax gap from 35 percent in 1995-96, the paper concludes, "It seems clear that the tax collection process itself has been improved over the period." The results rely on "simplified calculations of the liabilities based on the most important elements, namely employment income, pensions and average separate tax rates for property income where applicable, though we also took account of the most important deductions, personal allowances and tax credits." The OECD's finding of a much lower NIC gap suggests that results are very sensitive to how the tax system is simplified. The trend in the personal income tax gap may also be sensitive to simplification.

[14] OECD (2002), "Measurement of the Non-Observed Economy: A Handbook," page 52.

[15] The independent estimate of the shadow economy includes, as well as the difference between independent and official GDP, the estimate of the shadow economy in official GDP.

[16] The reliance on assumptions is sometimes disguised by reference to estimates from other studies. However, these studies are generally also indicator studies and can only give the size of the shadow economy by assumption.

[17] Breusch (2005) also makes several more technical objections to the use of the latent variable method to estimate the shadow economy. The main objections as summarised by Schneider et al (2010, page 19, footnote 23) are "(i) instability in the estimated coefficients with respect to sample size changes, (ii) instability in the estimated coefficients with respect to alternative specifications, (iii) difficulty of obtaining reliable data on cause variables other than tax variables, and (iv) the reliability of the variables grouping into "causes" and "indicators" in explaining the variability of the shadow economy". These technical objections appear to be largely accepted. However, this paper does not discuss the technical objections further. The conclusion that the latent variable method has no value for tax gap estimates does not require technical arguments.

[18] For the currency demand studies the reliance is direct and for the MIMIC studies indirect through their use of other studies that rely on the assumption.

[19] For an acknowledgement of the limitations see Del'Anno and Schneider (2006). "Estimation is particularly challenging, researchers are forced to use some kind of 'imagination' because existing estimation procedures are not convincing and complications are still numerous and available" (page 2). "Shadow economy estimates are never very stable and absolute, and there is always space for questions, discussion and critique" (page 16).

[20] In Denmark an attempt to apply the Italian labour input method showed no discrepancies (Rockwool, 1998: page 72). In a 2008 UNECE survey Italy had the largest non-observed economy (NOE), 15 percent of the total economy, of any OECD country. Excluding Italy and Mexico, the seven OECD countries in the survey with NOE estimates averaged 3 percent. The NOE is broader than the shadow economy, including also informal and other activities, such as household production for own use, omitted from NSIs' basic data collection programmes. For most OECD countries differences in the NOE are likely to reflect differences in the shadow economy.

[21] The survey is based on the publications in the references and correspondence with colleagues in Denmark, New Zealand, Sweden and the USA.

[22] A 2008 report by the OECD Forum on Tax Administration, which as well as OECD countries covered 15 non-members, found that only four countries, Denmark, Sweden, Chile and Mexico, had estimated top-down direct tax gaps.

[23] The Swedish tax administration, the Tax Gap Map (2008).

[24] The National Audit Office in its 2008 report, Tackling the Hidden Economy, states "In our work comparing how tax authorities tackle the hidden economy we found that none of them has yet found a reliable way to estimate the shadow economy."

[25] Pedersen (2003) estimates the Swedish personal tax gap as 3.3 percent of GDP in Sweden and 1.7 percent in the UK, but attributes much of the difference to the scope of taxation.

[26] Even the initial assumption that the shadow economy can be traced in cash demand is questionable, at least for the UK (Franklin, 2010).

[27] Not all studies are explicit about these assumptions. For example, other studies may be used to give an estimate of the shadow economy in a base year. The base year estimate enables the shadow economy in later years to be calculated from estimated changes, but the studies used for the base year make assumptions about another base year.

[28] Declaration of the ISWGNA (2006). The members of the ISWGNA are the European Commission (Eurostat), the International Monetary Fund, the Organisation for Economic Co-operation and Development, the United Nations and the World Bank.

[29] The uncertainty of the estimates, particularly for ghosts, makes a more definite statement on relative importance impossible. HMRC estimates the 2008/09 tax gap due to ghosts as £1.3bn with a lower limit of £0.3bn and an upper limit of £3.9bn and due to moonlighters as £1.9bn with a range of £1.2bn to £3.6bn.

The Individual Income Reporting Gap: What We See and What We Don't

Brian Erard, B. Erard & Associates and Jonathan Feinstein, Yale University

1. Introduction

Tax agencies are continually making decisions about the allocation of their resources across activities to promote tax compliance and combat evasion. The quality of these decisions is limited by their capacity to measure the overall level of compliance with taxpayer filing, reporting, and payment obligations, the frequency with which various types of transactions are misreported, and the characteristics of those who are responsible. In this paper, we provide an overview of how the IRS attempts to measure the degree to which filers of federal individual income tax returns properly report their incomes from various sources using data from the National Research Program (NRP).

The 2001 NRP provides a direct and nationally representative assessment of how much noncompliance IRS auditors are able to identify on individual income tax returns. However, willful tax evaders often undertake considerable efforts to conceal their misreporting, and NRP examiners are not always successful in uncovering this activity. The IRS therefore attempts to estimate not only the portion of the tax gap that we see from the NRP audit results, but also the portion that we don't.

Detection Controlled Estimation (DCE) is a statistical methodology that was initially developed by Feinstein (1990, 1991) to account for imperfections in examination processes (such as audits) to fully uncover violations (such as tax noncompliance). Under this methodology, one jointly models the detection process along with the underlying violation of interest. Under contract with the IRS, we have refined and generalized this methodology for application with National Research Program data to develop estimates of detected and undetected income underreporting for use in tax gap estimation. A key feature of the approach is that it accounts for differences among examiners in their ability to uncover noncompliance on tax returns. Intuitively, the methodology permits one to scale up the audit findings of less successful examiners to represent something closer to what the most successful examiners would have uncovered had they audited the returns.

A previous version of our methodology was employed in the development of the IRS estimates of the Tax Year 2001 individual income tax underreporting gap. Under that version, separate multiplier estimates were produced for "low visibility" and "high visibility" sources of income for each of two return categories ("business" and "nonbusiness"). Each multiplier represented an estimate of the ratio of the actual amount of underreporting present within that income source and return category to the amount that was detected during the NRP examinations. More recently, we have extended the methodology to produce more disaggregated estimates of detected and undetected underreporting by income line item in support of future tax gap estimates.

The DCE approach represents a significant departure from the earlier methodology employed by the IRS in developing its estimates of the tax gap. Under this earlier approach, an *ad hoc* adjustment was made to the portion of noncompliance on TCMP examinations that was identified by examiners without the aid of third-party information documents. While a common adjustment factor was applied to a wide range of income items, the value of this factor (3.28—meaning that there was an estimated $228 in undetected unreported income for every $100 of underreporting that was detected by examiners without the aid of third-party documents) had been derived based on findings from retrospective analysis of the random audits conducted under the Taxpayer Compliance Measurement Program (TCMP) for Tax Year 1976, the last year in which auditors in that program did not have the taxpayer's information documents available during the audit.

2. What We See

The NRP results provide details of what the taxpayer reported on each line item of the tax return as well as the NRP examiner's conclusion as to how much should have been reported for each line item. Since the returns were randomly selected, the results provide an indication of how much additional income (and tax) would have been detected if all federal individual income tax returns in the 2001 tax year population had been examined under the NRP process. They also provide an indication of what specific income and deduction items are commonly associated with compliance problems.

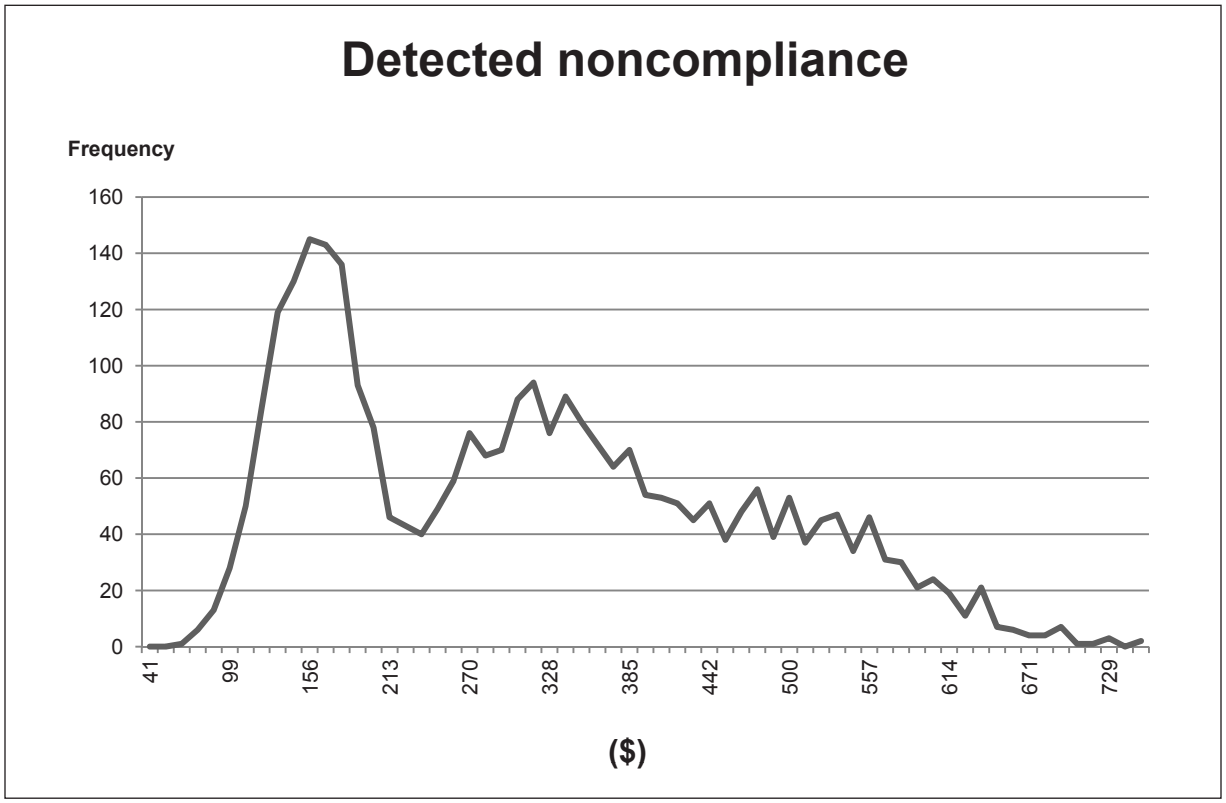
3. What We Don't See

In many instances taxpayers undertake considerable effort to conceal their tax transgressions from the tax authority. In such cases, it can be difficult for examiners to fully uncover all misreporting that is present. In general, one would expect that audit adjustments would allow us to observe many of the unintentional errors that taxpayers make in reporting their taxes, but only a portion of the deliberate cheating. Therefore, the raw NRP examination results are likely to provide an incomplete picture of the compliance landscape.

3.1 How to Measure What We Can't See

Intuitively, examiners will tend to vary in their experience and their skill at uncovering noncompliance. Some examiners may be globally superior at uncovering noncompliance with respect to all return issues; others may have a comparative advantage at uncovering noncompliance on particular issues. If we knew the relative abilities of different examiners to uncover noncompliance with respect to a particular tax issue or line item, we could "scale up" what was detected by a given examiner to approximate what the best examiner would have found in the audit.

FIGURE 1. Illustration of Distribution of Detected Noncompliance

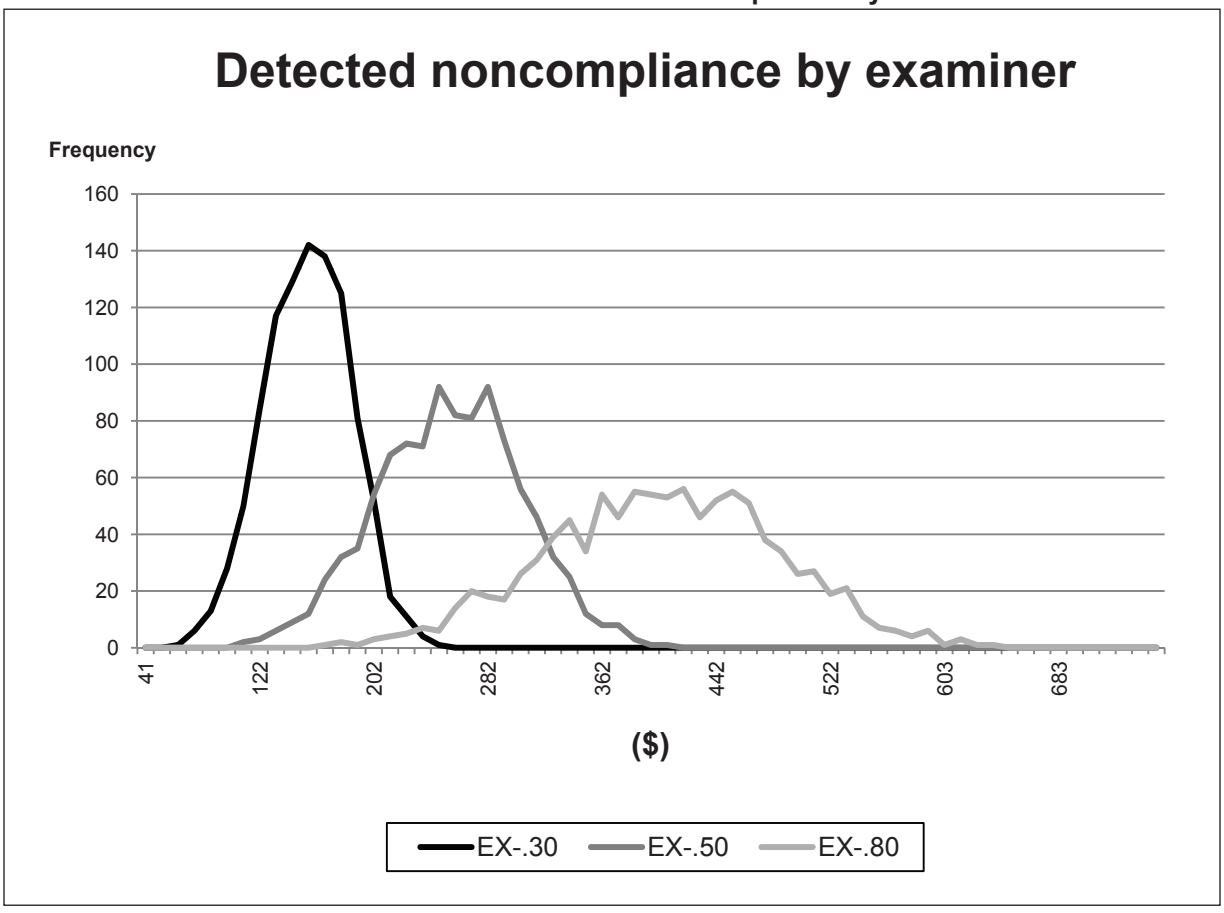


To gain a sense of how the DCE approach works, it is helpful to consider the following scenario. Suppose that we are shown the following plot of the distribution of detected underreporting with respect to a given income source based on a random audit study:

Based on our discussion so far, we recognize that the actual distribution of noncompliance may differ from the above detected distribution, but how can we account for noncompliance that has gone undetected during the study?

Imagine that you were told that three different examiners had been randomly assigned to audit a share of the returns included in this study and that it was possible to identify the detected amounts of noncompliance that were attributable to each examiner. Suppose that a more detailed plot that illustrates the distribution of noncompliance detected by each examiner looks as follows:

FIGURE 2. Illustration of Distribution of Detected Noncompliance by Examiner

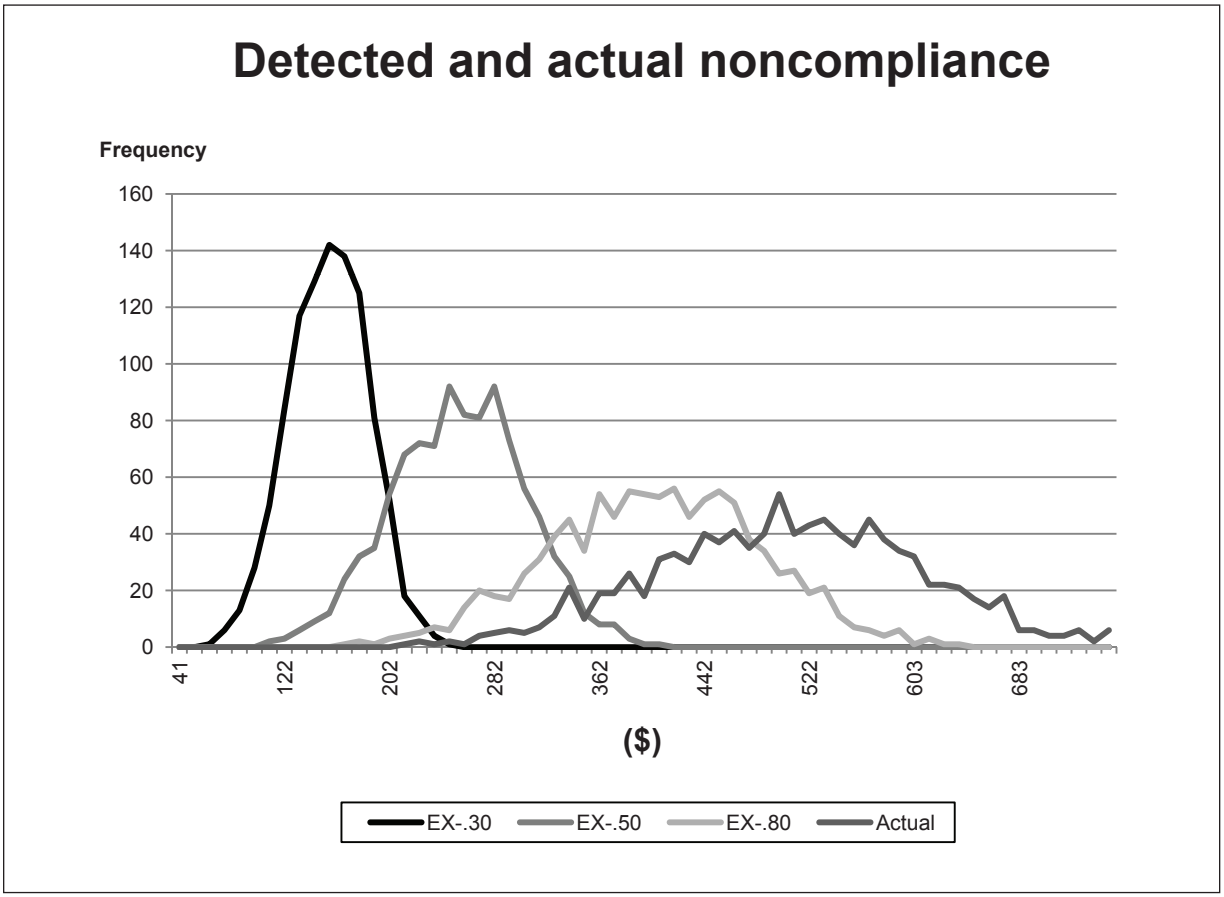


We now recognize that each of the examiners had a fairly unique detection pattern. In particular, the examiner associated with the darkest line tended to detect relatively modest levels of noncompliance, while the examiners associated with the lighter gray lines each tended to uncover progressively larger amounts of noncompliance.

Using the observed relative detection rates for the three examiners, one can anticipate how the results for the examiners with the lower detection rates might be scaled up to approximate what the examiner with the highest detection rate would have uncovered had he been assigned to perform the audits in their place.

The actual distribution of noncompliance (including both the detected and undetected amounts in the population) is superimposed in Figure 3:

FIGURE 3. Illustration of Distributions of Actual and Detected Noncompliance



These results are based on a simulation in which the three examiners were able to detect, on average, about 30, 50, and 80 percent of noncompliance on a given return, respectively. After scaling up the results for the two examiners with the lower detection rates, one might have a predicted distribution of overall noncompliance in the population similar to the lightest gray line in the above figure. Certainly, this is much closer to the actual distribution of noncompliance than was the detected distribution presented in Figure 1. However, a comparison with the line representing the actual distribution of noncompliance indicates that it still somewhat under represents the true mean and variance of noncompliance in the population. If one knew something about the shape of the actual distribution of noncompliance, one might be able to make a more refined estimate of the overall distribution of noncompliance that improves on the lightest line.

Essentially, this is what the DCE approach does. It compares the relative detection performances of different examiners and combines this information with what is assumed about the distribution of noncompliance (e.g., that it has a skewed shape similar to the lognormal distribution) to scale up the examination results for a given line item to better represent the true level of noncompliance on a given return. Under such an approach, results for examiners with a relatively low detection rate on a given line item (when compared against examiners with similar levels of experience) receive a more substantial adjustment than those with a relatively high detection rate. Typically, examiners with the very highest detection rates receive only a very modest adjustment, suggesting that they were able to fully uncover nearly all noncompliance that was present for the line item on the returns that they examined.

4. NRP Data

We have adapted and refined the DCE methodology to estimate noncompliance on all key income sources on individual income tax returns using NRP data. This database contains the results of examinations of a stratified

random sample of approximately 45,000 tax returns. About 10 percent of these returns were either accepted as filed or subjected only to a correspondence examination targeting a small number of issues on the return. In past analysis, we have found that such returns have a rather limited potential for significant amounts of undetected noncompliance. Consequently, the focus in our research has been on those returns that were subjected to a more intensive face-to-face examination. Our estimation sample includes approximately 38,000 returns.

A key feature of the NRP face-to-face examinations is that not all line items on a given return were examined. Prior to the audit, an experienced IRS examiner known as a "classifier" reviewed the return as well as other available information known to the IRS (such as third-party information returns and prior tax return filings) and made some decisions regarding what line items should be examined. Some income sources were routinely examined (such as Schedule C and Schedule F when such schedules were filed with the return). Other line items were subject to the classifier's discretion and were "classified" for examination or not on the basis of his experience and judgment in light of the information available to him. So, for instance, if the amount of wages and salaries reported on a return was consistent with the amount shown on the W-2 forms available to the classifier, this line item might not be classified for examination. When a line item was classified for examination, the NRP examiner was instructed to conduct an audit of that item. In cases where a line item was not classified, the NRP examiner in most cases did not audit the item. However, the NRP examiner did have the discretion to audit an unclassified item if noncompliance was suspected. For instance, if an initial probe uncovered potentially unreported income on a given line item, the examiner was free to pursue this issue.

The TY2001 NRP data also include a "calibration sample" of approximately 1,200 returns that were subject to more thorough examination—something closer to the detailed line-by-line audit process employed under the predecessor Taxpayer Compliance Measurement Program that was in place through tax year 1988. We incorporate returns from the calibration sample in our analysis of selected income items; see section 5.2 below for details on how these calibration sample returns have been employed.

5. DCE Models

We employ DCE models to develop estimates of income underreporting by line item for most income sources reported on individual income tax returns.[1] These income sources fall broadly into one of two categories:

1. Income items that are fairly routinely classified for an NRP examination (at least when a nonzero amount is reported for the item on the return). This category consists of income items that are not subject to a high degree of third-party information, such as self-employment income and rental income.
2. Income items that are not routinely classified for examination. This category includes income items that are subject to substantial third-party reporting, such as wages and salaries and interest income.

For the first category of income items, we have developed a DCE model that incorporates equations describing the likelihood and magnitude of noncompliance as well as the propensity for noncompliance to be discovered during an examination. For the second category, we have extended this model to account for the classification process and for discretionary examinations of unclassified income items.

5.1 Model for Line Items with Routine Classification

Line items for income sources that are fairly routinely classified for NRP examination include:

1. Schedule C net nonfarm self-employment income;
2. Schedule F net farm self-employment income;
3. Schedule D net long-term capital gains;
4. Schedule D net short-term capital gains;
5. Schedule E net rents and royalties;
6. Schedule E other net income (partnerships, s-corporations, estates, trusts, etc.);
7. Form 4797 net supplemental gains; and
8. Form 1040 other income.

When a nonzero amount is reported for one of the above income sources, the line item or schedule is generally classified for examination. In our discussion, we focus on the portion of our model that addresses this situation. We also have developed and estimated a DCE specification for the case in which no income was reported for these income sources. However, we omit discussion of this case for the sake of brevity.

For the above line items, we specify a DCE model with three building-block equations. In this model, we distinguish between the actual level of noncompliance associated with an income source (N) and the detected level of noncompliance for the income source (as measured by the NRP examiner's adjustment A). If detection were perfect, the actual level of noncompliance for the income source would be equal to the detected amount (i.e., $N=A$); however, our model accounts for the possibility that detection is imperfect, in which case the adjustment (A) will understate the true level of underreporting on the line item (N) by some unobserved amount.

In addition to accounting for undetected noncompliance, our model allows for the fact that many taxpayers make fully compliant reports with respect to any given income item on a return (i.e., $N=0$). We do this by modeling the true level of noncompliance using a two-part specification:

$$P^* = \beta_P' x + \varepsilon_P$$
$$\ln N = \beta_N' x + \varepsilon_N.$$

This two-part specification accounts for two of the three building block equations in our DCE model. In this specification, P^* represents is a latent variable describing the propensity for noncompliance with respect to the income source being modeled. The propensity for noncompliance is assumed to depend on a set of taxpayer and tax return characteristics (x) as well as a random disturbance term (ε_P). The term β_P represents a set of coefficients of the explanatory variables that we estimate. If P^* is less than zero (implying a relatively low propensity for noncompliance), then the income source is fully reported on the return and noncompliance (N) is equal to zero. On the other hand, if P^* is greater than zero (implying a relatively high propensity for noncompliance), then the income source is underreported on the return to some extent, meaning than N is greater than zero. In that case, the magnitude of noncompliance is determined by the second equation of the model, which relates the natural log of N to our set of explanatory variables (x) and an error term (ε_N). The term β_N represents a second vector of coefficients that we estimate.

We employ the following rather standard two-part modeling assumptions:

1. ε_P and ε_N are independently distributed;

2. ε_P follows the standard normal distribution (mean zero and standard deviation one);

3. ε_N is normally distributed with mean zero and standard deviation σ_N.

Under these assumptions, the conditional magnitude of noncompliance (when it is present) follows the lognormal distribution. This modeling structure is consistent with experience, which suggests that many taxpayers make fully compliant reports and, among those who do understate their income, many do so by relatively modest amounts, while a small minority underreport by very large amounts. We note that we have experimented with alternative distributional assumptions (such as the generalized gamma); however we have found that the lognormal distribution performs reasonably well and makes estimation somewhat more straightforward.

In a standard two-part model, one observes the values of the dependent variable. If we wanted to assume that detection was perfect, we could in fact estimate our above specification by setting true noncompliance (N) equal to the examiner adjustment A. Although we still would not observe the latent noncompliance propensity P^*, we would in this case observe the noncompliance indicator P, defined by the expression:

$$P = \begin{cases} 1 & P^* > 0 \quad (A > 0) \\ 0 & P^* \leq 0 \quad (A = 0) \end{cases}$$

Having observable measures of the dependent variables *P* and *N* (as well as the set of explanatory variables *x*) would make it feasible to estimate our two-part specification, which would then permit us to examine how taxpayer compliance behavior on a given income source is associated with various taxpayer and tax return characteristics *x*. However, if our objective was simply to measure the detected level of underreporting within the tax return population with respect to the income source, we would not even need to estimate a model. Rather, we could just aggregate the individual NRP examiner adjustments *A* on each return using the NRP sample weights.

Thus, the fundamental reason for the complexity in our approach is that we want to account for the fact that NRP examiners are not always successful at uncovering noncompliance, meaning that actual noncompliance *N* is sometimes greater than the adjustment *A*. Our model accounts for imperfections in the NRP detection process via the third building block equation of our model:

$$D^* = \beta_D' x_D + \varepsilon_D,$$

where D^* represents the propensity of the examiner to uncover noncompliance when it is present, x_D is a set of explanatory variables (including dummy variables for different NRP examiners, an indicator for whether the examination was conducted in the field rather than in the office, and the GS grade of the examiner), β_D is a vector of coefficients that we estimate, and ε_D is an error term assumed to follow the normal distribution with mean zero and standard deviation σ_D. We assume that this error term is independent of error terms in the first two equations. Let the detection rate (the fraction of noncompliance that the examiner is able to uncover be represented by *D*. Then the detection rate has the following relationship to the detection propensity D^*:

$$D = \begin{cases} 1 & (Perfect \quad Detection) \quad D^* \geq 1 \\ D^* & (Partial \quad Detection) \quad 0 < D^* < 1 \\ 0 & (Non - Detection) \quad D^* \leq 0. \end{cases}$$

Our DCE model then consists of three equations that respectively describe the likelihood of noncompliance with respect to an income source on the tax return, the magnitude of noncompliance if it is positive, and the extent to which any noncompliance has been detected. We jointly estimate the parameters from all three equations of our model jointly using the method of maximum likelihood.[2] Although incorporating imperfect detection into our two-part specification of noncompliance significantly complicates the likelihood function, it adds an important sense of realism to the specification while still keeping it tractable to estimate.

The likelihood function for a model is defined in terms of the conditional probability distribution of the observed dependent variables given the control variables (*x)*. So although the three equations of our model are defined in terms of the unobservable variables *P*, N,* and *D*,* the likelihood function must be defined in terms of the observed dependent variable *A*. In other words, we must derive the conditional distribution of *A* from the specified joint conditional distribution of these three unobserved response variables. Observe that the adjustment *A* is related to the actual level of noncompliance *N* and the detection rate *D* according to the following expression:

$$A = N * D.$$

Therefore, we can assess the conditional probability of an adjustment in the amount *A* by combining together the conditional probabilities associated with the various combinations of variables *N* and *D* that produce that value for *A*. To better understand how this process is carried out, it is useful to consider separately the cases where *A* is zero and where A is positive.[3]

When the adjustment *A* is equal to zero, there are two possibilities to consider:

1. The taxpayer was fully compliant in reporting the income item (i.e., *N=0*); or
2. The taxpayer understated the line item, but the examiner did not detect any of the noncompliance that was present (i.e., *N*>0 and *D*=0).

Observe that each of these cases will yield:

$$A = N^{*}D = 0.$$

The likelihood associated with the first case is defined by the probability that $P^{*} \leq 0$ (zero noncompliance). The likelihood associated with the second case is defined by the joint probability that *P*>0* (some noncompliance is present) and $D^{*} \leq 0$ (none of it was detected). The overall likelihood expression when the adjustment *A* is zero is computed as the sum of these two probabilities. Equivalently, it can be expressed as one minus the joint probability that *P*>0* and $D^{*} > 0$.

When the adjustment *A* is greater than zero, there are also two possibilities to consider:

1. All noncompliance was detected (i.e., *A=N*); or
2. Noncompliance was only partially detected (i.e., *A<N*).

The likelihood associated with the first of these cases is defined by the joint probability that *P*>0* (some noncompliance is present) and $D^{*} \geq 1$ (detection is perfect) multiplied by the probability density function for *N*, evaluated at *N=A*. Observe that for this case, we have:

$$A = N * D = N * 1 = N.$$

To determine the likelihood associated with the second case, one has to account for the fact that the detection rate *D* can take any value between 0 and 1. Therefore, it is necessary to integrate the joint probability density function for the adjustment *A* and the detection rate *D* over this range of values for *D*. This result is then multiplied by the probability that *P*>0*. Observe that for this case we have:

$$A = N * D < N; \;\; 0 < D < 1.$$

The overall likelihood expression when the adjustment *A* is positive is computed as the sum of the likelihood values associated with these two cases.

Estimation of the model yields estimates of the coefficients β_P, β_N, and β_D as well as the standard deviation terms σ_N and σ_D. Using these parameters, we are able to predict the conditional probability and magnitude of undetected noncompliance for a line item on a return given the NRP examiner's adjustment (detected noncompliance) *A*.

5.2 Model for Line Items with Non-Routine Classification

Line items for income sources that are not routinely classified for NRP examination include:

1. Wages and salaries;
2. Taxable interest;

3. Dividends;
4. State and local tax refunds;
5. Taxable pensions and IRAs;
6. Gross social security benefits;[4] and
7. Unemployment compensation.

For these income items, we extend our previous DCE model to incorporate a classification equation:

$$C^* = \beta_C' x_C + \varepsilon_C,$$

where C^* represents the propensity of a classifier to assign the line item to be examined on a return. In this probit specification, x_C represents a set of explanatory variables (including a set of classifier dummy variables as well as some variables measuring the discrepancy between the amount reported for a line item and the information shown on third-party reports) , β_C is a vector of coefficients to be estimated, and ε_C is a disturbance term that is assumed to follow the standard normal distribution. The return is classified for examination if and only if C^* is greater than zero. An important feature of our model is that we allow for the possibility that the classifier may observe some relevant information about the taxpayer (such as details from prior year tax returns) that is unavailable to us. We do so by allowing nonzero correlations between the classification equation error term ε_C and the noncompliance equation error terms ε_P and ε_N. These correlations account for factors not observed by us that may make it possible for the classifier to more effectively select which returns should be examined for a given line item.

When a line item is classified for examination in our model, an examination takes place and the three equations from our prior model continue to describe the probability and magnitude of noncompliance and the extent of detection. When a line item is not classified, we account for the possibility that the examiner uses his discretion and elects to audit the item. We do this by specifying the probability of an unclassified line item being examined as:

$$\Pr(Audit \mid Not \quad Classified) = \frac{\exp(\alpha_0 + \alpha_1 N)}{1 + \exp(\alpha_0 + \alpha_1 N)}.$$

Under this logistic probability expression, the likelihood that an unclassified income item is examined is depends positively on the level of noncompliance with respect to the item (N). What we have in mind here is that examiners who decide to audit an unclassified line item probably have uncovered some signal that significant noncompliance is likely to be present. As a consequence, the unclassified returns they choose to audit will tend to be the ones with relatively large levels of noncompliance for the line item. The parameters α_0 and α_1 are coefficients that we estimate along with the other parameters of our model.[5]

The introduction of a classification equation and a logistic specification for the likelihood of an audit of an unclassified item complicate the likelihood function. To avoid an overly technical discussion, we will not provide a detailed explanation of the derivation of the likelihood function for this case. However, we do note that the likelihood function now involves a distinct expression for each of the following cases:

1. Classified, Positive Adjustment;
2. Classified, Zero Adjustment;
3. Not Classified, Examined, Positive Adjustment;
4. Not Classified, Examined, Zero Adjustment; and
5. Not Classified, Not Examined.

With the introduction of a classification equation, the model becomes more difficult to identify; specifically, it can be challenging to reliably estimate the correlation terms between the errors of the classification and noncompliance equations in a model of this sort. To improve identification, we have incorporated observations from the calibration sample into our analysis. For the calibration sample observations, we assume that each of the line items was examined on all returns. Since there is no classification issue for these observations, they provide an independent source of information about the noncompliance equation parameters, thereby making it easier to distinguish between the coefficients of these equations and the correlation terms of the model.

Estimation of the model yields estimates of the coefficients β_P, β_N, β_D. β_C, α_0, and α_1 as well as the standard deviation terms σ_N and σ_D. Using these parameters, we are able to predict the conditional probability and magnitude of undetected noncompliance for a line item on a return given the classification and examination outcomes that have been observed.

5.3 Need for Joint Estimation of Line items

A key feature of our methodology is that it exploits heterogeneity among examiners in their ability to detect noncompliance. To do this effectively, one needs to have a reasonable number of examiners who have each audited a given line item on a significant number of returns (say, 15 or more). While this condition is satisfied for Schedule C and Schedule F reports, it is not generally satisfied for the remaining income items that are the subject of our analysis.

We have therefore undertaken a joint estimation strategy for estimating groups of income items subject to a common detection equation. Essentially, our approach assumes that a given examiner has the same potential for detecting noncompliance (when it is present) on any of the line items included in the group. However, the specification continues to allow for differences in detection abilities across examiners and across groups of line items. It is important to note that our joint estimation strategy does not restrict either the level or the rate of undetected noncompliance to be the same for different members of a group of income items. The level and rate can vary across group members, both because neither the likelihood nor the magnitude of noncompliance have been constrained to be the same for different income sources and because the sets of examiners that have audited each source do not perfectly overlap.

We have two distinct groups of income items that are employed under our joint estimation strategy. The first is our set of seven income items that are subject to a high degree of third-party information reporting:

1. Wages and salaries;
2. Taxable interest;
3. Dividends;
4. State and local tax refunds;
5. Taxable pensions and IRAs;
6. Gross social security benefits; and
7. Unemployment compensation.

Recall that we employ our DCE model for return line items with non-routine classification as described in Section 5.2 for this group of income sources.

The second group includes the following six income items, which are subject to less substantial third-party information reporting:

1. Schedule D net long-term capital gains;
2. Schedule D net short-term capital gains;
3. Schedule E net rents and royalties;

4. Schedule E other net income
5. Form 4797 net supplemental gains; and
6. Form 1040 other income.

Recall that this group of income sources is estimated using our DCE model for return line items that are subject to routine classification as described in Section 5.1.

Since we have sufficient examiners who have each audited a significant number of Schedule C and Schedule F returns, we estimate our DCE model (for return line items subject to routine classification) separately for these income sources without grouping them with other line items.

6. Results

In general, the estimated detection rates for each of our models indicate significant heterogeneity across examiners, ranging from very low (sometimes single digits) to near-perfect rates.

Table 1 presents the average implicit DCE multiplier for several categories of income items. The implicit multiplier represents the conversion factor to produce an estimate of overall noncompliance (detected plus undetected) from an estimate of detected noncompliance.[6]

The high third-party information reporting group includes wages and salaries, taxable interest, dividends, state and local tax refunds, taxable pensions and IRAs, gross social security benefits, and unemployment compensation. For this group, the overall implicit DCE multiplier is 2.52, indicating that there is approximately $152 in undetected noncompliance on these line items for every $100 that is detected. Table 1 also breaks down the implicit multipliers for the cases where an income item was classified for examination and where the item was not classified. When items in the group were classified for examination, the implicit multiplier was only 1.46; however, it was much higher (5.37) when items were not classified. Recall that the examination of unclassified income items was at the discretion of the NRP examiner, and in the majority of cases no examination was conducted. The higher multiplier for unclassified income items accounts for undiscovered noncompliance on the significant portion of returns that went unexamined for the line items as well as undetected noncompliance on the smaller portion that were examined.

Table 1 also provides the implicit DCE multiplier for the group of six income items (net short-term and long-term capital gains, net rental and royalty income, other Schedule E income, Form 4797 net supplemental gains, and Form 1040 other income) that were routinely classified for examination (when reported on the return). Although our description of the DCE specification in Section 4.1 focused on the case where these income items were reported on the return, we mentioned that we also estimated an econometric specification for the case where they were not reported on the return. The overall DCE multiplier (3.26) for this group presented in the table accounts for both cases, indicating that approximately $226 in noncompliance went undiscovered for every $100 that was detected for these items. Intuitively, this is higher than the multiplier for the high third-party information return income category as it is more difficult to detect noncompliance in the absence of comprehensive third-party information reporting.

For this group of income sources, Table 1 breaks the DCE multiplier down for the cases where the income items were and were not reported on the tax return. Examination was fairly routine within this group when the income items were reported on the return. For this case, the multiplier was relatively low (2.86). In contrast, examinations were less common when the income items were not reported. Intuitively, the multiplier was much larger for this case (4.80).

TABLE 1. Implicit DCE Multipliers by Income Category

Income Category	Multiplier	Income Category	Multiplier
High 3rd Party Information Reporting*		**Schedule C**	
Classified	1.46	Schedule Reported	2.92
Not Classified	5.37	Scheduled Not Reported	16.4
Overall	2.52	Overall	3.4
Routinely Classified*		**Schedule F**	
Items Reported	2.86	Schedule Reported	3.18
Items Not Reported	4.80	Schedule Not Reported	20.0
Overall	3.26	Overall	3.41

*These implicit multipliers are averaged over several line items, which were estimated separately.

Table 1 also presents implicit DCE multiplier estimates for Schedule C (nonfarm) and Schedule F (farm) net self-employment income. These estimates reflect the cases where these schedules were and were not filed with the return. As we would expect, the implicit multipliers for Schedule C (3.47) and Schedule F (3.41) are high relative to the other income categories as these income sources are subject to a very low degree of third-party information reporting.

Table 1 breaks down the overall multipliers for Schedule C and Schedule F to compare the cases where the schedules were and were not filed with the income tax return. In the former case, examination was routine and the multipliers were relatively low (2.92 and 3.18, respectively). In the latter case, examination was at the discretion of the examiner and often was not undertaken. The implicit DCE multipliers for this case were much higher (16.4 and 20.0, respectively), suggesting that a significant number of taxpayers failed to report their self-employment earnings and escaped detection during the NRP examination process.

Table 2 presents the estimated Net Misreported Amounts (NMAs) and Net Misreporting Percentages (NMPs) associated with the same income categories covered in Table 1. The estimated NMA represents the difference between the amount of income that federal individual income tax filers are required to report on their tax returns and what they actually report. Our estimated NMA accounts not only for the noncompliance on returns subject to face-to-face audits in our study, but also returns that were accepted as filed or subject to a correspondence examination. For the face-to-face audit cases, the results are based on our DCE estimates. For the other cases, our estimates rely on the additional tax recommended by the NRP examiners without any adjustment for undetected noncompliance. Based on prior analysis, we do not believe that the magnitude of undetected noncompliance for such cases is likely to be very substantial. Across all of the income sources we have analyzed in our study, the overall NMA is estimated to be $805 billion.[7]

TABLE 2. Net Misreported Amount and Net Misreporting Percentage by Income Source

Income Category	Net Misreported Amount ($B)	Net Misreporting Percentage
High 3rd Party Information Reporting	88	1.6%
Routinely Classified	359	28.5%
Schedule C	330	54.8%
Schedule F	27	51.4%
All Income Categories Combined	805	10.9%

The Net Misreporting Percentage (NMP) is also a measure of aggregate reporting noncompliance for a given income source, but is expressed as a rate. It is computed as the ratio of the NMA for the income source to the sum of the absolute values of the amounts that should have been reported across returns.[8] Consistent with prior IRS research, the estimated NMP for income items in the high third-party information reporting category of returns is very low (less than 2 percent). In contrast, the NMP for the group of six income items subject to routine classification (and less substantial information reporting) is much higher (28.5 percent). Again, this is an indication that income sources that are subject to less comprehensive third-party information reporting

tend to have significantly greater potential for noncompliance. For Schedule C and Schedule F, which are subject to very little third-party reporting, the estimated NMPs are even higher (54.8 and 51.4 percent, respectively). Across all of the income categories in our analysis, the overall NMP is just under 11 percent, suggesting that, as a group, U.S. federal individual income tax filers do report a very substantial portion of the income they are required to report on their returns.

Table 3 breaks down the shares of the overall estimated NMA attributable to different income sources. Overall, a very large share of the gap (41 percent) is attributable to underreporting of Schedule C net income. When Schedule F underreporting is included, the share of the overall NMA attributable to understated net self-employment earnings amounts to 44.4 percent. A similar share of the overall gap (44.6 percent) is attributable to the six income items in our category of income sources that are generally subject to only a modest degree of third-party information reporting. In contrast, only about 11 percent of the overall NMA is attributable to income items that are subject to reasonably comprehensive third-party information reporting.

TABLE 3. Share of Overall Income NMA by Income Category

Income Category	Net Misreported Amount ($B)	Share of Overall NMA
High 3rd-Party Information Reporting	88	11.0%
Routinely Classified	359	44.6%
Schedule C	330	41.0%
Schedule F	27	3.4%
All Income Categories Combined	805	100.0%

7. Directions for Future Research

There are several directions for future research. First, the IRS is working to incorporate our line item DCE estimates of income underreporting into its model for developing updated estimates of the individual income tax gap—the difference between individual income taxes that should have been paid and the amount actually paid on time without enforcement action. Second, it would be desirable to incorporate the DCE results into a micro-simulation model that would permit the IRS to analyze a wide variety of "what-if" questions regarding changes in the composition of the taxpayer population or the tax treatment of various line items on the return. Such a model would be particularly useful if it also incorporated other taxpayer information, such as taxpayer burden estimates. A third avenue for future research concerns how best to adapt the DCE methodology to account for the new NRP sampling design, which involves annual audits of stratified random samples of approximately 13,000 individual income tax returns per year. A fourth direction for further research relates to the application of the DCE methodology to other NRP data sources, such as the recent NRP studies of S-corporation and employment taxes.

References

Feinstein, Jonathan S. (1990) "Detection Controlled Estimation," *Journal of Law and Economics,* Vol. 33, No. 1, April, pages 233-76.

Feinstein, Jonathan S. (1991) "An Econometric Analysis of Income Tax Evasion and Its Detection," *RAND Journal of Economics*, Vol. 22, No. 1, Spring, pages 14-35.

Endnotes

[1] We employ our DCE analysis to control for undetected misreporting with respect to income sources. Although various offsets (such as itemized deductions and credits) are also subject to misreporting on tax returns, the burden of proof for such items is on the taxpayer to justify the amounts claimed. Consequently, our working assumption is that examiner adjustments for offset items are a reasonably accurate reflection of noncompliance with respect to these items.

[2] In some cases, the number of parameters to be estimated precludes simultaneous estimation of all parameters owing to computer memory limitations. In such cases, we divide the parameters into groups and employ an iterative stepwise maximization procedure that converges to the global maximum over all parameters.

[3] In our analysis, we treat negative audit adjustments (i.e., cases where the examiner has determined that income has been over reported) as an assessment of zero.

[4] We model underreporting with respect to gross rather than taxable social security benefits to focus on cases of direct misreporting. In other words, we wanted to focus on cases where taxable social security benefits were understated as a result of understating gross benefits. The degree to which gross social security benefits are taxable depends on the amount of income one has from other income sources. We wanted to exclude from our analysis cases of indirect misreporting of taxable social security benefits that were solely attributable to understatements of other forms of income.

[5] Since this probability expression depends on the unobserved true level of noncompliance N, it is necessary to integrate over the possible values that N may take given the observed examiner adjustment A when estimating our specification.

[6] Note that these implicit multipliers are not comparable to the aforementioned 3.28 multiplier derived from the 1976 TCMP study. That multiplier (because of the way it was derived) was applied only to amounts detected without the aid of information documents; it was not applied to all adjustments that examiners made as these implicit multipliers could be. The overall implicit multiplier corresponding to the 3.28 figure, which would be comparable to these figures, was therefore significantly less than 3.28. Of course, it is also important to keep in mind that the predecessor TCMP program involved intensive line-by-line audits of the entire tax return. As discussed, NRP audits are more selective and therefore may have greater potential for undetected noncompliance.

[7] Note that this is not a tax gap estimate but rather an estimate of aggregate income underreporting. It would be necessary to apply a tax calculator to assess the degree to which this income underreporting translates into an understatement of tax liability. In a full tax gap analysis, one would also account for the tax implications of misstatements of adjustments, itemized deductions, and credits.

[8] The denominator of the NMP measure was computed by adding the aggregate estimated level of underreporting with respect to a given income source to the sum of the absolute values of the amounts actually reported for this source across returns. This approach is somewhat different than the official IRS approach of first combining the estimates of reported and unreported income before taking absolute values. Consequently, the NMP reported in this study will tend to differ from the official IRS measure.

The Income Tax Position of Persons Not Filing Returns for Tax Year 2005

Joshua Lawrence and Michael Udell, Ernst & Young LLP,
and Tiffany Young, Yale University[1]

This research focuses on the federal income tax position of persons resident in the United States who do not file an income tax return. For Tax Year 2005, about 38.6 million persons resident in the U.S. did not file or appear on a filed income tax return. A lack of sufficient income is the main reason that persons do not file a federal income tax return. However, 11.8 million of these persons did not file, or appear on a filed tax return but had sufficient income that required them to file. A significant number of persons who did not file an income tax return should have filed and owed tax. Of the estimated $21.1 billion of total income tax liability associated with the 38.6 million persons who did not file a return, $13.8 billion was unpaid. Many persons who do not file a tax return still pay taxes, however, and many of these persons would be due a refund of some or all of the income taxes withheld had they filed. Among the 38.6 million persons who did not appear on a filed income tax return, we estimate that if they didn't have any income not already reported to the IRS, then $3.8 billion of income tax that was withheld by a third party on their behalf could have been refunded had they filed a tax return. In addition, many persons who did not file an income tax return could have been eligible for refundable tax credits such as the earned income tax credit and the child tax credit. As a result of not filing, as much as $4.9 billion of refundable credits was not claimed. Importantly, all of these estimates are based upon amounts of income reported by a third party on behalf of a person and do not account for any unreported income.[2]

The goal of this research was to understand the federal income tax position of persons not filing an income tax return. A secondary objective was to rely, to the greatest extent possible, exclusively upon information that was filed with IRS as part of tax administration. The lone source of data not meeting this requirement was family structure information (married household with and without children; single head of household with children; and single household) necessary to infer the size of the resident U.S. population and filing status. The analysis focuses on Tax Year 2005, but can be repeated for any year in which third-party tax return information is filed on behalf of individuals.

This research began with an estimate of the U.S. residential population for 2005 from the March 2006 Current Population Survey (CPS).[3] This population was then compared with the residential population shown on filed income tax returns for Tax Year 2005. Approximately 38.6 million persons resident in the U.S. did not appear on a filed tax return for Tax Year 2005.[4]

An estimated 30.8 million out of the 38.6 million persons who did not appear on a filed income tax return had a third-party information return filed on their behalf. This was estimated by drawing a random sample of all information returns filed for all persons for Tax Year 2005 and removing those information documents associated with a filed income tax return. By adding these 30.8 million persons to the 256 million persons who filed and appeared on income tax returns for 2005 and were resident in the U.S., we estimate that approximately 97 percent of the 296.5 million persons resident in the U.S. during 2005 were represented in the federal income tax system. Of the 38.6 million persons who did not appear on a filed income tax return, the federal tax system had some information from third-party filed information returns on 30.8 million and no information through third-party filed information returns on another 7.8 million.

If all 38.6 million persons were to have appeared on income tax returns, we estimate that there would have been an additional 22.8 million tax returns with the following income and tax characteristics:

$511.5	billion of total income from information documents
$233.7	billion of total income subject to tax
$21.1	billion of income tax liability
$11.2	billion of tax payments made
$13.8	billion of income tax due before either child or earned income tax credits
$3.8	billion of refunds due before either child or earned income tax credits
$4.7	billion of potential earned income credit
$2.8	billion of potential child tax credit
$13.0	billion of income tax due after both child and earned income tax credits
$8.7	billion of refunds due after both child and earned income tax credits.

The Data

The data set created for this project consists of a random sample of all information returns filed for Tax Year 2005 for persons regardless of whether a tax return was filed. Some of the approximately 256 million persons reported on a filed tax return had no information returns filed on their behalf. For example, many children do not have information returns but most appear on a filed tax return, whether timely-filed or late-filed. No-return persons (i.e., those who do not appear on a timely-filed or late-filed tax return) who had an information report filed on their behalf totaled 30.8 million, of which 30.4 million contained information about income and an additional 0.4 million contained information about attendance at an educational institution.

The 30.4 million no-return persons with income on information returns include:

16	million with social security benefits
10	million with wage income
8	million with interest income
6	million with pension, annuity, or IRA income

Together, the 256.0 million persons who appeared on filed tax returns and the 30.8 million no-return persons who had information reports provided by a third party indicate that the income tax system for Tax Year 2005 accounted for 286.8 million persons out of a total U.S. population of approximately 296.5 million (i.e., 97% of the U.S. population).

The information documents for the research consist of a sample of all information documents filed for individuals for Tax Year 2005. The sample included all of the information documents filed on behalf of a person and was based upon a random sample of social security numbers. Included in the sample were 1,256,000 information documents of which 1,121,000 (89 percent) were associated with timely filed income tax returns; 52,000 (4 percent) were associated with late-filed income tax returns; and 83,000 (7 percent) were associated with persons who did not appear on a tax return. In all, 28 different information documents were sampled including all Forms W-2, all Forms 1099 and 1098, all Forms K-1, and all Forms 5498.

The information returns report amounts of wages, interest, dividends, pensions and annuities, social security benefits, gambling winnings and partnership, S-corporation, and trust distributions, as well as gross proceeds from sales of capital assets and certain real estate transactions. In addition, any amounts of income tax withheld at the source and any amounts of estimated taxes paid were also included in the analysis. Some of the forms don't show amounts of income but rather amounts of interest paid on a mortgage or attendance at an institution of higher learning. The most important source of income absent from the data set was net income from a business operated as a sole-proprietorship, since this is a self-reported amount and no information return was filed to document that income.[5] Thus, to the extent that business income is not reported through third-party information reporting, this analysis will understate income.[6] The last piece of information provided for the data set was the age and gender of the person for whom the information report was filed.

Constructing Tax Families for No-Return Persons

Concepts

Information documents are filed for individual persons but tax returns can be filed for a number of living and support arrangements, including families (joint filing and head-of-household filing) as well as single persons.[7] No information about the living and support arrangements for persons who did not appear on a filed income tax return is available from third-party information returns. Yet this information is necessary to determine filing status and ultimately whether a no-return person should have filed a tax return.

To bridge this information gap, hypothetical tax returns were constructed by placing each of the 38.6 million no-return persons into a family structure using the extensive living arrangement information contained annually in the March Current Population Survey (CPS). The CPS does not identify the tax filing status of a person. However, it is useful as a national sample of households that contains extensive information about the relationships among persons residing in a household. For example, married couples are identified as well as their dependents residing in the household. This research exploits these relationships to identify the family structures that could result in the tax filing status as married filing jointly, head of household, or single.[8]

In using the CPS data, the analysis focuses on the person in each household who is the CPS survey's reference person and analogizes that person as the primary person on a filed tax return. In a few cases there were households with multiple families each of which could result in a possible tax return. In these instances multiple tax returns could result from within a single CPS household. The demographics from these multiple tax returns—ages and number of persons in each family—would be used to identify a person like a CPS reference person. This newly identified reference person would also form the basis of a family structure from which a tax filing unit would be created from the relationship information in the CPS. These constructed CPS reference persons, along with the official CPS reference persons and their family structures, were used to match against the primary persons identified in the information returns sample by the IRS.

Each family structure within a CPS household that could result in a tax filing unit was matched to the primary person and their family structure from filed tax returns to account for all 256 million resident persons who appeared on filed income tax returns for 2005.

All of the no-return persons who had information documents were also eligible to be matched as primary persons to the CPS reference persons.[9]

Implementation

The matching of CPS data to information return data used a predicted mean matching algorithm as follows: First, the entire CPS was sorted by age and income of persons, and then all of the persons in the entire information returns file (timely filers, late filers, and no-return persons) were sorted by age and income. This sort by age and income results in some no-return persons interleaved among timely and late filer persons as a result of their greater amounts of income. For example, a 36-year-old primary filer from a timely filed return might have had total income on information documents of $10,000, while a 36-year-old no-return person had total income on information documents of $50,000. After both files were sorted they were matched to one another by relative ranking order according to total income.[10] As a result, the no-return person would be matched with a 36-year-old CPS reference person with greater CPS income than would the 36-year-old timely filer person who had $10,000 of total income.

The IRS had provided the family sizes and ages of dependents for primary persons who filed a tax return and who were included in the sample of information returns. A primary person on a tax return in our sample (and their filing family size, ages of spouse and dependents, and their income) was matched to a similar CPS reference person (and their family size, age of spouse and dependents, and their income). The matched CPS reference person and their family members were then removed from the CPS file. The remaining CPS persons provide a demographic portrait of the no-return population of 38.6 million persons.

The predicted mean matching algorithm had also matched CPS reference persons to some no-return persons. The family structures and the ages of persons related to the CPS reference persons were then replicated from the inventory of no-return persons by drawing without replacement from the no-return persons who were not matched to CPS reference persons. These persons became the tax family members of the no-return person who was matched to the CPS reference person. The incomes of these persons from the information returns data became part of the total income of the tax return for the family. Tax families could result in married filing joint returns; head-of-household returns; or single filer returns. The CPS had 38.6 million persons while the no-return file accounted for 30.8 million persons, so 7.8 million additional persons were needed to complete the modeling of the CPS family structures. No information was created for the 7.8 million persons other than their ages.

One desirable feature of this algorithm is that not all of the CPS persons who had income were matched to persons who were either timely-filers or late-filers, if a no-return person had greater income than either of these types of filers. This was the case as many persons file tax returns when they do not have a filing requirement, and some persons do not file tax returns even when they do have a filing requirement. The approach used here allows for no-return persons to reach well up into the income distribution reported on the CPS. What is most important in this predicted mean matching application is that the relative incomes of persons from the CPS be matched with the relative income distribution from the tax return information. The absolute difference in income between a CPS family and a tax return constructed family does not have a role.[11] Total income for a return was calculated as the sum of all of the reported amounts on information returns for persons in the constructed family.[12] Total income was used to test whether there was a requirement to file a tax return.

The 38.6 million no-return persons were distributed to 22.8 million simulated tax returns as follows:

TABLE 1. Filing Status and Number of Simulated Returns of No-Return Persons, TY 2005

Filing Status	Married Filing Jointly	Single	Head of Household	Total
Number of Returns	7,265,506	13,663,951	1,856,469	22,785,926
Number of Persons:				
Total	19,939,576	13,663,951	5,029,472	38,633,000
21 or younger	5,274,781	—	2,693,728	7,968,509
14 or younger	3,956,874	—	1,879,345	5,836,219

Source: Ernst & Young LLP analysis of IRS data
NOTE: Details may not sum to totals due to rounding.

Constructing Tax Liability for No-Return Persons

Total income is but a first step in determining tax liability. Both adjustments to income and deductions are necessary in order to calculate liability. However, information returns generally report amounts of income received, not amounts paid that would be allowable as adjustments or deductions. One source of data within the IRS to model amounts of adjustments and deductions on tax returns is the Statistics of Income Division individual income tax file for Tax Year 2005. These data include information about income, adjustments and deductions for filed tax returns and were used to impute total amounts of adjustments and deductions on each return.[13]

Separate imputations were performed for adjustments and deductions for each of the filing statuses created as a consequence of placing no-return persons into family structures: married filing joint; head-of-household; and single. The SOI individual income tax file was first censored by removing tax returns with total income amounts greater than the largest no-return person tax return. This exclusion avoided out-of-scope information from influencing the imputations of adjustments and deductions. Since filed income tax returns include 1) returns that were required to be filed as well as 2) returns that were not required to be filed (due to insufficient income), there is some similarity between the filed tax returns and the potential returns of no-return persons, justifying them as a basis for imputing adjustment and deduction amounts.

Adjustments were imputed beginning with the frequency and dollar amounts of total adjustments from filed returns. Adjustments were calculated for each return type (married filing jointly, head-of-household, and single) and age of primary taxpayer and income range using the following two-step procedure. First a logit model was estimated as to whether there were adjustments or not, and second, conditional on the logit parameters, an amount of adjustments was estimated. Adjustment amounts were quite small for the no-return person potential tax returns.

Imputing itemized deductions followed a similar approach with a few exceptions. Unlike adjustment amounts that lack a third-party information return, there was/is information on state and local taxes withheld and on mortgage interest paid. Both are significant components of itemized deductions. Following in the same manner as adjustments, first a logit model on the choice to use the standard deduction or the itemized deduction was estimated using these data. Conditional on the logit parameters, the amount of total itemized deductions was imputed. Itemized deduction amounts were more frequent than adjustments, largely due to the presence in the no-return data of many home mortgage interest payments on Form 1098-T and payments of state and local income taxes on Form W-2.

With filing status, total income, adjustments, and deductions determined, tax liability was calculated. Offsetting tax liability were amounts of taxes withheld and estimated tax payments made (as provided by the IRS). The result was either a balance due or refund amount, or no balance.

The Tax Position of No-Return Persons

General Results Before Child Tax Credit and Earned Income Credit

Table 2 shows that of the 22.8 million simulated tax returns representing 38.6 million no-return persons, 5.2 million (23%) had a filing requirement, a tax liability of $21.1 billion, refunds due of $2.9 billion, and tax due of $13.8 billion. Of the $233.7 billion of total income potentially subject to tax, $196.7 billion (84%) was on these returns. Total income potentially subject to tax excludes amounts of non-taxable income, consisting mostly of social security income and some pension and annuity income.

TABLE 2. Tax Position of Simulated Tax Returns of No-Return Persons Before Child Tax Credit and Earned Income Tax Credit by Filing Requirement for TY 2005

[Amounts in Millions of Dollars]

Tax Position	Number of Returns	Number of Persons	Total Income	Distribution of Total Income	Total Tax Liability	Total Refund	Total Balance Owed
Filing requirement	**5,180,474**	**11,822,423**	**$196,694**	**85%**	**$21,148**	**-$2,909**	**$13,811**
Balance due	2,489,302	4,898,088	$127,778	55%	$17,891	$0	$13,811
No balance nor refund	670,967	1,425,122	$6,236	3%	$1	$0	$0
Refund	2,020,205	5,499,213	$62,680	27%	$3,256	-$2,909	$0
No filing requirement	**17,605,452**	**26,810,577**	**$37,041**	**16%**	**$0**	**-$917**	**$0**
No balance nor refund	14,496,123	20,550,448	$23,516	10%	$0	$0	$0
Refund	3,109,329	6,260,129	$13,525	6%	$0	-$917	$0
Total	**22,785,926**	**38,633,000**	**$233,735**	**100%**	**$21,148**	**-$3,826**	**$13,811**

Source: EY analysis of IRS data.

Note: Details may not sum to totals due to rounding.

The remaining 17.6 million potential returns (77%) were not required to be filed for income tax reasons, and accounted for 26.8 million no-return persons (69%) but only $37 billion (16%) of total income potentially subject to tax. These returns had no income tax liability, but would have been eligible for $0.9 billion in refunds prior to calculating the earned income or child tax credits.

TABLE 3. Tax Position of Simulated Tax Returns for No-Return Persons Before Child Tax Credit and Earned Income Tax Credit by Total Income for TY 2005

[Amounts in Millions of Dollars]

Total Income	Number of Returns	Number of Persons	Total Income	Disribution of Total Income	Total Tax Liability	Total Refund	Total Balance Owed
Less than $0	46,620	84,378	-$525	0%	$0	-$15	$0
$0–$1,000	10,246,223	14,156,661	$1,420	1%	$0	-$163	$0
$1,000–$5,000	4,682,284	7,438,069	$12,388	5%	$0	-$241	$0
$5,000–$10,000	2,578,042	4,393,008	$18,460	8%	$13	-$342	$5
$10,000–$20,000	2,069,468	4,078,717	$29,548	13%	$520	-$626	$282
$20,000–$30,000	1,062,587	2,431,272	$26,064	11%	$1,080	-$615	$525
$30,000–$40,000	686,449	1,776,836	$23,747	10%	$1,347	-$478	$583
$40,000–$50,000	383,297	1,038,386	$17,106	7%	$1,284	-$265	$563
$50,000–$100,000	798,718	2,485,362	$54,210	23%	$5,093	-$807	$2,145
Over $100,000	232,237	750,310	$51,316	22%	$11,811	-$273	$9,709
Total	22,785,926	38,633,000	$233,735	100%	$21,148	-$3,826	$13,811

Source: EY analysis of IRS data.

NOTE: Details may not sum to totals due to rounding

As shown on Table 3, the simulated tax returns for no-return persons can be sorted into three bins according to income and tax liability.

The lowest income bin, spanning up to $10,000 of total income potentially subject to tax, accounts for 26.1 million (67%) of the no-return persons, with a total tax liability of $13 million and a balance due of $5 million. This same group of persons, however, is eligible for $1.4 billion of refunds (36%).

The middle bin, spanning income between $10,000 and $50,000, accounts for 9.3 million (24%) of the no-return persons. These returns had an estimated total tax liability of $3.7 billion and a balance due of $1.7 billion but an almost equal amount of refunds at $1.4 billion.

The 3.2 million persons (8%) associated with potential returns having total income potentially subject to tax in excess of $50,000 account for $16.9 billion (80%) of the $21.1 billion in tax liability, $11.9 billion of tax due, and $1.1 billion in income tax refunds before the child tax credit. *Within this group are the largest 1,000 unweighted returns in the sample, shown on Table 4. These returns, weighting to 0.588 million in the population and accounting for only 1.9 million no-return persons (5%), had total income subject to tax of $80.0 billion (34%), total liability of $14.8 billion (61%) and the lion's share of tax due of $11.0 billion (80%).*

Nevertheless, there are still sizeable refunds even for this narrowly defined group of high-income simulated returns, as nearly $0.7 billion in refunds (17%) are to be found here. What do we know about these returns? Income on these returns is overwhelmingly from wage and salary compensation reported on the Form W-2, at $61 billion (76%), followed by rental income, royalty income, and real estate income reported on Form K-1 at $8.6 billion (11%), followed by business income of $6.4 billion (8%) reported on Forms 1099-MISC and Form K-1. Only 4% of the income of the top 1,000 records in this data set is from taxable social security benefits ($2.6 billion).

Social Security Beneficiaries

Many no-return persons are 62 or older and eligible for social security benefits. Another big story in these data is the tremendous amount of social security benefits that are not subject to income tax and not included in total income subject to tax on line 22. Almost 40% of no-return persons (15.139 million) are greater than 61 years old, but have potentially taxable income of only $43.0 billion (18% of total) despite having social security income totaling $151.3 billion. *Of this amount only $2.6 billion, or less than 2%, is includable in total income subject to tax.* Total tax liability is $2.1 billion, which is just 10% of the total tax liability of no-return persons ($21.1 billion). Table 5 shows the income distribution of the elderly no-return persons:

TABLE 4. Total Income on Largest 1000 Simulated Tax Returns (Unweighted) of No-Return Persons by Total Income for TY 2005

[Amounts in Millions of Dollars]

Total Income	Number of Returns	Number of Persons	Total Income	Distribution of Total Income	Total Tax Liability	Total Refund	Total Balance Owed
Less than $0	—	—	$0	0%	$0	$0	$0
$0–$1,000	—	—	$0	0%	$0	$0	$0
$1,000–$5,000	—	—	$0	0%	$0	$0	$0
$5,000–$10,000	—	—	$0	0%	$0	$0	$0
$10,000–$20,000	—	—	$0	0%	$0	$0	$0
$20,000–$30,000	—	—	$0	0%	$0	$0	$0
$30,000–$40,000	—	—	$0	0%	$0	$0	$0
$40,000–$50,000	—	—	$0	0%	$0	$0	$0
$50,000–$100,000	356,025	1,135,513	$28,666	12%	$2,946	-$379	$1,319
Over $100,000	232,237	750,310	$51,316	22%	$11,811	-$273	$9,709
Total	588,262	1,885,823	$79,982	34%	$14,757	-$652	$11,028

TABLE 5. Total Income for Primary or Spouse Filers 62+ Years of Age on Simulated Tax Returns of No-Return Persons by Total Income for TY 2005

[Amounts in Millions of Dollars]

Total Income	Number of Returns	Number of Persons	Total Income	Distribution of Total Income	Total Tax Liability	Total Refund	Total Balance Owed
Less than $0	19,865	37,369	-$225	0%	$0	$0	$0
$0–$1,000	5,803,448	7,373,401	$742	0%	$0	-$61	$0
$1,000–$5,000	2,693,236	3,877,368	$7,211	3%	$0	-$79	$0
$5,000–$10,000	1,482,875	2,237,739	$10,472	4%	$1	-$87	$1
$10,000–$20,000	653,687	1,039,777	$8,872	4%	$67	-$179	$51
$20,000–$30,000	122,977	223,012	$2,977	1%	$79	-$59	$41
$30,000–$40,000	64,373	136,550	$2,200	1%	$97	-$60	$36
$40,000–$50,000	29,334	55,959	$1,323	1%	$76	-$36	$27
$50,000–$100,000	64,197	123,010	$4,177	2%	$379	-$57	$185
Over $100,000	19,025	35,095	$5,288	2%	$1,359	-$89	$1,241
Total	10,953,017	15,139,280	$43,037	18%	$2,058	-$707	$1,584

Source: EY analysis of IRS data.

NOTE: Details may not sum to totals due to rounding.

But for the 19,025 tax returns with income in excess of $100,000, there would be only $0.3 billion of tax liability (3%) associated with 15.1 million (39%) no-return persons at least 62 years old. Most of the refund amounts are generated from small W-2 wage withholdings that, when taken in consideration with the general exemption of social security income, are overwithheld.

Age Profile of Persons on Simulated Tax Returns

The age profile of the primary person on the simulated tax returns for no-return persons on Table 6 also shows a skewed distribution toward the elderly. Approximately 38% of no-return persons are 62 or older. The no-return elderly have a dearth of taxable income accounting for just 17% ($40.2 billion) of the $233.7 billion of taxable income. Of course, this amount excludes approximately $150 billion of social security benefits that the

elderly received in 2005. In contrast, 83% ($11.5 billion) of the total balance owed ($13.8 billion) of no-return persons is associated with primary filers between the ages of 30 and 61, who also account for 71% of total income subject to income tax ($165.6 billion). This is where the bulk of taxable income is, as can be seen on Table 6, both of wages and other sources, and is where the bulk of the tax liability and balance due amounts occur.

Table 6. Age of Primary Filer on Simulated Tax Returns of No-Return Persons for TY 2005
[Amounts in Millions of Dollars]

Age	Number of Returns	Number of Persons	Total Income Subject to Tax	Distribution of Total Income	Total Tax Liability	Total Refund	Total Balance Owed
Under Age 18	158,492	162,396	$36	0%	$0	$0	$0
Age 18-22	619,402	829,568	$4,312	2%	$211	-$89	$91
Age 23-29	2,224,276	3,924,616	$23,560	10%	$1,570	-$415	$969
Age 30-39	2,770,975	6,690,577	$59,360	25%	$7,264	-$955	$5,496
Age 40-49	3,179,851	7,124,998	$61,831	26%	$5,680	-$982	$3,216
Age 50-61	3,071,198	5,197,422	$44,389	19%	$4,785	-$720	$2,837
Age 62-79	6,267,311	8,875,243	$24,868	11%	$907	-$446	$518
Age 80+	4,494,420	5,828,178	$15,378	7%	$731	-$219	$685
Total	22,785,926	38,633,000	$233,735	100%	$21,148	-$3,826	$13,811

Source: EY analysis of IRS data.

Note: Details may not sum to totals due to rounding.

Returns with Wages

Potential returns with wage income account for 46% of all no-return persons (17.8 million), 77% of total income potentially subject to tax ($178.8 billion), 68% of balance due amounts ($9.4 billion) and 85% of refunds ($3.2 billion) as shown on Table 7. More revealing, however, are the 0.176 million returns (less than 1.0% of simulated returns) with 605,601 persons (less than 1.6% of no-return persons) and total income subject to tax greater than $100,000. These returns include $36.9 billion of total income (16% of the total) and account for 45% of the balance due amount ($6.2 billion). *Less than 2% of no-return persons account for almost half of all balance due amounts, and all of the simulated returns for these persons include wage income.*

Earned Income Credit

A surprisingly large number of no-return persons, almost 12.3 million, are associated with simulated returns potentially eligible for the earned income tax credit. The information documents alone are not sufficient to establish eligibility for the earned income tax credit, and this analysis relies mainly upon income (as established with information documents) and family structure (as identified with the CPS data) to estimate an upper bound of eligibility and credit amounts. The 12.3 million represent 32% of no-return persons. These persons account for only 19% of total income subject to tax ($45.1 billion), 4% of tax liability ($0.8 billion) and 3% of tax due ($0.4 billion), but fully 34% of income tax refunds ($1.3 billion). *In addition, the amount of "unclaimed" earned income credit could have increased the total amount of earned income credit provided in TY 2005 by an additional $4.7 billion (14%) from $34.5 billion actually claimed to $39.2 billion.*

TABLE 7. Simulated Tax Returns of No-Return Persons with Wages by Total Income for TY 2005

[Amounts in Millions of Dollars]

Total Income	Number of Returns	Number of Persons	Total Income Subject to Tax*	Distribution of Total Income	Total Tax Liability	Total Refund	Total Balance Owed
Less than $0	9,841	21,115	-$135	0%	$0	-$15	$0
$0–$1,000	1,023,440	1,738,294	$454	0%	$0	-$17	$0
$1,000–$5,000	1,668,576	3,135,996	$4,492	2%	$0	-$152	$0
$5,000–$10,000	1,098,868	2,166,717	$8,109	3%	$11	-$260	$4
$10,000–$20,000	1,442,981	3,051,646	$21,106	9%	$432	-$495	$203
$20,000–$30,000	947,791	2,231,718	$23,243	10%	$969	-$564	$431
$30,000–$40,000	620,856	1,650,094	$21,475	9%	$1,213	-$462	$468
$40,000–$50,000	347,518	955,335	$15,485	7%	$1,177	-$248	$471
$50,000–$100,000	704,904	2,268,897	$47,668	20%	$4,431	-$758	$1,577
Over $100,000	176,845	605,601	$36,913	16%	$8,296	-$266	$6,227
Total	8,041,620	17,825,414	$178,809	77%	$16,530	-$3,236	$9,380

Source: EY analysis of IRS data.

NOTE: Details may not sum to totals due to rounding.

* Income Subject to Tax is the amount shown on line 22 of the Form 1040.

TABLE 8. Simulated Tax Returns of No-Return Persons with Earned Income Credit But Before Child Tax Credit by Total Income

[Amounts in Millions of Dollars]

Total Income	Number of Returns	Number of Persons	Total Income Subject to Tax*	Total Tax Liability	Total Refund Before EIC	Balance Owed Before EIC	Total EIC	Balance Due After EIC
Less than $0	6,999	15,748	-$61	$0	-$1	$0	-$3	$0
$0–$1,000	1,287,474	2,103,774	$518	$0	-$20	$0	-$394	$0
$1,000–$5,000	1,909,177	3,346,594	$5,070	$0	-$147	$0	-$768	$0
$5,000–$10,000	1,156,124	2,141,339	$8,478	$11	-$244	$4	-$964	$0
$10,000–$20,000	785,156	2,158,467	$10,638	$60	-$279	$30	-$1,591	$0
$20,000–$30,000	381,059	1,390,943	$9,353	$121	-$336	$41	-$724	$0
$30,000–$40,000	175,815	765,966	$5,913	$133	-$197	$40	-$140	$0
$40,000–$50,000	25,209	115,201	$1,113	$44	-$33	$12	-$35	$0
$50,000–$100,000	39,199	185,242	$2,624	$179	-$54	$60	-$48	$13
Over $100,000	7,283	32,248	$1,413	$294	-$5	$234	-$8	$226
Total	5,773,495	12,255,521	$45,059	$842	-$1,315	$420	-$4,674	$239

Source: EY analysis of IRS data.

NOTE: Details may not sum to totals due to rounding.

* Income Subject to Tax is the amount shown on line 22 of the Form 1040.

Focusing only on returns with total income less than $10,000, $11 million of the $13 million of total income tax liability on all simulated returns with income below this amount is on earned income tax credit returns. The income tax refunds on these same simulated returns prior to the earned income credit would have been an estimated $0.4 billion and almost entirely from wage withholding. The earned income tax credit added an additional $2.1 billion of potential refunds. These returns are not eligible for a refundable child tax credit (because adjusted gross income will not be greater than $11,000) and so the earned income credit would be the primary source of a refund through the tax system for these persons.

The earned income credit reduces the balance due amount of $420 million by 43% to $239 million on a total tax liability of $842 million and increases refunds from $1.3 billion to almost $6.0 billion. *The net effect on the tax position of all no-return persons as a result of the earned income tax credit is to reduce the balance due amount from $13.8 billion to $13.6 billion and increase refunds from $3.8 billion to $7.5 billion.*

Child Tax Credit

Only 7 percent of the simulated returns, with nearly 7 million persons (18%) could be eligible for the child tax credit. As with the earned income tax credit, there is not sufficient information to determine eligibility for the credit solely from the information returns. Assuming that every primary taxpayer on a simulated return was the primary source of support for a child claimed, Table 9 summarizes the potential universe of no-return persons with potential child tax credit eligibility.

TABLE 9. Simulated Tax Returns of No-Return Persons with Child Tax Credit But Before Earned Income Tax Credit by Total Income

[Amounts in Millions of Dollars]

Total Income	Number of Returns	Number of Persons	Total Income Subject to Tax	Total Tax Liability	Total Refund Before Child Tax Credit	Balance Owed Before Child Tax Credit	Total Child Tax Credit	Refundable Child Tax Credit	Balance Owed After Child Tax Credit	Total Refund After Child Tax Credit
Less than $0	842	3,367	-$15	$0	-$2	$0	-$2	-$2	$0	-$3
$0–$1,000	–	–	$0	$0	$0	$0	$0	$0	$0	$0
$1,000–$5,000	–	–	$0	$0	$0	$0	$0	$0	$0	$0
$5,000–$10,000	–	–	$0	$0	$0	$0	$0	$0	$0	$0
$10,000–$20,000	365,860	1,367,659	$5,676	$7	-$161	$3	-$229	-$222	-$3	-$383
$20,000–$30,000	351,305	1,369,285	$8,614	$85	-$322	$29	-$535	-$454	-$52	-$776
$30,000–$40,000	270,559	1,125,490	$9,384	$260	-$281	$87	-$540	-$323	-$129	-$604
$40,000–$50,000	164,466	713,231	$7,333	$333	-$166	$123	-$357	-$122	-$112	-$288
$50,000–$100,000	415,354	1,818,572	$28,512	$2,148	-$534	$803	-$895	-$86	-$6	-$620
Over $100,000	125,128	558,572	$27,975	$6,265	-$119	$5,229	-$274	-$2	$4,957	-$121
Total	1,693,513	6,956,084	$87,481	$9,098	-$1,585	$6,275	-$2,831	-$1,212	$4,655	-$2,796

Source: EY analysis of IRS data.

NOTE: Details may not sum to totals due to rounding.

The child tax credit reaches higher incomes than the earned income credit (which phases out at just above $37,000 of income in 2005). The 13.8 million persons on simulated returns eligible for the child tax credit are associated with $9.1 billion of income tax liability, as compared with 12.3 million persons with $0.8 billion of income tax liability on simulated returns potentially eligible for the earned income tax credit. The $2.8 billion of child tax credits on these returns reduces the balance due amount by 25% from $6.3 billion to $4.7 billion and increases refunds by 76% from $1.6 billion to $2.8 billion.

The net effect on the tax position of all no-return persons as a result of the child tax credit and the earned income tax credit is a reduction in the balance due amount from $13.8 billion to $13.6 billion to $13.0 billion, and the increase in potential refunds from $3.8 billion to $7.5 billion to $8.7 billion.

Returns with Both Earned Income Credit and Child Tax Credit

One-half of the simulated returns potentially eligible for the child tax credit would also be eligible for the earned income credit. These 767,164 returns had one-fifth of the income shown on all child tax credit returns (or $17.5 billion of $87.5 billion), and one-fiftieth (or 2%) of tax liability ($0.2 billion of $9.1 billion). However, these same returns account for 37% ($3.2 billion) of total refunds of $8.7 billion.

TABLE 10. Simulated Tax Returns of No-Return Persons with Child Tax Credit and Earned Income Tax Credit by Total Income

[Amounts in Millions of Dollars]

Total Income	Number of Returns	Number of Persons	Total Income Subject to Tax	Total Tax Liability	Total Refund Before CTC and EIC	Balance Owed Before CTC and EIC	Total Child Tax Credit	Refundable Child Tax Credit	Total EIC	Balance Owed After CTC and EIC	Total Refund After CTC and EIC
Less than $0	–	–	$0	$0	$0	$0	$0	$0	$0	$0	$0
$0–$1,000	–	–	$0	$0	$0	$0	$0	$0	$0	$0	$0
$1,000–$5,000	–	–	$0	$0	$0	$0	$0	$0	$0	$0	$0
$5,000–$10,000	–	–	$0	$0	$0	$0	$0	$0	$0	$0	$0
$10,000–$20,000	310,853	1,175,396	$4,821	$5	-$139	$3	-$195	-$190	-$1,067	$0	-$1,396
$20,000–$30,000	296,610	1,158,502	$7,291	$73	-$273	$25	-$458	-$389	-$617	$0	-$1,279
$30,000–$40,000	158,626	698,709	$5,329	$116	-$180	$32	-$341	-$234	-128	$0	-$543
$40,000–$50,000	1,076	4,829	$45	$1	-$2	$0	-$2	-$2	0	$0	-$4
$50,000–$100,000	–	–	$0	$0	$0	$0	$0	$0	$0	$0	$0
Over $100,000	–	–	$0	$0	$0	$0	$0	$0	$0	$0	$0
Total	767,164	3,037,436	$17,487	$196	-$594	$60	-$997	-$815	-$1,812	$0	-$3,222

Source: EY analysis of IRS data.

NOTE: Details may not sum to totals due to rounding.

Conclusion

Of the 38.6 million persons who did not appear on a filed income tax return (no-return persons) for Tax Year 2005, 30.8 million had at least one third-party information document filed on their behalf. This research created tax returns for no-return persons relying solely upon a sample of these information returns filed by employers, businesses, financial institutions, governments, real estate professionals, educational institutions, and trustees. Families were constructed for no-return persons using a predicted-mean-matching algorithm to map family structures (shown in the March Current Population Survey for the residential U.S.) to tax system information documents lacking that information. Tax returns were created relying upon these family structures. Tax liability was determined by imputing adjustments to income and itemized deductions, using information documents (for the home mortgage interest deduction found on the Form 1098, as well as state and local income taxes withheld), supplemented by actual levels of adjustments and itemized deductions found on filed tax returns.

The 38.6 million no-return persons resulted in 22.8 million tax filing units with $233.7 billion of income potentially subject to tax. Of the 22.8 million tax filing units, 4.5 million had a tax liability amounting to $21.1 billion; 2.5 million returns accounted for an estimated $13.8 billion in tax due; and 5.1 million returns accounted for $3.8 billion in unclaimed refunds. Using imperfect data to determine eligibility of the child tax credit and the earned income tax credit, an upper bound estimate of the impact of both of these credits for no-return persons would have reduced the amount of tax due from $13.8 billion to $12.9 billion and increased the amount of unclaimed refunds from $3.8 billion to $8.7 billion.

This research should not be construed as providing an estimate of the tax gap associated with persons who did not file a tax return. First, most income from self-employment does not have third-party information reporting and is not included in the data used in this analysis. As a result, total income for no-return persons is likely understated.[14] Second, no upward adjustment to reported income is used to account for unreported income in this analysis. The IRS, in preparing tax gap estimates, adjusts detected amounts of unreported income to account for undetected income. This analysis does not include this kind of adjustment. If it did, it is likely that nonfiling gap would be larger than suggested by this study—not just because of the tax on the addition income, but also because that income would likely reduce the amount of credits among the nonfilers. With these caveats, this analysis shows that substantial amounts of income tax are due by persons who do not file tax returns and that substantial amounts of refundable tax credits are likely unclaimed as a result of not filing a tax return.

Endnotes

[1] Josh Lawrence, Michael Udell, and Tiffany Young performed this research for the IRS Office of Research, Analysis and Statistics pursuant to contract TIRNO09-K-0053. Each of the authors worked for the Quantitative Economics and Statistics group (QUEST) at Ernst & Young LLP. Tiffany Young is at the Yale School of Management. Michael Udell is the corresponding author at Michael.udell@ey.com. Views in this paper are those of the authors and do not reflect those of Ernst & Young LLP or Yale University. This research is a continuation of the research program introduced in "Attaching the Left Tail: A New Profile of Income for Persons Who Do Not Appear on Federal Income Tax Returns," by Jacob Mortenson, James Cilke, Michael Udell, and Jonathon Zytnick, in Proceedings of the 102nd Annual Conference on Taxation, National Tax Association.

[2] In addition, this analysis does not identify how much of unpaid taxes were, or would be collected through IRS enforcement actions.

[3] The CPS is a random sample of households resident in the United States. Unlike tax returns, it does not include the residential population in institutions, such as prisons and college dormitories, and it does not include U.S. citizens resident overseas, so it does not cover the identical federal tax return population.

[4] In arriving at this estimate, the research program distinguished between persons who filed an income tax return late—as late as the end of 2008—from persons who never filed a return. The estimates presented here are only for persons who failed to appear on a filed tax return for Tax Year 2005 by 2008. The great majority of so-called late filers, persons who appeared on a tax return for Tax Year 2005 after October 15 of 2006 (when the second extension for timely filing of a 2005 return had eclipsed) had filed by 2007.

[5] While net income from a sole-proprietorship is not reported by a third-party source, some gross income of sole-proprietors is provided by a third party. Generally, payments made to a trade or business that exceed $600 per year are required to be shown on a Form 1099-MISC. Specifically, payments to attorneys, payments for fish purchased directly from fisherman, and most payments for health care made through an insurance company are reported on a Form 1099 and were included in the data set.

[6] The IRS, when estimating the tax gap, uses various factors to account for unreported income. None of these factors, or unreported income multipliers, was used in this analysis so the estimates presented here will not be directly comparable to an estimate of the tax gap associated with no-return persons.

[7] This research did not create returns for the married filing separate filing status because the CPS lacks information that could be used to create that status.

[8] The CPS does not include information about the extent of support that one person provided to another in the household. Support tests are an important part of the ability to claim a dependent exemption for tax purposes. For this analysis it was assumed that if a person had a family relationship with a person then they would be in that person's family. The analysis also looked to the amount of income reported on the CPS for a person to determine whether they would have sufficient income on their own to be required to file a tax return. Having sufficient income and being older than 21 would prevent a person from becoming a dependent on another person's tax return for this analysis.

[9] IRS Office of Research staff provided filing status, primary person, family size and age, and total income information from filed income tax returns for the 1,121,000 information reports in the sample associated with timely filed 2005 tax returns and for the 52,000 information reports in the sample associated with late filed 2005 tax returns.

[10] During this matching process sometimes a record from one file is split into pieces with identical information but different weights so that the record weight for each person from the two files is the same. One of the challenges in implementing predicted mean matching algorithms is to define groups of returns from the two files with equal or nearly equal total weights so that record splitting does not result in unused records. When records from one file are not matched to records from the second file in a predicted mean matching exercise, the marginal distributions from the first file are not preserved, which defeats the purpose of the algorithm.

[11] The nature of CPS non-response of income amounts is not investigated in this research. Future analysis of the relationship between CPS non-response of income amounts and IRS third-party information reported amounts of income would refine the accuracy of the predicted mean matching algorithm by reducing a key source of variability between the CPS and the IRS tax return information.

[12] The reporting rules for income associated with dependent filers were modeled from IRS publications. In some cases, small amounts of income associated with a dependent on a return would not be included on the parents' return. In addition, some information returns that only report gross proceeds, such as the Form 1099-B (for gross proceeds from the sale of securities) and the Form 1099-S (for certain gross proceeds from the sale of real estate) were adjusted to reflect a "net" amount of income. For the Form 1099-S this meant that no amounts of income were reported in total income due to the

[13] This assumes that no-return persons are eligible to claim adjustments and deductions of the same magnitude as did similar people who filed returns.

[14] The understatement may not be as serious as it suggests because some self-employment income is provided on third-party information reports and the gross amounts of these sources of income are included in total income, rather than the net amounts that would be included in self-employment income.

4

New Disclosure and Regulation Issues

Contos ◆ Guyton ◆ LaVelle ◆ Myers
Brody

Use of Assets in Large and Mid-Size Corporations: An Econometric Analysis of the Manufacturing Sector Using Schedule M-3 Data

George Contos and John Guyton, Office of Research, Internal Revenue Service
Jean LaVelle and Deborah Myers, Large Business and International Division, Internal Revenue Service

In this paper we examine the relationship between the difference in the rates of return to assets reported to the shareholders and the IRS and the firm's ownership type, its capital structure, foreign activity, and other characteristics. Studying the book-tax difference in terms of rates of return on assets is important because, over the past two decades, the income reported by firms to the IRS has decreased while the financial income and the rates of return on assets reported to the shareholders has increased. The widening of the book-tax differences (BTD) in the 1990's and early 2000's is a possible indication of earnings management by corporations.

The prior literature studying the BTD in terms of income has found evidence of additional tax sheltering activity since the widening of the BTD can not be fully attributed to temporary and permanent differences or differences in consolidation. However, these works have been hampered by the lack of data. The financial statements provided by firms to the public and the tax returns reported to the IRS did not provide information on why the tax and book income increasingly differed. In a recent article Eames and Luttman (2010) discuss the problems with trying to compute effective tax rates for corporations with publicly available information. They find that when analyzing tax rates estimated using Compustat, aggregate SOI, and GAO data can result in contradictory and incomplete results.

New tax reporting requirements for business entities have improved the potential for research using this data to reveal more about book-tax differences. The new Schedule M-3, required for corporations with total assets of $10 million or more, reconciles the information on financial statements and the tax returns. On the M-3 firms are asked to report both worldwide consolidated net income and tax net income, amounts that may not be the same due to different book and tax consolidations, as well as to varying disclosures of which income and deduction items are temporary versus permanent.

Current literature using preliminary M-3 data has found that, on aggregate, the BTD exists and has reported some first information on the nature of the differences. A 2008 article by Boynton, DeFilippes, and Legel using preliminary tax year 2005 data found in the aggregate a large positive BTD. A 2007 article by Lisowsky and Trautman using firm level Schedule M-3 data for publicly traded companies found evidence of a strong positive relationship between differences in the rates of return on assets reported in the book and tax consolidations and the differences in the capital structure measures reported in the financial statements and the tax returns. Finally, a 2008 GAO report found that foreign controlled corporations reported lower tax liabilities than domestically controlled corporations.

While the book-tax literature is replete with evidence of an increasing divergence between what public corporations report for book versus tax purposes, nearly all early studies of this subject were seriously hampered by the quality and lack of data. The literature on trends in corporate effective tax rates, multinationals, and corporate tax planning suggests that temporary and permanent differences in book-tax treatment explain only part of the total difference and suggest further analysis of the differential use of the underlying assets. By choosing to perform an analysis of manufacturing based on a prior seminal work by Mills and Newberry (2001), we employ new data sources and expand the analysis to all manufacturing firms with end-of-year total

assets of $10 million or more. The research goal is to broaden our understanding of the relationship between assets and income, and how this relationship differs between book and tax reporting.

For this work we use tax return data for tax years 2005 to 2008. As mentioned above, this allows us to take advantage of the first years of Schedule M-3 data. The challenge with using this data is that tax year 2005 was a year of record profits for this sector (largely due the Oil and Gas major group); during tax year 2007 many firms began experiencing the effects of the recession, which started officially in December of 2007 and did not end until June of 2009,[1] making atypical changes to asset holdings and valuations. Graham et al. (2010) have shown that overall economic activity is closely related to BTD (Graham et al., 2010, 21) therefore our finding of earning management should be fairly robust.

The rest of the paper proceeds as follows. Section 2 discusses the existing asset structure literature, in particular the BTD literature, and identifies findings and deficiencies in the literature. Section 3 discusses the data currently available and associated issues. Section 4 reports on the replication of the Mills and Newbury (2001) analysis for all manufacturing firms and describes our extended framework. Section 5 presents summary statistics for the full sample and by income and ownership type. Section 6 presents the results of the analysis. Finally, Section 7 presents our conclusions and offers recommendations for further research.

1. Literature Review

A large body of research has examined the differences in the asset structure of firms for evidence of tax sheltering activities. These works have tested the relationship between differences in the nature and physical location of firms' assets and differences in their revenues, rates of return on assets, and tax liability. Based on their scope and the empirical approach used, the asset structure literature can be divided into several streams. The common thread of the literature is the need to account for and interpret the widening of the book-tax differences.

Tax income is calculated following the Internal Revenue Code (IRC) while financial (book) income is calculated under the General Accepted Accounting Principles (GAAP) issued by the Financial Accounting Standards Board (FASB). The book-tax differences arise because income and expenses are treated differently under the IRC and the GAAP, but also because the two measures are reported for different purposes. Tax income is reported to the IRS and is used to determine the firm's tax liability, so corporate management has an incentive to report lower taxable income. Financial statements, on the other hand, are reported to the firm's shareholders and often to the public, for reporting corporate profits. Here, corporate management has a clear incentive to report higher financial income.

A number of authors have analyzed the causes of the BTD (see Plesko (2000, 2004), Boynton, DeFilippes, and Legel (2005 and 2008)). They found, as mentioned earlier, that the BTD widened significantly over the 1990s and the early 2000s.[2] The BTD differences come from a variety of sources. They include temporary differences, such as differences in allowable depreciation, which eventually reverse, and permanent differences, which do not reverse.

Temporary differences in allowable depreciation result from the use of different depreciation schedules for book and tax purposes. For example, under the IRC a firm can typically accelerate its depreciation deductions relative to the treatment under GAAP. The difference in timing of the available depreciation deduction reduces the firm's tax income in the short run. As long as the firm's depreciable assets are increasing, this book-tax difference should increase.

Permanent differences are more important to the IRS and the shareholders than temporary differences because they never reverse and permanently reduce tax income while increasing book income. An example of a permanent difference is non-qualified stock options that are expensed for tax purposes but not accounted for in financial statements. Firms that use stock options as a form of employee compensation can effectively keep significant amounts of wage expenses out of their financial statements without increasing their tax liability.[3] As a second example, tax exempt interest is included in the financial income but not in the tax income. Third, differences may arise from the treatment of discontinued operations and other large "extraordinary" write-offs and expenses. The treatment of nonqualified stock options has received perhaps the most attention (see Manzon and Plesko, 2002; Desai, 2003; and Hanlon and Shevlin, 2005). Although these works found that firms

used stock options extensively in the 1990's they found evidence of additional tax sheltering activity since the widening of the BTD can not be fully attributed to temporary and permanent differences.

Another class of differences relates to the rules for consolidation of subsidiaries (see Mills, Newberry, and Trautman, 2002; Mills and Plesko, 2003; and Plesko, 2004). For financial purposes, a parent company must include in its consolidated financial statements all domestic and foreign subsidiaries for which it owns an interest of 50 percent or more. Under tax rules, however, domestic subsidiaries must be 80 percent or more owned to be included in the parent's tax return and foreign subsidiaries cannot be consolidated.

Another reason financial and tax data may differ is off-balance-sheet financing. Some firms have used special-purpose entities (SPEs) to keep debt outside their consolidated financial statements. SPEs that are 80 percent or more owned are included in the tax consolidation but not in the financial consolidation. Such entities are classified under FIN 46R and they do have to be included in the book consolidation (Lisowsky and Trautman 2007). Mills and Newberry (2005) find "that these financial reporting effects occurred primarily during 1994-1999." (Mills and Newberry 2005, 251) These works found evidence that the financial statements of large firms for that period underreported both interest paid and debt, and inflated book income while leaving tax income unchanged.

Finally, the BTD may exist because some companies in their tax returns did not properly eliminate intercompany assets, such as accounts payables and receivables, and dividends. By improperly eliminating intercompany assets, firms increase the assets reported on their tax returns while keeping tax income unchanged and hence understate their rates of returns on tax assets. By improperly eliminating intercompany dividends firms overstate book income while leaving tax income unchanged. Mills, Newberry, and Trautman (2002) report anecdotal evidence of such reporting issues.

Taken together, the BTD literature finds evidence that the BTD widened but does not fully explain all the reasons behind its growth. Data limitations have hampered this research. Previously, the financial statements provided by firms to the public and the tax returns reported to the IRS did not provide information on why the tax and book income increasingly differed. As Weiner (2007) points out, "the real problem with the book-tax income gap was not that it existed but that no one really knew why it existed and why it was growing" (Weiner 2007, 853).

To improve the quality and expand the information available to tax administrators starting in Tax Year 2004, the IRS introduced a new reconciliation schedule. The new Schedule M-3, required for corporations with total assets of $10 million or more, further details the reconciliation of the information on financial statements and the tax returns. On the M-3, firms are asked to disclose if the book-tax differences for a large number of income and deduction items are temporary or permanent. Firms are also asked to reconcile their worldwide consolidated net income, reported on the SEC Form 10-K, with the net income for the tax consolidation.

As mentioned earlier, a 2008 article by Boynton, DeFilippes, and Legel, using tax year 2005 data, found that the BTD equaled $15,440 million. For returns with a positive BTD they found a return on assets of 5.1 percent, while for returns with a negative BTD the return on assets was 1.7 percent. For tax year 2006 they found that the BTD was significantly larger at $146,411 million and the return on assets was 3.5 and 0.6 percent respectively. We have found similar results for the manufacturing industry during this time period.

The Lisowsky and Trautman (preliminary 2007) article used tax year 2004 Schedule M-3 data. They examined the relationship between the book and tax rates of return on assets on 4,346 large public companies and their book and tax financial statements. They found at the firm level evidence of a strong positive relationship between differences in the rates of return reported in the book and tax consolidations and the differences in the capital structure measures reported in the financial statements and the tax returns.

2. Data Availability

The current study probes book-tax differences by using tax return data not previously available to researchers in the preceding decade. The sample is composed of companies that were selected by the SOI sampling process in each tax year from 2005 to 2008. The sample selection process is set up in such a manner that any corporation selected into the sample in a given year will be selected again the next year so it is possible to form both

unbalanced and balanced panels from the annual cross-sections. The advantage of the SOI data over other IRS datasets is that it provides more complete statistics. For example, all income statement, balance sheet, and Schedule M-3 fields are edited. In addition, SOI processing edits the data into consistent, standardized form that minimizes any "taxpayer reporting" noise in the data for analyses such as BTD.

To confine the data to medium and large sized manufacturing firms, we excluded all returns in five categories. First, we excluded all firms with end-of-year total assets of less than $10 million. Second, we excluded all foreign-controlled firms so that all entities in the sample have similar characteristics and incentives. Third, foreign corporations that engaged in trade or business in the U.S. or had income from sources within the U.S.—which are required to file Form 1120F—were also excluded because SOI does not collect balance sheet information for such returns. Fourth, part-year returns were excluded because they report income and deduction amounts for the short tax period but the balance sheet data reflect full year amounts. Finally, we exclude from our sample a small number of firm-year observations with extreme values.[4] The final dataset consisted of 17,794 firm-year observations for 6,480 corporations.

3. Econometric Modeling Framework

Mills and Newberry Framework

In their 2001 paper, Mills and Newberry (MN) developed an econometric model that utilizes many of the variables of interest. The authors had access to the IRS Coordinated Examination Program (CEP) data (primarily companies with end-of-year assets of $500 million or more) for manufacturing firms covering tax years 1981-1996. CEP companies know that they can expect more audit scrutiny than the general population. Therefore, the authors argue that the CEP dataset allows for a more homogenous group of taxpayers in terms of reporting behavior; these taxpayers should be more inclined to reduce book-tax reporting differences since they are under constant audit. Effects found to be significantly different from zero for this group should provide strong evidence of the effects in the broader population of returns.

The CEP has been replaced by the Coordinated Industry Case (CIC) program, which operates in the same manner as CEP but includes a greater range of firms. There are seven criteria used for the identification of CIC Program returns, per IRM Exhibit 4.46.2-2, including total gross assets, total gross receipts, total number of operating entities, number of multiple industries, total foreign assets for all Forms 5471, total related transactions on Form 5471/5472, and total foreign taxes. Approximately, 8 percent of all firms in our sample are in the CIC program, while for firms with end-of-year assets of $500 million or more 37 percent are CIC. Firms know that as their size, complexity, and foreign activity increase their chances of being in the CIC program. On the other hand, this population has passed through a set of filters that attempt to identify firms justifying a continual examination process. In meeting these criteria, there may be reason to suspect that BTD would be greater, despite the additional scrutiny. As such, it is unclear what to expect as to the direction of the effect of being in the CIC program.

MN's work used book income reported on the Schedule M-1 of the Form 1120 corporate tax return, providing self-reported book income for both private and public firms. As mentioned previously, the IRS has collected book data since 2004 on a more comprehensive Schedule M-3, which we used for our analysis.

We had several objectives in choosing to replicate MN's results. First, several control variables used by the authors were also identified in earlier analyses as important data partitions (Batson et al. 2010). Using a similar model, we could explore the importance of these variables in a multivariate regression framework. Second, the Schedule M-3 contains enormous detail of book income for both public and private firms.[5] While various sources exist to provide book data for public firms, researchers were often left to use costly and onerous surveys to collect this data for private firms. Even book data reported on public financial statements may differ from the book data for public firms as reported on Schedule M-3 for several reasons (e.g. different consolidations rules, or foreign subsidiaries). To the extent that Schedule M-3 data could be used to identify known effects, researchers should be able to exploit the greater detail provided to explore new relationships among book-tax reporting differences. Finally, by conducting our analysis on data from very recent tax years (2005-2008) and comparing it to the MN results using a decade or older data, it would be informative to see if

the significance levels and magnitudes of the effects have changed.[6] The model, which draws directly from the MN specification, can be expressed linearly for firm *i* in year *t* as follows:

$$BTD_{i,t} = \beta_0 + \beta_1 Public_{i,t} + \beta_2 Debt_{i,t} + \beta_3 Debt_{i,t} * Public_{i,t} + \beta_4 Debt_{i,t} * Distress_{i,t}$$
$$+ \beta_5 FTC_{i,t} + \beta_6 CapInt_{i,t} + \beta_7 Size_{i,t} + \beta_8 Distress_{i,t} + \beta_{9\text{-}11} Year_{i,t} + \beta_{12\text{-}32} Major$$
$$Industry_{i,t} + e_{i,t}$$

BTD is the pretax book income from Schedule M-3[7] less taxable income[8] (before Net Operating Loss deduction), scaled by end-of-year total assets.[9] Public is a dummy variable equal to 1 if public.[10] Debt is the ratio of long-term debt to total assets. FTC is the foreign tax credit claimed by income firms scaled by assets. CapInt is a measure of capital intensity (net property, plant and equipment to assets). Size is the log of assets. Distress is the probability of bankruptcy using tax data to compute Ohlson's predictor model[11]. BTD can be thought of as the difference in the book return-on-assets (ROA) versus the tax ROA for the tax consolidation .

4. Our Framework

The model again draws directly from MN's specification but we take advantage of the availability of Schedule M-3 data. We seek to better understand differences in ROA as reported for book and tax purposes. We test whether the relationship between BTD and ownership type and capital structure is the same even after controlling for temporary and permanent differences of key income and deduction items. We also test the sensitivity of the findings to other specifications of BTD. The model can be expressed linearly for firm *i* in year *t* as follows:

$$BTD_{i,,t} = \beta_0 + \beta_1 Public_{i,t} + \beta_2 Debt_{i,t} + \beta_3 Debt_{i,t} * Public_{i,t} + \beta_4 Debt_{i,t} * Distress_{i,t}$$
$$+ \beta_5 Amortization_{i,t} + \beta_6 Amortization_{i,t} * Public_{i,t} + \beta_7 Depletion_{i,t} + \beta_8 Depletion_{i,t} *$$
$$Public_{i,t} + \beta_9 Depreciation_{i,t} + Tax\ Exempt\ \beta_{10} Depreciation_{i,t} * Public_{i,t} + \beta_{11} FTC_{i,t}$$
$$+ \beta_{12} CapInt_{i,t} + \beta_{13} Size_{i,t} + \beta_{14} Distress_{i,t} + \beta_{15} Federal\ Deferred\ Taxes_{i,t} + \beta_{16} Tax$$
$$Exempt\ Interest_{i,t} + \beta_{17} CIC_{i,t} + \beta_{18} Consolidated_{i,t} + \beta_{19} CIC*Consolidated_{i,t} \beta_{20\text{-}23}$$
$$Year_{i,t} + \beta_{24\text{-}44} Major\ Industry_{i,t} + e_{i,t}$$

BTD is the pretax book income from Schedule M-3 less taxable income (before the Net Operating Loss deduction), scaled by total assets.[12] Amortization is temporary or permanent amortization of goodwill as reported on the 2005 Schedule M-3 Part III, Line 26. Depletion is temporary or permanent depletion as reported on Schedule M-3 Part III, Line 30. Depreciation is temporary or permanent depreciation as reported on Schedule M-3 Part III, Line 31. Federal Deferred Taxes are temporary or permanent deferred taxes as reported on Schedule M-3 Part III, Line 2. CIC variable is a dummy equal to one if the firm was a CIC firm. Consolidated is a dummy equal to one if the firm had consolidated subsidiaries.

We anticipate that both temporary and permanent differences in amortization of goodwill, depletion, and depreciation deductions are positively related to BTD, for both income and loss firms. As reported by Mark Jackson (2009), temporary depreciation differences are driven by earnings management and permanent depreciation differences by economic factors. The magnitude of the effects should be greater for temporary than permanent deductions since all three are typical examples of temporary differences that eventually reverse their effect on BTD.

For the tax exempt interest we again anticipate a positive relationship with BTD but with the permanent deductions having a greater effect since tax exempt interest permanently reduces taxable income. Finally, we are unsure about the relationship between the U.S. deferred income tax expense and BTD. A positive (negative) net deferred income tax expense, as reported on Schedule M-3, decreases (increases) book income while leaving tax income unchanged, hence decreasing (increasing) BTD. The effect of deferred tax expenses on BTD would depend on whether firms have net deferred tax liabilities or net deferred tax assets.[13]

In this set of regressions we also exclude the size variable because it is collinear with the debt variables, and we introduce two dummy variables that differentiate firms into three groups with similar incentives and ability to manage their book reporting. The first variable is a dummy equal to one if the firm was a CIC firm. The second variable is also a dummy equal to one if the firm had consolidated subsidiaries.[14] The third variable is the interaction of the two dummies. Again here we are unsure of the relationship between the variables and BTD. If CIC firms, as discussed by MN, are more inclined to reduce book-tax reporting differences, since they are under constant audit, the relationship will be negative. Else, if being selected in the CIC program is a reason to suspect earnings management by the firm, then the relationship will be positive.

5. Summary Statistics

The following two tables report summary statistics for the key variables for all Large Business & International (LB&I) manufacturing firms in the sample, covering tax years 2005 to 2008. Table 1 shows summary statistics for the full sample. The average book-tax difference over total assets for tax years 2005 to 2008 is -0.012, therefore the firms in the sample report on average negative book-tax differences. Separating BTD in its two components, we find that book ROA is -0.2 percent and tax ROA is 1.0 percent, with book ROA being more volatile than tax ROA. Finally, the firms in our sample have a fairly low debt-to-asset ratio, low probability for bankruptcy, and are capital intensive. We should caution the reader that these statistics are for the full sample, which includes firm-year observations with extremely high (low) ROA and debt-to-asset ratios.

TABLE 1: Summary Statistics for LB&I Manufacturing Firms, Tax Years 2005–2008†

Variable	Mean	Minimum	25th Pctl	Median	75th Pctl	Maximum
BTD/ASSETS	-0.012	-1	-0.024	-0.002	0.016	0.939
BOOK INCOME/ASSETS	-0.002	-2.927	-0.034	0.028	0.094	2.038
TAX INCOME/ASSETS	0.01	-2.707	-0.026	0.029	0.091	2.17
DEBT/ASSETS	0.28	0	0.05	0.192	0.413	5.859
FOREIGN TAX CREDIT	0.001	0	0	0	0	0.326
DISTRESS	0.07	0	0.001	0.005	0.026	1
CAPITAL INTENSITY	0.202	0	0.07	0.156	0.289	1
SIZE	17.824	16.118	16.591	17.276	18.546	27.82
TOTAL ASSETS*	1,054	10	16	32	113	**

* In millions of dollars. ** Amount deleted to prevent disclosure of taxpayer-specific information.

† Dataset consisted of 17,794 firm year observations for 6,480 corporations

Table 2 shows the average values for all key variables by income and ownership type. On average, the BTD of income firms is positive, 0.003, and for loss firms negative, -0.046. BTD is significantly higher for public than private income firms, 0.011 versus -0.001. The difference is primarily due to tax ROA, which is lower for public firms, 9.6 percent, than for private firms, 11 percent. For loss firms, book ROA, -24.0 percent, is significantly lower than tax ROA, -19.6 percent. Finally, income firms have lower long-term debt over assets ratios than loss firms, 0.246 versus 0.382, and lower probability of bankruptcy, 0.013 versus 0.204.

TABLE 2: Summary Statistics for LB&I Manufacturing Firms, by Income and Ownership Type‡

	Income Firms				Loss Firms			
Variable	All Mean	Public Mean	Private Mean	Public/ Private Test of Differences	All Mean	Public Mean	Private Mean	Public/ Private Test of Differences
FIRM YEARS*	10,205	3,203	7,002		5,340	1,771	3,569	
BTD	0.003	0.011	-0.001	9.87**	-0.046	-0.066	-0.37	-6.71**
BOOK INCOME	0.108	0.107	0.109	-0.43	-0.242	-0.252	-0.239	-1.35
TAX INCOME	0.105	0.096	0.110	-5.36**	-0.196	-0.186	-0.202	1.90*
DEBT	0.246	0.242	0.247	-0.33	0.382	0.326	0.41	-6.26**
DISTRESS	0.013	0.009	0.016	-5.57**	0.203	0.186	0.212	-2.79**
CAPINT	0.196	0.149	0.212	-19.84**	0.181	0.14	0.195	-12.11**
SIZE	18.567	20.263	17.733	83.20**	18.093	18.922	17.682	31.83**
TOTAL ASSETS***	2,096	5,834	112		558	1,202	82	

‡Approximately 2,249 firms year observations did not satisfy the income or loss firm criteria and were excluded from these statistics.

*The t-statistic for a test of differences in means is significant at the 0.05 level.

**The t-statistic for a test of differences in means is significant at the 0.01 level.

***In millions of dollars.

6. Findings

One limitation we faced in replicating MN's work was the lack of a long time-series dataset. The authors' 15 years of data allowed for a more robust analysis than we were able to undertake with four years of data. In contrast, we take advantage of the Schedule M-3 data. In addition, the SOI data we use provides statistics that are consistent and in a standardized form.

For replicating MN's analysis we chose to restrict the data to all Manufacturing sector firms with end-of-year assets of $500 million or more.[15] As discussed earlier, such firms would typically have been included in the CEP sample (Mills and Newberry, 2001).

Like MN, we find significant effects for several key variables. Our results, shown with the original results, are summarized in Table 3 below. Overall, our findings for income firms match their findings fairly closely, while for loss firms we find weaker relationships. Public firms in a positive book and tax income position have larger BTD than do private firms, when controlling for capital structure, foreign activity, size, capital intensity, and probability of bankruptcy. This result supports the theory that capital market pressures provide financial-reporting incentives for managers of public firms with assets greater than $500 million to report relatively higher book income during profitable periods. It is important to mention here that BTD in this framework can be thought of as the difference in ROA reported to the shareholders versus ROA reported to the IRS. However, for firms in loss positions, public firms do not report larger book losses than tax losses, as they did in the 1980's and early 90's. Therefore, managers of loss firms do not seem to maximize book losses. For the debt variable, we confirm that evidence exists that more leveraged income firms report larger book income (or smaller book losses) than tax income (or loss). This is most likely due to the additional scrutiny these firms face from creditors. It should be noted that the magnitudes of the Public and Debt coefficients for income firms are significantly higher than the coefficients estimated by MN. They interact debt and public, finding significant results, indicating that the effect of debt on BTD is reduced for public firms relative to private firms (as private firms face more binding constraints from debt.) MN also interact debt with the probability of bankruptcy and they find a positive relationship with BTD. We do match the authors' impact and significance levels for both interaction terms fairly well. We also find a strong negative effect for firms with foreign tax credits. As firms repatriate foreign earned income, taxable income increases, shrinking the book-tax reporting difference. Our year dummies support our earlier finding that over the 2005 to 2008 time period BTDs have decreased. This is in accordance with the findings of Graham, Raedy, and Shackelford (2010). They report that over the 1993 to 2008 time period book income exceeds tax income for all years, except 2001 and 2008.

Overall, we are encouraged to have found the anticipated relationship for the key variables in this broader group of firms.[16] We find this particularly encouraging given the unique exogenous influences taking place in the broader economy, and in the manufacturing sector particularly, during our observation years.

TABLE 3: Comparison of Regression Output to Mills and Newberry (MN)

	Income Firms			Loss Firms		
Variable	Predicted Sign	Current $500 Million or more	MN	Predicted Sign	Current $500 Million or more	MN
		Coefficient (T-statistic)	Coefficient (T-statistic)		Coefficient (T-statistic)	Coefficient (T-statistic)
Ownership type:						
PUBLIC	+	0.028** (4.49)	0.0008** (3.08)	-	0.010 (0.58)	-0.025** (-2.91)
Debt constraints:						
DEBT	+	0.045** (3.06)	0.020** (2.65)	?	0.013 (0.48)	0.036 (1.94)
DEBT*PUBLIC	-	-0.060** (-3.78)	-0.023** (-2.72)	?	-0.035 (-1.15)	0.040* (2.44)
DEBT*DISTRESS	+	0.079* (1.82)	0.185** (8.45)	?	0.096 (1.26)	-0.007 (-0.35)
Control variables						
FTC	-	-1.090** (-2.98)	-1.332** (-30.51)		NA	NA
CAPINT	?	0.010 (0.90)	0.010* (1.99)	?	0.061** (2.34)	0.002 (0.09)
SIZE	?	0.005** (3.52)	0.001* (2.00)	?	-0.007* (-1.85)	-0.009** (-4.79)
DISTRESS	?	-0.090** (-3.21)	-0.254** (-10.86)	?	-0.339** (-4.1)	-0.134** (-8.24)
YEAR 2006		-0.022** (-5.78)	Not reported		0.008 (0.94)	Not reported
YEAR 2007		-0.017** (-3.64)	Not reported		0.004 (0.37)	Not reported
YEAR 2008		-0.017** (-3.93)	Not reported		-0.046** (-3.34)	Not reported
Industry Dummies		Yes	Yes		Yes	Yes
Observations		2,021	4,956		566	820
Firms		779			346	
Adjusted R-squared		0.08	0.26		0.23	0.23

**,* The t-statistic is significant at the 0.01, 0.05 level respectively. OLS standard errors corrected for clustering by firm.

Next, we proceed to expand our analysis to the full sample of manufacturing firms. Table 4 reports findings for the full sample and also for firms with end-of-year assets less than $500 million. The public coefficients for both groups of income (loss) firms are positive (negative) and statistically significant. Therefore, these results provide evidence that capital market pressures provide financial-reporting incentives for managers of even smaller sized public firms to report relatively higher book income during profitable periods, and to maximize book losses during loss periods. We also find a strong negative effect for firms with foreign tax credits. For the debt variables, the coefficients have the expected sign but only the debt interacted with the probability of bankruptcy is statistically significant. Therefore, in the 2005 to 2008 time period it is not evident that highly leveraged firms report larger book income (or smaller book losses) than tax income (or loss).

Next we report the regression results of the expanded framework where, in addition to the ownership type, capital structure, and foreign activity controls, we also take advantage of the temporary and permanent

differences of key items as reported on Schedule M-3. Table 5 reports the findings for two groups of firms: the full sample and firms with end-of-year assets of $500 million or more. As expected, the majority of temporary differences of amortization of goodwill, depletion, and depreciation are positively and significantly related to BTD while permanent differences have the expected sign but are not statistically significant. Extrapolating from the Jackson (2009) finding, this could indicate significant earnings management by manufacturing firms. The opposite is true for tax exempt interest income, which showed permanent differences only, as expected, and are statistically significant. Moreover, both temporary and permanent U.S. deferred tax expenses for income firms have a negative and statistically significant effect on BTD, so income firms have net deferred tax liabilities. For loss firms, the signs differ but the coefficients are statistically significant at the 5 percent level.

TABLE 4: Regression Output for All Firms and Firms with Assets of Less Than $500 Million

		Income Firms			Loss Firms	
Variable	Predicted Sign	Full Sample	Less than $500 Million	Predicted Sign	Full Sample	Less than $500 Million
		Coefficient (T-statistic)	Coefficient (T-statistic)		Coefficient (T-statistic)	Coefficient (T-statistic)
Ownership type:						
PUBLIC	+	0.011** .(4.03)	0.009** (2.86)	-	-0.023** (-2.91)	-0.024** (-2.91)
Debt constraints:						
DEBT	+	0.001 (0.29)	-0.0003 (-0.08)	?	-0.008 (-0.90)	-0.007 (-0.73)
DEBT*PUBLIC	-	-0.010 (-1.32)	0.002 (0.13)	?	-0.001 (-0.06)	-0.006 (-0.25)
DEBT*DISTRESS	+	0.009 (0.82)	0.009 (0.78)	?	0.026* (1.76)	0.026* (1.70)
Control variables						
FTC	-	-1.300** (-5.35)	-1.420** (-4.45)		NA	NA
CAPINT	?	0.001 (0.20)	0.0004 (0.09)	?	0.023* (1.75)	0.023* (1.79)
DISTRESS	?	-0.026 (-1.54)	-0.021 (-1.19)	?	-0.172** (-12.57)	-0.170** (-12.27)
SIZE		0.003** (4.86)	0.002* (2.33)		-0.006** (-3.53)	-0.010** (4.03)
TAX YEAR DUM.		YES	YES		YES	YES
INDUSTRY DUM.		YES	YES		YES	YES
Observations		10,203	8,182		5,339	4,773
Firms		4,405	3,710		3,063	2,766
Adjusted R-squared		0.06	0.06		0.13	0.13

**,* The t-statistic is significant at the 0.01, 0.05 level respectively. OLS standard errors corrected for clustering by firm.

Finally, we find evidence that CIC income firms, as well as CIC firms that have consolidated subsidiaries, report higher BTD. We find that consolidated firms in general report less BTD than non-consolidated firms. This is an interesting finding because it seems to contradict the MN assumption that firms in CEP/CIC type programs behave as if they expect to be audited more frequently and closely than the rest of the corporate population. This finding may be due to the fact that large firms have more opportunity to manage their BT reporting (Dyreng et al., 2008) and that CIC firms are in the program because they have greater need for thorough examination. For firms in loss positions the CIC and the interaction coefficients are not statistically significant, while the consolidated firms reported higher book tax differences than non-consolidated firms.

TABLE 5: Regression Output of Expanded Framework for All Firms and Firms with Assets of $500 Million or More

Variable	Income Firms			Loss Firms		
	Predicted Sign	Full Sample Coefficient (T-statistic)	$500 Million or more Coefficient (T-statistic)	Predicted Sign	Full Sample Coefficient (T-statistic)	$500 Million or more Coefficient (T-statistic)
Ownership type:						
PUBLIC	+	0.013** (5.58)	0.027** (4.11)	-	-0.013* (-1.73)	0.036* (2.13)
Debt constraints:						
DEBT	+	0.0004 (0.11)	0.045** (2.78)	?	0.002 (0.26)	0.003 (0.17)
DEBT*PUBLIC	-	-0.011 (-1.48)	-0.065** (3.61)	?	-0.017 (-0.86)	-0.043 (-1.62)
DEBT*DISTRESS	+	0.008 (0.68)	0.044 (0.88)	?	0.016 (1.12)	0.049 (0.75)
Temporary						
AMORTIZATION	+	0.532** (5.69)	0.159 (0.35)	+	0.884** (19.15)	0.887** (6.28)
AMORT*PUBLIC		0.003 (0.02)	0.421 (0.81)		0.029 (0.27)	-0.032 (-0.2)
DEPLETION	+	0.836** (9.19)	0.893** (9.24)	+	-0.609 (-1.11)	-60.133 (-0.46)
DEPL*PUBLIC		1.503 (0.90)	-2.586 (-1.26)		-2.282 (-0.14)	-63.776 (-0.49)
DEPRECIATION	+	0.545** (8.00)	0.263* (1.73)	+	0.914** (6.74)	0.870** (2.53)
DEPR*PUBLIC		-0.45* (-2.13)	-0.041 (-0.17)		-0.150 (-0.58)	0.037 (0.07)
TAX EXEMPT INTEREST	+	0.050 (0.20)	0.232 (1.56)	+	0.857 (1.07)	4.724 (1.25)
DEFERRED TAXES	-	-0.714** (-3.16)	-0.634** (-3.60)	-	0.206* (1.74)	-2.74* (-1.78)
Permanent						
AMORTIZATION	+	0.815** (6.45)	1.643** (5.49)	+	1.027** (21.05)	1.048** (5.29)
AMORT*PUBLIC		-0.454 (-0.94)	-0.788 (-1.60)		-0.334* (-2.15)	-0.111* (-0.54)
DEPLETION	+	0.573 (1.24)	-0.497 (-0.14)	+	-4.672 (-1.21)	-337.15** (-10.26)
DEPL*PUBLIC		-0.326 (-0.46)	0.475 (0.13)		11.371** (3.26)	341.71** (10.56)
DEPRECIATION	+	-0.010 (-0.01)	1.686 (0.66)	+	6.306* (2.24)	23.227** (2.66)
DEPR*PUBLIC		-23.68 (-1.25)	1.337 (0.10)		-4.587 (-1.13)	-19.697* (-1.91)
TAX EXEMPT INTEREST	+	0.422* (1.78)	1.639* (1.94)	+	0.454* (2.02)	0.172 (0.24)
DEFERRED TAXES	?	-0.505** (-3.66)	-1.059** (-4.60)	?	-0.370* (-2.30)	0.260* (0.53)

Footnotes at end of table.

TABLE 5: Regression Output of Expanded Framework for All Firms and Firms with Assets of $500 Million or More—Continued

Variable	Income Firms			Loss Firms		
	Predicted Sign	Full Sample	$500 Million or more	Predicted Sign	Full Sample	$500 Million or more
		Coefficient (T-statistic)	Coefficient (T-statistic)		Coefficient (T-statistic)	Coefficient (T-statistic)
Control variables						
FTC	-	-1.320** (-5.36)	-1.019** (-3.42)	-	NA	NA
CAPINT	?	-0.003 (0.75)	-0.006 (-0.39)	?	0.01 (0.94)	-0.005 (-0.18)
DISTRESS	?	-0.028 (-1.68)	-0.098** (-2.93)	?	-0.131** (-10.54)	-0.224** (-3.28)
CIC	?	0.007* (1.84)	0.007 (1.61)	?	-0.007 (-0.48)	-0.029* (1.80)
CONSOLIDATED	?	-0.002 (-1.36)	-0.006 (-1.15)	?	0.010** (2.65)	-0.006 (-0.75)
CIC*CONSOL.	?	0.023** (3.67)	0.015* (2.25)	?	-0.0003 (-0.75)	0.009 (0.66)
TAX YEAR DUM.		YES	YES		YES	YES
INDUSTRY DUM.		YES	YES		YES	YES
Observations		10,203	2,021		5,339	566
Firms		4,405	779		3,063	346
Adjusted R-squared		0.10	0.13		0.32	0.52

**,* The t-statistic is significant at the 0.01, 0.05 level respectively. OLS standard errors corrected for clustering by firm.

7. Conclusion

We find confirmation that use of Schedule M-3 data for both publicly traded and privately held companies illuminates differences in the return on assets ratios estimated using book and tax data. We also qualitatively replicate MN's primary findings using M-3 data and extend the primary relationship beyond the Manufacturing CEP population to cover the entire Manufacturing sector of the LB&I population. We further extend this work by attempting to identify the contributions of identifiable temporary and permanent differences along with the contribution of the firms' financial structure and other characteristics in estimating otherwise unexplained book-tax differences. We think these three findings are important for better understanding book-tax differences.

First, the Schedule M-3 potentially provides enormous detail about private firms' book income, previously achievable only through surveys. Moreover, the designation of ownership can be more strictly applied, eliminating the difficulties of matching to public datasets. By combining Schedule M-3 data with a proven econometric model to confirm many well-developed theories in the book-tax literature, future researchers should be encouraged to use this dataset for more detailed research. However, as we discovered, Schedule M-3 reporting is only as good as taxpayer reporting and SOI cleaning. As researchers move away from totaled fields (particularly ones that should match the Form 1120 income statement), the data become sparse. In fact, out of the 24 lines on the Schedule M-3, Part II that are available for taxpayers to report book income, for every one dollar reported in itemized lines 1-24, ten dollars are reported in the catch-all "Other income items" field[17]. Based on descriptions supplied by taxpayers, many of the large items reported on this line match up to a line already on the Schedule M-3[18]. With the availability of M-3 electronic data for most LB&I firms, we believe much value could be added to this dataset if SOI could identify and allocate items shown in "other income" to more appropriate fields. SOI applies similar perfection to "other" fields found on income and deduction statements as well as balance sheets. Additional perfection of this sort should ensure future researchers have a richer dataset.

Second, we qualitatively replicate the earlier work by MN on the entire LB&I population, rather than only on CEP/CIC returns. It is encouraging that the broad conclusions of MN were found to exist in the more

diverse LB&I manufacturing population. We hope that as more years of data become available, research will be conducted on minor industries within this group, incorporating additional variables into the model that highlight unique reporting features.

Third, we provide evidence that even after controlling for temporary and permanent differences of key items as reported on Schedule M-3, the ownership type and capital structure finding are still robust. We also provide evidence that CIC income firms, as well as CIC firms that have consolidated subsidiaries, report higher BTD.

We've demonstrated that the current state of research on book-tax differences can be furthered by revisiting the research on publicly traded companies and by expanding the scope of the research to privately held LB&I business entities. Our study opens the door for future analysis of asset structures to reveal why BTD exists. We've discussed several reasons to use new data sources, in particular to draw out differences in public vs. private firm behaviors.

What else might matter? Other unexplored but compelling "predictors" of BTD are expected to further explain the reporting differences. We briefly discuss three areas we believe are ripe for further research. These areas cover temporary and permanent differences, as well as an attempt to increase our understanding of remaining and otherwise unexplained differences:

1. Intangibles/assets ratio. The literature suggests that rates of return to assets may not be meaningful indicators of what tax should fall out from income streams that are based in part on book value of intangible assets (Hulten and Hao, 2008). Among chemicals firms for example, as much as 10% of total assets are amortized intangibles. The taxable amount is difficult to discern from tax or book reporting, although Schedule M-3 data help. One possible way to get a better idea of the role that intellectual property plays in BTD is to examine patent assignor/assignee data and collaboration or marketing agreements between U.S. companies and foreign subsidiaries.

2. Foreign control of assets and extent of firm's foreign operations. Empirical evidence of income shifting by multinational companies provides impetus to explore these relationships further. Preliminary analysis of Physical Plant & Equipment (PP&E) return on assets by chemicals firms suggests that a considerable amount of tax revenue is at risk for underreporting by multinationals heavily engaged in moving profits offshore**.** Profit-maximizing companies "park" both real and intangible assets to take advantage of lower tax rates in other countries. The trend among large U.S. drug companies is rapid growth in profits earned offshore in relation to domestic profits due to transfer pricing **(**Martin Sullivan, 2008**).** This presents a compelling area for further analysis.

3. Aggressive tax behavior. Finally, BTD analysis begs the question of how far these differences are manifested in aggressive tax behavior. More needs to be done to get to the root cause of behaviors that lead to a wider corporate tax gap. Hanlon and Heitzman (2009) bemoan the "lack of good structural models of book-tax differences and effective tax rates" pointing out that studies are plagued by the perpetual chicken-and-egg story: do taxes paid impact decisions as first-order "drivers" of behavior, or are they simply the byproduct doomed to remain buried in the residual term of regression analyses?

Acknowledgments

The authors wish to acknowledge comments and assistance from Mark Batson, Charles Boynton, Thomas Burger, Charles Christian, Eurry Kim, Petro Lisowsky, Mark Lowe, Lillian Mills, George Plesko, Kenneth Szeflinski, and William Trautman. Any errors are solely the responsibility of the authors.

References

Altman, E, 1968, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," The Journal of Finance, Vol 23, No. 4: 589-609.

Batson, Mark, George Contos, John Guyton, Mark Lowe, Jean LaVelle, Deborah Myers, "Use of Assets in Large and Mid-Size Corporations: An Econometric Analysis of the Manufacturing Sector Including a Focus on the Chemical Industry," Working Paper.

Boynton, Charles, Portia DeFilippes, and Ellen Legel, 2005, "Prelude to Schedule M-3: Schedule M-1 Corporate Book-Tax Difference Data 1990-2003," Tax Notes, December 19: 841-870.

Boynton, Charles, Portia DeFilippes, and Ellen Legel, 2006, "2004 Schedule M-3 Reporting by Large Corporations," Tax Notes, September 11: 943-81.

Boynton, Charles, Portia DeFilippes, and Ellen Legel, 2008, "A First Look at 2005 Schedule M-3 Corporate Reporting," Tax Notes, November 3: 563-89.

Desai, 2003, "The Divergence Between Book and Tax Income," in James Poterba, ed., Tax Policy and the Economy 17, Cambridge: MIT Press.

Dyreng, Scott, Michelle Hanlon, and Edward Maydew, 2008, "Long-Run Corporate tax Avoidance," The Accounting Review, Vol 83: 61-82.

Graham, John, Jana Raedy, and Douglas Shackelford, 2010, "Research in Accounting for Income Taxes," NBER, Working Paper 15665.

Hanlon, Michelle, and Terry Shevlin., 2005, "Book-Tax Conformity for Corporate Income: An Introduction to the Issues," NBER, Working Paper 11067.

Hanlon, Michelle, and Shane Heitzman, 2009, "Review of Tax Research," University of Rochester – Bradley Policy Research Center, Working Paper FR09-23.

Hulten Charles, and Janet Hao, 2008, "Intangible Capital and the Mark to Book Value Puzzle," Conference Board Economics Program, Working Paper 08-02.

Jackson, Mark, 2009, "Book-Tax Differences and Earnings Growth," Doctoral Dissertation, University of Oregon.

Lisowsky, Petro, and William Trautman, preliminary 2007, "Book-Tax Consolidation Differences, Rates of Return, and Capital Structure," Working Paper.

Manzon, George B, and George Plesko, 2002, "The Relation between Financial and tax reporting Measures of Incomes," Tax Law Review, Vol. 55: 175-214.

Mills, Lillian, and K. Newberry, 2001, "The Influence of Tax and Nontax Costs on Book-Tax Reporting Differences: Public and Private Firms," Journal of the American Taxation Association (Spring): 1-19.

Mills, Lillian F., Kaye J. Newberry, and William B. Trautman, 2002, "Trends in Book-Tax Income and Balance Sheet Differences," Tax Notes, August 19: 1109-1124.

Mills, Lillian F., and Kaye J. Newberry, 2005, "Firms' Off-Balance Sheet Financing: Evidence from their Book-Tax Reporting Differences," Journal of Accounting Research, Vol. 43:251-282.

Mills, Lillian F., and George Plesko, 2003, "Bridging the Reporting Gap: A Proposal for More Informative Reconciling of Book and tax Income," National Tax Journal, Vol. 56, No. 4: 865-893.

Ohlson, J. 1980, "Financial Ratios and the Probabilistic Prediction of Bankruptcy". Journal of Accounting Research, Vol. 18 (Spring, 1980): 109–131.

Plesko, George, A., 2000, "Book-Tax Differences and the Measurement of Corporate Income," Proceedings of the Ninety-Second Annual Conference," Washington D.C.: National Tax Association – Tax Institute of America: 171 -176.

Plesko, George, A., 2004, "Reconciling Corporate Book and Tax Net Income, Tax Years 1995-2001," IRS, Statistics of Income Winter 2004-2005 Bulletin, Publication 1136:103-108.

Plesko, George, A., 2006, "Estimates of the Magnitude of Financial and Tax Reporting Conflicts," University of Connecticut Working Paper.

Poterba, James, 2007, "The Significance and Composition of Deferred Tax Assets and Liabilities," MIT, NBER. SSRN# 962750.

Sullivan, Martin, A., 2008, "Why Reported Effective Tax Rates Are Falling," Tax Notes, March 3, 2008.

U.S. General Accountability Office Report to Congressional Requesters, 2008, "Tax Administration, Comparison of the Reported Tax Liabilities of Foreign- and U.S.- Controlled Corporations, 1998-2005," GAO-08-957.

U.S. Government Accountability Office, 2008, "Comparison of the Reported Tax Liabilities of Foreign- and U.S.-Controlled Corporations, 1998-2005," GAO-08-957.

U.S. Department of the Treasury, 1999, "The Problem of Corporate Tax Shelters: Discussion, Analysis, and Legislative Proposals," Washington, D.C.: U.S. Government Printing Office.

Weiner, Joann M., 2007, "Closing the Other Tax Gap: The Book-Tax Income Gap," Tax Notes (May 28): 849-861.

Endnotes

[1] US Business Cycle Expansions and Contractions. www.nber.org/cycles.html

[2] Hanlon and Shevlin (2002) and Plesko (2004) reported a negative book-tax gap in 2001 but the gap turned positive again in 2002 and reached new heights in 2003.

[3] As of December 31, 2004 FASB 123R no longer allows this practice.

[4] This included 16 observations with debt to asset ratios greater than 6 and 131 observations with BTD greater than end-of-year total assets

[5] Additionally, the Schedule M-3 specifically asks about whether the firm is traded on an exchange or filed a 10-K. Previous studies relied on matching firms by EIN to public datasets as the way to determine ownership. It is quite common to have many mismatches, classifying public firms as private. Schedule M-3 data should allow for a more accurate determination of ownership type. There are, of course, reporting issues by taxpayers on these fields. For example, in the four-year period we examined, about 4% of firms "changed" ownership status at least once. Much of this could be real, but some could also be artifacts in the data (e.g., taxpayers simply not checking the 10-K box or listing a CUSIP). In this exercise we coded firms as public in all four years if they declared public in one. We also reran the data, coding firms as public only in the year for which we had information. The results were nearly unchanged.

[6] Pretax book income is the net book income for tax entities (filed on Schedule M-3 Part I, Line 11) plus U.S. current income tax expense (filed on Schedule M-3 Part III, line 1).

[7] Net Income from Form 1120, line 28 is reduced by special deductions. Special deductions eliminate the total dividends reported on Schedule C, line 19, and have been brought forward to page 1, line 4, so they are included in gross income and tax net income. Since taxable income cannot be negative and since tax net income does not include special deductions we elected to use as the main tax income variable the difference between tax net income and special deductions.

[8] End-of-year total assets is the only measure of assets reported on corporate tax returns. Total assets are reported on Form 1120 Schedule L balance sheet, line 15(d). Per taxpayer instructions all Schedule L items should be calculated under GAAP and they should follow the tax rules for consolidation of subsidiaries. Items reported on Schedule L, Income Statement, and Schedule M-3 should reconcile.

[9] Following the convention of Boynton, DeFilippes, and Legel (2008), a firm is labeled public if on the M-3 they either indicate they are publicly traded or filed a 10-K.

[10] Ohlson (1980) uses a logistic function to model the probability of bankruptcy, using several income and asset variables from current and prior years. Many users of this model apply the original coefficients derived from the work (using a 1970s dataset), to updated variables. Another standard bankruptcy model calculates a Z-score, based on work by Edward Altman in 1968. Hillegeist (2004) describes many of the concerns with using the decades-old estimates from these models. For consistency of replication, we used followed Mills' and Newberry's use of Ohlson's model, but future work should find a better predictor.

[11] We also tested an alternative specification for BTD, where BTD is equal to the sum of temporary and permanent difference reconciliation totals reported on Schedule M-3 lines 30 (b) and (c) plus U.S. current income tax expenses (reported on Schedule M-3 Part III, line 1 (b) and (c)) scaled by total assets. In addition, we separated BTD into temporary and permanent, again using Schedule M-3 lines 30(b) and (c), scaled by total assets and used as our dependent variable.

[12] A deferred tax liability is a tax obligation that will be paid in future taxes, while a deferred tax asset is a tax benefit that will occur in future years.

[13] For this variable we take advantage of information provided on Schedule M-3.

[14] As mentioned earlier, the CEP program has been replaced by the CIC program, which includes firms from a greater range of size. That, in addition to having only four years of Schedule M-3 data available, leads us to expand our dataset to Manufacturing sector firms with end-of-year assets of $500 million or more. We believe that this group of firms is the closest approximation to the CEP firm sample used by MN, for the 2005 to 2008 time period.

[15] The findings remain robust even when we use the alternative definition of BTD. When we separate BTD into temporary and permanent, the findings are much weaker, particularly for the temporary BTD. We show the output from the latter regressions in Appendix A.

[16] Note that this excludes, of course, amounts associated with cost of goods sold.

[17] Boynton, DeFilippes, and Legel (2008) investigate and allocate some of the largest amounts found in this line. This process, however, is tedious and dependent on each researcher creating a dictionary of terms to search for and move. They also note several large, recurring items in the "other" field that indicate the need for new M-3 lines.

[18] Graham et al. in their 2010 paper that reviews the research on accounting for income taxes, which includes the MN paper, show that if researchers rely on ordinary least squares (OLS) and White standard errors, they risk underestimating standard errors and overestimating the statistical significance of the coefficients. The reason is the OLS assumption of identically and independently distributed errors, which is violated when residuals are correlated through time or across firms. In our research, we adjust for this correlation by clustering standard errors.

Appendix A:
Regression Output for Firms with Assets of $500 Million or More

	Income Firms			Loss Firms		
Variable	Predicted Sign	Temporary Coefficient (T-statistic)	Permanent Coefficient (T-statistic)	Predicted Sign	Temporary Coefficient (T-statistic)	Permanent Coefficient (T-statistic)
Ownership type:						
PUBLIC	+	0.003 .(0.62)	0.018** (3.27)	-	0.035* (1.84)	0.002 (0.11)
Debt constraints:						
DEBT	+	0.010 (0.91)	0.033* (2.09)	?	0.052 (1.32)	-0.019 (-1.06)
DEBT*PUBLIC	-	-0.013 (-0.95)	-0.046** (2.75)	?	-0.062* (-1.53)	-0.019 (-0.92)
DEBT*DISTRESS	+	0.067* (2.34)	-0.004 (-0.13)	?	-0.162 (-1.32)	0.182** (3.44)
Control variables						
FTC	-	0.052 (0.35)	-1.596** (-5.86)		NA	NA
CAPITAL INTENSITY	?	-0.005 (-0.34)	-0.008 (-0.66)	?	-0.061 (-1.08)	0.346* (2.07)
DISTRESS	?	-0.034 (-1.27)	-0.034* (2.000)	?	-0.84* (1.78)	-0.247** (-3.51)
CIC		0.005 (1.19)	0.005 (1.49)		-0.002 (-0.15)	-0.001 (-0.11)
CONSOLIDATED		-0.004 (-1.35)	0.003 (0.54)		-0.14 (0.95)	-0.003 (-0.23)
CIC*CONSOL.		0.005 (0.15)	-0.001 (-0.22)		-0.004 (-0.19)	-0.0006 (0.10)
TAX YEAR DUM.		YES	YES		YES	YES
INDUSTRY DUM.		YES	YES		YES	YES
Observations		2,022	2,022		567	567
Firms		779	779		347	374
Adjusted R-squared		0.04	0.13		0.23	0.21

**,* The t-statistic is significant at the 0.01, 0.05 level respectively. OLS standard errors corrected for clustering by firm.

Sunshine and Shadows on Charity Governance: Public Disclosure as a Regulatory Tool

Evelyn Brody, Chicago-Kent College of Law, Illinois Institute of Technology [1]

The federal tax treatment of tax-exempt organizations, as I once wrote, is a photo-negative of Congress's treatment of those who pay taxes.[2] As a substantive matter, this favors charities: The higher the tax rate on for-profit corporations, the higher the relative value of the charity's income-tax exemption; the higher the individual income-tax rates, the lower the price of charitable giving (and the greater the interest-rate savings to charities from issuing tax-exempt bonds); and charities are big defenders of the estate tax, under which donations are fully deductible. By contrast, exempt organizations might lament the reversal of the presumption of privacy: In stark contrast to the strict protections enjoyed by taxpayers under Internal Revenue Code section 6103, exempt-organization filings are publicly available under Code section 6104.

Indeed, in the last twenty years, the annual information return filed by federally tax-exempt organizations—the Form 990—has become not only the public face of individual charities, but also the most readily available data source for potential donors, state regulators, the media, and researchers, as well as the charity's governing board, staff, and volunteers. So important is this filing that in 2008 the Internal Revenue Service took into account the interests of these various stakeholders in radically redesigning the Form 990, which now makes available, among other information, a detailed picture of the organization's governing structure, policies, and related-party transactions. Moreover, simply by asking questions about the existence of perceived "best practices," the Service sends a strong signal of their desirability. Meanwhile, the emergence of the third-party online database of Forms 990 maintained by GuideStar—itself a private, nonprofit organization—completes the goal of transparency: Aside from any oversight actions of regulators, any member of the public (including competitors, nonprofit and for-profit) can scrutinize filings without the charity's knowledge of who is looking, when, or why.

Sunlight, of course, creates both clarity and shadows. Knowing that detailed information about charity structure and practices will be available to the public can—as no doubt intended—influence charity behavior. However, requiring charities to disclose information to the IRS is a separate question from requiring charities to disclose their IRS filings to the public. Since 1987 exempt organizations have operated under a statutory obligation to provide their Forms 990 upon demand.[3] That significant development has long made me wonder about the effect on the nonprofit sector from mandated public disclosure of tax filings.[4] In a March 2010 letter, then-ranking member (and former chair) of the Senate Finance Committee, Charles Grassley, praised the 2008 redesign of the Form 990 in declaring: "The best way I know to increase voluntary compliance is to inject transparency."

Meanwhile, pursuant of its obligation to administer and enforce the requirements for federal tax exemption, the IRS has long kept its hand in issues of sound charity governance. In the 1990's, the public was treated to peeks at the IRS's view of appropriate governance through the release of a few otherwise confidential "closing agreements" that the IRS entered into on the condition that the organization agree to publish them.[5] More systematically, in 1996, Congress involved the IRS in charity governance by adopting the "intermediate sanctions" statute designed to deter charity insiders from engaging in "excess benefit transactions" with charities.[6] The legislative history to Internal Revenue Code section 4958 suggested that administrative guidance could protect financial transactions entered into between charities and their insiders if the approval process assured independent decision-making, obtained comparable data, and maintained documentation. Treasury Regulations issued under section 4958 detail the process for qualifying for such a "rebuttable presumption of reasonableness."[7]

More recently, charity governance writ broadly has emerged as a fundamental focus in the regulation of federally tax-exempt organizations. In 2004, the staff of the Senate Finance Committee produced a white paper proposing a broader role for the IRS in charity governance; the nonprofit sector responded with studies and proposals to improve nonprofit governance, including recommendations for self-regulation. Both when chair and subsequently as ranking member of the Senate Finance Committee, Grassley demanded and posted online massive amounts of information (including emails and correspondence, some labeled "privileged and confidential") from specific organizations whose governance practices he questioned.[8]

Relying on public disclosure, however, puts pressure on the IRS to ensure that the form asks the "right" questions, and allows the filer to present complete and accurate answers. The IRS itself benefited from a transparent process in its Form 990 redesign, having posted online drafts of the form (and schedules) and the thousands of comments it received, all still available on the IRS website.[9] That exposure process allowed the IRS not just to rework misleading questions but also recast the questions both to produce a better picture of the organization and to steer the sector to good governance structures and practices. Notably, the IRS acceded to a storm of pleas to remove the most "prejudicial" (and uninformative) lines from the all-important new summary page (Part I). (Compare the 2007 draft, on the first page of the Appendix, with the final version, on the third page.) Line 6 of draft Part I had asked: "Enter the number of individuals receiving compensation in excess of $100,000 (Part II, line 2)"; while this line, like all the others in the summary page, draws from a question elsewhere on the form, what valid information does it convey by including it on the front of the form? Similarly, the IRS removed the three "efficiency ratio" questions, which, while used by some charity watchdog groups and rating agencies, have long been criticized as oversimplified and unhelpful metrics.

To give another example, consider the 2007 draft Form 990's question in Part III (Statements Regarding Governance, Management, and Financial Reporting) on conflict of interest transactions (see the second page of the Appendix):

3a Does the organization have a written conflict of interest policy?

b If "Yes," how many transactions did the organization review under this policy and related procedures during the year?

What is the preferred answer to question 3b? If the organization answers "zero," is this good (because there were no conflict of interest transactions to review) or bad (because the organization was blind to the interested transactions that occurred)? Commentators pointed out the problems with this and other governance questions of the draft. Substantially revised (and renumbered) Part VI not only addresses the suggestions (see the last page of the Appendix), but also states at the outset: "Governance, Management and Disclosure (Sections A, B, and C request information about policies not required by the Internal Revenue Code.)" (On the 2010 version of the Form 990, this disclaimer has been moved to the beginning of Part B (Policies).). Moreover, the 2007 draft did not provide an opportunity for the organization to provide attachments to the form. In response to complaints—including the argument that it is unconstitutional to deny a filer subject to mandatory disclosure the opportunity to explain yes/no and other short answers—the final form includes a Schedule O for extensions of responses and supplemental narration. The final conflict of interest questions read:

12a Does the organization have a written conflict of interest policy? *If "No," go to line 13.*

b Are officers, directors or trustees, and key employees required to disclose annually interests that could give rise to conflicts?

c Does the organization regularly and consistently monitor and enforce compliance with the policy? *If "Yes," describe in Schedule O how this is done.*

(Incidentally, contrary to the suggestion in line 12a, an organization could have a conflicts-of-interest policy, and engage in effective monitoring, without reducing the policy to writing.)

Separately, since 2003, the Internal Revenue Service has become subject to public disclosure obligations of its own. As a result of Freedom of Information Act suits brought against the IRS, the agency's views on a range of issues can, at least informally, be gleaned through the release of rulings denying or revoking exemption. (By law, the IRS redacts these rulings to hide the names of and other identifying information about the charities

and other taxpayers.) Most helpful for the nonprofit sector would be for the Service to take the now-substantial database of denial and revocation letters and develop from it formal guidance on which indicators of governance structure and policies the IRS would like to impose as conditions for exemption.

The simultaneous developments of substance and process—of increased federal interest in charity governance and in the tool of disclosure—threaten to conflate an examination of the relative merits of each. It might be appropriate, for example, to require reporting of certain information to state regulators or to the Internal Revenue Service without also requiring that information to be made publicly available. It is important to recognize, however, that a large percentage of exempt organizations file forms other than the Form 990, or do not even file a substantive return. Because of statutory exemptions, it can be difficult, if not impossible, to obtain much information on churches and on smaller charities. Moreover, hundreds of thousands of charities will fall below the governance radar when the cutoff between charities required to file the Form 990 and the simplified Form 990-EZ is fully phased in beginning in 2010: The definition of "small" doubled from $25,000 or less in gross receipts to $50,000 or less. Finally, arguably the most important disclosures take place internally within the organization.

To explore these thoughts, Part I begins with the desirability of information flow to key decision makers in the organization, including the board, and considers the possible dilatory as well as salutary effects of public disclosure on governance practices. Part II covers reporting to the states and to the IRS as regulator of the federal tax-exemption regime. Part III, the longest of this essay, compares disclosure of filings with the regulators (the charities' transparency) and disclosure of enforcement activities (the regulators' transparency). Part IV looks at voluntary public disclosure by the organization, as well as disclosure by third parties, notably the media and third-party "watchdogs." Throughout, I not only describe criticisms of required disclosure, but also suggest areas in which the current levels or types of disclosure are not enough.

And now the dark side of sunshine. The largest practical impediment to relying on public disclosure is the unfortunately widespread assumption that providing charity is a free good—and so general overhead, much less fund raising expenses, should be zero or close to it.[10] One of the great lost opportunities of the September 11th experience was the failure of charities to defend the costs of wisely allocating charitable resources. More broadly, charities resist increased standardized disclosures because they worry that the public will misunderstand or misinterpret the information. A public that does not understand cost constraints cannot perform effective oversight. A public whose oversight focuses on the wrong considerations induces charities to adopt inefficient and ineffective behaviors.

In this climate, the solution to the problem of a misinformed public is more disclosure—nothing prevents an organization from providing a more positive narrative of its goals and accomplishments. While the competing demands of the various stakeholders cannot always be reconciled, all involved will better appreciate the challenges faced by a charity that reveals rather than hides its costs of fund raising and administration; explains why its executives merit their pay and why its reserves are necessary; and describes its limits as well as its potential in delivering services and addressing social needs. Finally, the sector as a whole should also weigh in, denouncing unacceptable practices.

Compare a recent U.K. report addressing whether public confidence in charities would be affected by increased mandatory disclosure of expense reimbursements. The report opposed expanding mandated disclosure beyond current requirements, arguing, in part: "Greater disclosure might risk being at best, of little interest or, at worst, of misinterpretation and even suspicion, possibly leading to damage to public trust and confidence. This might risk elevating expenses to become an inappropriate measure of charity effectiveness and distract attention away from more appropriate measures, namely those relating to a charity's overall outcomes and impact. It might even lead to pressure to inappropriately drive down certain costs." Moreover, the report continued, focusing only on expenses ignores issues of greater accountability for "good governance and sound systems of internal control.[11] Rather, the lengthy report—which was based partially on a survey to which 575 registered charities responded—urged trustees to consider additional, appropriate voluntary public disclosure, in addition to ensuring the adoption, internal communication of, and compliance with an expense reimbursement policy.[12]

Indeed, the voluntary disclosure of information also serves charities that do not solicit donations. All nonprofits remain politically vulnerable—not just to the removal of subsidies, but also to the danger of unwise legislation and regulation.[13]

Regrettably, the most important information that both regulators and the public might want will continue to be unavailable—simply because performance measurement is an unsolved metric. As a society, we would want to be able to assess whether and which charities are producing favorable outcomes, but often we cannot even measure outputs because quality can be subjective. At the same time, while focusing on outputs (such as patient stays or unemployed trained) can lead to de facto quotas, focusing on outcomes (such as good health or jobs) holds nonprofits responsible for factors beyond their control. Thus, beyond the scope of this paper is the ultimate disclosure question: How do we challenge an organization that says it "does good"?

I. Sound Governance and Internal Disclosure

As described in the American Law Institute's project on Principles of the Law of Nonprofit Organizations, for which I am the reporter, a charity's governing board is responsible for "establishing appropriate procedures for internal controls, including financial controls, legal compliance, and information flow to the board."[14] Thus, as a matter of good governance, the board needs accurate and timely information from management, including financial reports, and accurate and timely information from board members themselves, such as when a transaction might present a conflict of interest for a particular fiduciary.

I am often asked if there are any limits on the information to which a board member is entitled, and the answer is almost always no.[15] More fundamentally, it seems that we cannot be too basic in explaining to nonprofit board members what information they should be seeking. A comment in the ALI project sets forth the documents which should be provided to every board member.[16] The availability of Forms 990 from GuideStar's website, of course, means that board members—and prospective board members—can learn a great deal about the organization even if management is not forthcoming

After all, organization formation and operations generally are private affairs. If the organization is itself a quasi-public entity, it might be subject to sunshine laws. Tax-exemption alone, however, does not convert a nonprofit organization into a public entity.[17] (Separately, the government as grant-maker might impose transparency as a condition of funding; state laws vary.) Nevertheless, a great deal of internal information becomes public information because it must be set forth on regulatory reports, as explained in Part III.

Nonprofit governance practices have long remained a mystery. In 2007, the first comprehensive survey was published, by the Urban Institute's Francie Ostrower.[18] Notably, she found that charities commonly enter into transactions for goods and service (beyond board services) with members of the governing body, and that these transactions grew with charity size; but she further found that it was not even always known to a particular organization whether a fiduciary was on the other side of a transaction.[19] She further found a serious lack of internal disclosure: "among those nonprofits that say they did not engage in transactions with board members or affiliated companies, however, fully 75 percent also say they do not require board members to disclose their financial interests in entities doing business with the organization, and thus, respondents may have been unaware of transactions that do exist."[20]

How has—and will be—nonprofit governance affected by the knowledge that internal information is public due to its presence on the Form 990 and other filings? (This topic is explored at length in Part III; specifically, see Part II for more discussion of the governance questions in the redesigned Form 990.) Will organizations change their decisions or pay more attention to documenting their decisions, providing additional explanation? Will organizations try harder to skew the information to what it perceives the public wants to see? There is a difference between perceived wrongdoing and actual wrongdoing. If the public misinterprets or demands the wrong "answers," charities can suffer a loss of trust.

To give a personal example, early in the Internet age, as I was about to write our family's charitable contribution checks, I realized I could and should consult the organizations' Forms 990 from my home computer. Back then—and sadly, still too often today—you couldn't expect to find this information on most charities' own websites, but rather you would have to sneak, feeling somewhat guilty, to GuideStar. There I discovered

that two organizations to which we had generously contributed reported high executive compensation (i.e., more than four times what I make) and high retained surpluses (i.e., an amount that overwhelms my intended contribution). Then I tried to get a grip on myself: "Hold on," I muttered. "You're a professional! Surely you appreciate why these important, well-run organizations need to pay the executive salaries and maintain the reserves they do." But if that was the reaction of "a professional," it's easy to see why charities are loath to report to the public at large.

Even before the 2008 redesign of the Form 990, advisors focused on the importance of having the board know what will appear in the organization's federal tax filing.[21] Attention to executive compensation, interested transactions, and relationships among fiduciaries will be even more important as exempt organizations file the redesigned Form 990. As described in Part II, the new version of the form contains numerous questions about organizational structure and governance practices. (See also the last page of the Appendix, which reproduces the governance part of the form.) Despite the disclaimer, described above, that this portion of the form "requests information about policies not required by the Internal Revenue Code,"[22] the expectation is that most organizations will want to answer "yes" to the questions. It will be interesting to see, as the next few years pass, the rise in adoption of the policies and practices asked about on the return.

More basically, if board members have not routinely been provided with the organization's Forms 990, they likely will now: One question reads: "Was a copy of the Form 990 provided to the organization's governing body before it was filed? All organizations must describe in Schedule O the process, if any, the organization uses to review the Form 990."[23] Not only will the typical board's role in preparing or reviewing the 990 change, but also the relationship between the board and management could change as the board focuses on reported structures and events as it might not have in the past.

At the extreme, a nonprofit might even be willing to forgo tax-exempt status in part to preserve the confidentiality of its activities,[24] given that corporate income tax returns are not subject to public disclosure.[25] More likely, a nonprofit might use a for-profit affiliate to carry out charitable activities for which tax exemption would be available,[26] especially when taxable profits are expected to be nonexistent or low. While an organization might sacrifice some support (from employees, donors, or others) in forgoing exemption, other advantages of the for-profit form include the ability of raising equity capital; avoiding an IRS inquiry into whether the nonprofit has sufficient charitable purposes; and gaining some flexibility in providing levels and types of compensation.[27]

II. Regulatory Registration and Reporting

This Part looks at filings received by nonprofit regulators. Part III's discussion of public disclosure includes the transparency of enforcement actions by the regulators.

A. State Registries; Constitutional Limits on State Regulation of Fundraising

A nonprofit corporation typically obtains its certificate of incorporation from the state secretary of state, and makes annual filings with that office. Outside the well-regulated area of charitable solicitation, described below, Marion Fremont-Smith's comprehensive survey chronicles the development—but lamentably limited extent—of attorney general registration and annual filing (7 states).[28] (Fremont-Smith separately found that in four states the attorney general must be notified when the nonprofit seeks tax-exemption.[29]) In 2011, the Uniform Law Commission adopted a "Model Protection of Charitable Assets Act"; the project addresses the authority of state attorneys general to protect charitable assets, to require annual filing and notice of specified "life-events," and to cooperate in interstate and multi-state cases and with the IRS.[30]

Most state oversight of charity deals with the solicitation of contributions. In the 1960s and 1970s, the desire to protect charities from "wasting" resources on fundraising led a total of 28 States and countless municipalities to impose ceilings on the percentage of annual revenues that could be spent on fundraising expenses. In the 1980s, however, a trio of Supreme Court decisions blocked these restrictions, on First Amendment free-speech grounds.[31] To the Court, Procrustean percentage limits on fundraising disproportionately impact new charities (with low name recognition and no established donor base) and unpopular causes (which require a

greater expenditure to raise a dollar). States may punish fraudulent fundraising speech after-the-fact, but, as the Court more recently confirmed, regulatory approaches seeking to equate fraud with fundraising efficiency are invalid.[32]

Conceding their inability to mandate fundraising limits, the states have concentrated their efforts on requiring charities to increase public disclosure using standardized forms. Almost all the states require registration; a charity soliciting in many states will welcome the Uniform Registration Statement accepted in most states requiring registration.[33] In addition, 35 states require annual filings, usually with the attorney general, for charitable trusts and nonprofit corporations that solicit charitable contributions; those states either require or accept the Form 990 in partial or complete satisfaction of that filing. Statutes, though, commonly exempt small entities, educational institutions, hospitals, and churches—and membership organizations—but variations abound. Some localities also regulate fundraising.

B. Federal Tax Filings: Governance Focus of Redesigned Form 990

Because of legislation enacted in 2006, the IRS will be able to clean up its Business Master File to weed out those nonfiling small charities that have simply ceased to exist: Effective for tax years beginning in 2007 small organizations that fail to file an annual notice of their continued existence (and minimal other information) for three consecutive years will have their exemption revoked.[34] As of 2009, the IRS records showed a total of 1,912,695 exempt organizations (1,238,201 million of which were exempt under section 501(c)(3)). As of mid-2011, the IRS had announced that the net total of automatic revocations exceeded 330,000.[35] To ascertain whether these organizations were "in fact defunct or just uninformed and/or confused about IRS regulations,"[36] researchers who had previously reached out to vulnerable Indiana organizations concluded that 27 percent of organizations "that we have reason to believe are still active" lost their exemption for failure to file.[37]

With the overhaul of the Form 990 effective for tax years beginning in 2008, we will finally have up-to-date information about organizational form for most large public charities.[38] Line K near the beginning of form asks the filer to identify the type of organization, with boxes provided for corporation, trust, association, and other (with space to describe). In a comment letter on the 2007 draft of the redesigned form, I suggested adding such a question.[39]

Not surprisingly, the Form 990 focuses largely on financial reporting and transactions—the Internal Revenue Service's core competency is, after all, tax collection, which is measured in dollars. The Form 990 is not limited to financial results, though, because it also has to reflect specific requirements and prohibitions in the tax laws. Thus we find many questions about relationships among fiduciaries and conflict-of-interest transactions, as well as questions about two additional concerns of federal tax exemption for charities: unrelated business activity and lobbying and political activity.

The most striking feature of the 2008 redesigned Form 990 is the new first page that highlights key information set forth elsewhere on the form. This summary page will make the form more accessible to donors, the press, and state regulators—not to mention to board members themselves. The form also adds a full page of questions about organizational structure and governance practices.[40] (See the Appendix for the 2007 exposure draft and the 2008 final versions of those two pages.) I strongly supported this focus on governance in my comment letter on the draft redesign. Indeed, I proposed replacing the draft half-page of questions with a full page of my own. As I explained: "It seems to me that most useful for the Service, potential donors, the press, and anyone else who reviews the Form 990 would be a series of questions that describe the governance structure of the organization and that determine whether the organization has in place procedures to support good governance." I added: "At the same time, it is important to recognize that these organizations are private entities, whose obligation to make public disclosures must be based on the requirements of the Code. I agree with those who have urged you make clear—on the Form itself and not just in the instructions—which of these items are legally required, so that readers do not draw inappropriate adverse inferences."[41]

Tracking many of my suggestions, Part VI as finalized requires the disclosure of whether the organization has a voting membership; the identity of voting board members (and which ones are independent); whether and how certain documents, including the organization's Form 1023, Forms 990, and 990-T, financial statements, governing documents, and conflict of interest policies, are made available to the general public; and

whether the organization became aware during the year of an embezzlement or other material diversion of the organization's assets.

But the governance-focused part of the Form 990, which Steve Miller, then-Commissioner for the Tax Exempt and Government Entities Division (TE/GE), characterized as "the crown jewel" of the IRS's recent activity in the nonprofit governance area, has proven somewhat controversial. The Advisory Committee on Tax Exempt and Government Entities (the "ACT"), a high-level advisory body to TE/GE, issued a lengthy report in June 2008 focused on the IRS role in charity governance. The 2008 ACT report comments:

> We believe in large part the governance questions on the redesigned Form 990 for 2008 are appropriate and formulated in a relatively neutral manner, recognizing that true neutrality is an unattainable goal. The inclusion of the questions, however, inherently (and intentionally) suggests that the IRS supports adoption of specific governance policies and practices. The danger then is that organizations will take the path of least resistance and adopt the policies and practices whether or not they are appropriate for them, or effective in their context.[42]

The ACT concludes that the public availability of the Form 990 will induce organizations to adopt practices that they might not need, as discussed in Part I, above: "Thus, while disclosure and transparency play a valid role in promoting compliance with the tax laws and in encouraging appropriate nonprofit governance, they also can impact behavior in a manner that can be harmful to the sector, and inappropriately suggest to the public and watchdog groups that the absence of specific governance policies or practices is in effect misgovernance. Accordingly, the IRS should carefully consider the public disclosures it requires."[43]

C. What's Not Publically Available from Federal Tax Filings

As thorough as the redesigned Form 990 appears, we still have reporting holes.

1. Filing Exceptions

Separate from the filing exemption for churches, as mentioned above, the IRS phased in the requirement to use the new Form 990 or the simpler Form 990-EZ by the size of the organization.[44] The Service doubled the annual revenue threshold for filing Form 990 or Form 990-EZ from $25,000 to $50,000 beginning in 2010.[45] Thus many "small" organizations will shift to filing either the short Form 990-EZ or the bare-bones e-postcard, Form 990-N (which requires only such basic information as employee identification number, the name of a principal officer, a mailing address, and affirmation that gross receipts total less than the threshold). Although I am sympathetic to saving costs for small organizations as well as the IRS, both regulators and the public stand to lose valuable information on hundreds of thousands of small organizations. This latter issue is of particular concern to state regulators that accept the series Form 990 as its annual filing document.

2. Data on Form 990 That Are Unclear or Not Collected

Some of the ambiguities on the prior Form 990 will be cleared up by the redesigned Form 990. Consider the fundamental example of determining who is in charge of the organization—particularly who actually has power in those arts and cultural or educational institutions with multiple advisory positions (the proliferation of titles, like "life trustee," are uniformative). While the draft redesigned Form 990 asked simply for a listing of trustees or directors, the final form makes clear that it is looking for those with voting rights only.

As another example, my comment letter to the IRS noted the tendency of too many expenses winding up on the "other" line, which allows for the itemization of specific categories not listed above. In the redesigned Form 990, Line 24 of Part IX (Statement of Functional Expenses) of the Core Form cautions: "Expenses grouped together and labeled miscellaneous may not exceed 5% of total expenses . . . ."

Problems of inaccurate or incomplete filings will continue. The push to electronic filing will help with the latter problem if the system will not accept a return unless the fields are properly filled in. As to the former problem, Floyd Perkins, former Illinois charities bureau chief, commented, "People don't realize how poor the quality is."[46] (See Part I, above, for a discussion of the pressures to fudge numbers. Perkins added, though, that there are "not a lot of examples where people relied on phony reports.")

Is there a duty to amend a return discovered to contain a material misrepresentation? The tax system imposes no statutory duty to amend a tax returns, although filing an amended return stops the accumulation of penalties and interest (but for an exempt organization, interest on what?). By contrast, the federal securities laws require amendment of a filing if failure to amend would be materially misleading. The possibility of state-level enforcement of an inaccurate return, where the Form 990 satisfies the state filing requirement, can provide an incentive to file an amended Form 990 at both the federal and state levels.[47]

3. Group Returns

The tax rules provide not only for umbrella recognition of multiple related exempt organizations, but also permit the filing of group returns. By contrast, the IRS does not permit members of an affiliated group to file a consolidated return, as that term is understood in corporate tax. Group returns thus can be uniquely uninformative and nontransparent: The return includes all members of the group except the "parent," in contrast to a corporate consolidated return (and any member of the group can elect to file its own return); the transactions within the group are not netted, as they would be in a corporate consolidated return; and it is impossible to determine the finances and operations of any particular member of the group. The topic was the subject of the 2011 IRS ACT report, which urged the IRS to strengthen the group exemption requirements but disallow the filing of group returns.[48]

III. Public Disclosure of Regulatory Filings and Determinations

The discussion in this Part III examines the privacy interests of charities and relevant third parties; reviews what types of state and federal filings are made public; analyzes the possible rationales for public disclosure; and addresses the transparency (or not) of charity regulators.

In the federal tax system as a whole, Congress' overarching lodestar with regard to tax return information is confidentiality. While individuals and businesses are compelled to report their activities to the IRS, the IRS may not release taxpayer identifying information to the public—or even, except as specifically permitted by statute, to other governmental agencies.[49] Indeed, a taxpayer may recover damages from the government for unauthorized disclosure, and severe penalties apply to IRS employees who improperly disclose return information. This presumption of confidentiality, however, is reversed for tax-exempt organizations.[50] Why does Congress only in the nonprofit view context view sunlight as an important disinfectant?

A. Privacy Interests of Charities and Their Supporters

By longstanding law and practice a charity's governance activities and operations are generally private affairs. Requiring regulatory filings and other information to be disclosed to the public intrudes even more than does reporting to regulators on the associational and operational autonomy of charities, and might even make board service or employment less attractive. (Thus the title for the talk from which this Article derives: "Governing in a Fishbowl.") Indeed, the most controversial portion of the IRS Form 990—and the primary reason for initial resistance by exempt organizations to requests for public disclosure—is the section reporting board member and executive compensation. (Often, the organization's own employees and volunteers are the most curious!) As discussed below, policy makers and observers have identified a variety of justifications for state and federally required public disclosures by charities, the levels and types of which seem only to increase. Importantly, the summary cover page of the redesigned Form 990 highlights certain information of particular importance to donors, the press, and state regulators—not to mention to the organization's board members.

Privacy interests are broader than the charity's, of course, and in certain situations public disclosure can lead to harm for the charity or to its donors, members, or those it serves. One category of sensitive information includes the types of trade secrets and personnel information protected from disclosure, as described below, by Freedom of Information laws. Narrower examples of sensitive information protected from disclosure include the address of a battered women's shelter (so that abusers cannot find clients) and the countries of operation of human rights organizations (note that Schedule F of the new Form 990 was revised to address this concern). Public disclosure of membership lists also can be sensitive, particularly for groups advocating on socially

contentious issues; usually, membership lists are not even required to be filed with regulators. Churches receive special protection by their exclusion from the requirement to file an application for recognition of federal tax exemption and Forms 990. (But see the lengthy discussion in a 2011 staff memorandum to Senator (and then-Ranking Member) of the Senate Finance Committee) Charles Grassley about the history and possible modification of this special treatment of churches.[51])

The identity of donors is an area of particular focus. The names of contributors to private foundation are not redacted from the Form 990-PF, which is required to be made publicly available in full. Donors to state-related nonprofit institutions, such as alumni-created foundations affiliated with state universities, are often unprotected as well.[52] By contrast, Congress exempts from public disclosure the names of donors reported on the list of major donors (Schedule B) to the Form 990 filed with the IRS by exempt organizations other than private foundations. As one result, only the IRS can fully review a charity's claim to be publicly supported, and thus not a private foundation.[53]

B. What Filings Are Subject to Public Disclosure?

The states typically make available—often online—corporate annual reports filed with the secretary of state, and annual reports filed with the attorney general in those states requiring reports, generally from those who solicit charitable contributions (see Part II, above). Confidential information can be protected from public disclosure. Uniquely, as far as I know, New Jersey requires that the audit submitted to the attorney general be accompanied by the auditor's management letter, if one was prepared, although the management letter will not be released to the general public.[54] Material supplied in the course of or subsequent to a state investigation remains confidential except as might be required under a state freedom of information law.

Specifically, Code sections 6104 and 6110 provide for disclosing applications for tax exemption, including supporting documents, and determination letters and rulings. All of these items are available from the IRS upon request. Moreover, the organization must make its exemption application, supporting documents, and determination letter or ruling available for public inspection without charge. Separately, the law obligates a charity to produce any of its last three tax returns upon request. Posting the Form 990 on the charity's website satisfies this obligation—but the posted return must be complete. Evidently, of greatest interest to the press, the public, competitors, and even other workers in the organization are the salaries and other compensation paid to the top executives and independent contractors, and a return provided without this information does not satisfy the disclosure obligation.

Even though the filings made with the IRS are available from the regulator (the same is true for some of the states), private groups revolutionized charity transparency. The searchable databases on GuideStar and the National Center for Charitable Statistics at the Urban Institute—themselves privately funded charities that work with each other and with the IRS—make this whole system work.[55] The IRS itself offers for sale (at no cost to the media and other government agencies) scanned copies of the last seven years of filed Forms 990 on DVD or CD-ROM.[56] It would be most helpful if the IRS provided usable data from these forms promptly to researchers.

A training program by the IRS Exempt Organizations Division explains some of the advantages of instantaneous, online disclosure: "Obtaining information from an organization had potential drawbacks if a requestor and the organization were not on friendly terms. Despite the requirements of the law, some organizations simply refused to allow access to their returns."[57] That article provides "a discussion of the more common errors made and an explanation of the reasons for some of the information requested." (Regrettably, in 2005, the EO division discontinued drafting these training materials, which has been a great loss to practitioners as well as to the Exempt Organization staff.[58])

Some information still remains private between the organization and the tax collector. The statute excludes from public disclosure the customary FOIA exceptions for "a trade secret, patent, process, style of work, or apparatus if the Service determines that the disclosure of the information would adversely affect the organization." In addition, as mentioned above, Schedule B to the Form 990, on which public charities report the identities of their large donors, is protected from mandatory disclosure.[59] Exemption applications are not public until exemption is granted; nor must withdrawn applications for exemption be disclosed.

Finally, the Pension Protection Act of 2006 requires an exempt organization to make public its Form 990-T, on which it reports and pays any tax due on unrelated business taxable income.[60] However, Congress did not impose a parallel requirement on the corporate returns of an exempt organization's taxable affiliates (business tax returns, like the returns of individuals, are not public documents), giving charities one more reason to spin off unrelated businesses into a separate for-profit corporation. Unfortunately, because of a glitch in the statute, the IRS cannot provide the Forms 990-T to GuideStar, so anyone curious about unrelated business activity of a particular charity will have to ask the organization for the form, and they will not be available in a searchable database of these forms.

1. Applications for Exemption—Form 1023

The application form used to file for recognition of federal tax exemption under section 501(c)(3) was significantly revised in 2004.[61] The 2008 ACT report on the IRS role in charity governance described the evolution of the IRS's approach to governance during the exemption application process: "While the Form 1023 prior to the current version asked questions regarding organization structure and governance, it principally focused on the charitable activities of the organization. In contrast, the 2004 (the most current) version places an increased emphasis on an organization's governance by focusing on board and management relationships (independence) as well as compensation and other potential opportunities for inurement."[62]

Commentator Jack Siegel praised the IRS for "attempting to identify those organizations that are likely to violate the rules governing Section 501(c)(3) organizations before granting tax-exempt status rather than relying on an audit process that is currently underfunded and spotty." However, Siegel cautioned future applicants who seek to abuse tax-exempt status to take care in filling out the application: "In the past, questions covering compensation, grant making, affiliations, and activities were very open-ended, permitting people who wanted to game the system to conveniently omit information without significant risk. The 2004 revised Form 1023 touches on all the same topics, but with very specific questions which will make it much more difficult to hide abusive arrangements without risking penalties of perjury.[63] Siegel added: "We also suspect that certain answers to questions may not cost an organization its requested exempt status, but may place the organization in a special queue for subsequent audits focused on potential violations under the intermediate sanctions."

Of course, failure to make full disclosure on the prior versions of the application form—which, like the Form 990, is filed under penalties of perjury—still had consequences. In an unusual case, the United States recently won a criminal conviction against a Muslim group that had failed to disclose on its Form 1023 what the Justice Department asserted were such terrorist activities as publishing newsletters and raising funds for jihad.[64]

2. Forms 990: Problems of Accuracy and Timeliness

Like other federal tax returns, the Forms 990 are self-reported. Many as filed contain errors, some materially misleading. Hopefully, compliance will improve as boards and top management become more involved in preparing the form. Even with the redesign, though, this document cannot provide much insight into the nature and quality of charity activities.

Moreover, many Forms 990 are filed under an automatic six-month extension. The blame for this commonly falls on the accountants, who can barely recover from having to prepare tax returns for individuals (due April 15) before gearing up to file Forms 990 (due May 15, for calendar-year organizations). No reputational sanction seems to follow from filing late, so many calendar-year exempt organizations file close to November 15. (You can set your calendar by all the news stories on nonprofit compensation that appear around Thanksgiving.) This means that events that occur in, say, January 2011 will likely not be disclosed to the public until November 2012, almost two years later.[65]

The IRS highlights the value of disclosure in describing its e-filing initiative: "E-filing reduces normal processing time and makes compliance with reporting and disclosure requirements easier."[66] Indeed, e-filing is mandatory for large charities: "For tax years ending on or after December 31, 2006, exempt organizations with $10 million or more in total assets may be required to e-file if the organization files at least 250 returns in

a calendar year, including income, excise, employment tax and information returns . . . . Private foundations and non-exempt charitable trusts are required to file Forms 990-PF electronically regardless of their asset size, if they file at least 250 returns annually."[67] Beginning in 2006, the Service started a federal/state filing system, and has begun working with individual states to test their systems.[68] In 2008, the Service processed 901,000 exempt-organization tax returns (mostly Forms 990, 990-EZ, and 990-PF); and in 2009 processed 1,132,000, an increase of 25.6% (presumably due to the filings of the e-Postcard, Form 990-N).[69] Of these returns, many were filed electronically: In 2008, exempt organizations filed 57,975 Forms 990; 44,362 Forms 990-EZ; and 292,002 Forms 990-N (which can only be filed electronically).[70]

C. Rationales for Governmentally Mandated Disclosure to the Public

This subpart considers four possible rationales for mandating public disclosure of charity finances and other activities.

1. Disclosure Without Judgment: "Disclose or Abstain"

While, as mentioned above, Congress provides for the confidentiality of tax returns, in regulating the securities issued by publicly traded companies, Congress has generally adopted a "disclose or abstain" model in lieu of prescriptive regulation. Under that approach, if the issuer makes honest (i.e., not materially misleading) public disclosures, we essentially leave investment decisions to the market. If a similar public disclosure rationale is chosen for charity regulation, what are nondisclosing nonprofits supposed to abstain from? Soliciting the public for contributions (state registration model)? Something else? After all, the typical private foundation or government-funded agency is not seeking or expecting contributions from the public. Interestingly, Congress required private foundations to make their Forms 990-PF available on request in 1969, but did not obligate publicly supported charities to make their Forms 990 available until 1987.

Incidentally, a disclosure model based on this rationale might be the only constitutional regulation permitted of corporate political speech after the Supreme Court's decision in Citizens United, a topic beyond the scope of this article.[71]

2. Condition of Tax Subsidies

Is the rationale for public disclosure instead that the "public" benefits through providing support for tax subsidies, and therefore tax filings should be made public? (Generally, imposing requirements conditions on tax-exempt status does not give rise to the argument of "unconstitutional conditions," because exemption is not a constitutional right.[72]) In 2000, the staff of the Joint Committee on Taxation released a congressionally-mandated study of the disclosure rules in the tax system, devoting a full volume to those that apply to exemption organizations.[73] The Joint Committee called for increased public disclosure of exempt-organization information, including the release of (1) complete private letter rulings and technical advice memoranda, without redaction of information identifying the entity and its transaction; (2) the results of all audits of tax-exempt organizations, also without redaction; (3) applications for exemption, not just exemptions once issued; (4) Forms 990-T (unrelated business income tax) and the returns of taxable affiliates; and (5) a description of lobbying activities, and amounts spent on self-defense lobbying and on nonpartisan research and analysis that includes a limited "call to action." Many of the Joint Committee's recommendations attracted strong criticism.[74] As mentioned above, Congress now requires disclosure of Forms 990-T (but not the returns of taxable affiliates); and as discussed below, the IRS must release determination letters denying or revoking exemption, although in redacted form.

The Joint Committee asserted the following rationale for public disclosure: "Disclosure of information regarding tax-exempt organizations also allows the public to determine whether the organizations should be supported—either through continued tax benefits or contributions of donors—and whether changes in the laws regarding such organizations are needed." That is, informing potential donors is one aspect, but only one, of this rationale. Just as important to the Joint Committee is allowing the public to judge the legitimacy of tax-exemption, and whether it should be altered.

3. Condition of Nonprofit (Epecifically, Charitable) Status

The Independent Sector, a leading trade association of charities, proposed an alternative rationale for transparency. In commenting on the Joint Committee's 2000 report, the Independent Sector declared: "IS believes that charities' public disclosure obligations derive from charities' fundamental nature as voluntary associations formed by private citizens to advance the public good—not from charities' receipt of favorable tax treatment."[75] After all, the Independent Sector observed: "Charities were recognized as separate entities with legal rights and responsibilities long before there was a federal income tax code. The need for disclosure stems from charities' unique social role. A charity must be transparent enough to make donors, volunteers, and partners confident that the charity will, in fact, advance public rather than private interests."[76]

As a general comment, the Independent Sector challenged the utility of counting on the Form 990, as it existed then, as the vehicle for informing the public: "Without an understandable user's guide—and no such guide exists—the public derives little benefit from much of the information already reported by charities. Thus, there is a deep need for tools to help the public understand the information that is already disclosed."[77] Independent Sector urged the IRS to revise the Form 990 "so that it highlights critical information and facilitates the reader's understanding of the significance of the information being presented. A top priority for the IRS in this regard should be providing, either directly or through nongovernmental intermediaries, on-line access to all Forms 990."

4. We Can't Think of a Better Alternative

Finally, we have to admit the possibility that we rely on public disclosure because we don't know what else to do (or who should do it). Betsy Adler nicely summarized the current regulatory approach with the acronym "FED": "funding, enforcement, disclosure."[78] In our laissez-faire system, we don't want government telling charities what to do and how to do it.[79] The absence of shareholders goes to why we disclose to regulators: By contrast, public disclosure seems driven by regulators' lack of resources, expertise, or inclination.

Nor should we discount the ceremonial value of sunshine. Public disclosure—even in the absence of enforcement action—is useful because knowing that information will be disclosed induces the fiduciaries to pay more (and better) attention not just to how they report, but also to what they do. At the same time, this leads to the possibility of fudging the reporting due to the pressures described in Part I. As the 2002 CPE text commented: "Several things must happen in order for this increased disclosure of Form 990 to be of maximum benefit to the public. First, the information entered on Form 990 must become more standardized and reliable. Second, potential users of the data must become more familiar with the requirements for proper completion of the return so that they will understand the data they are viewing."[80]

D. Disclosure of State and Federal Enforcement Activity

1. What Are the States Doing?

It is not easy to figure out how to spur nonprofit board members into performing better. Increasing monetary sanctions might make things worse: Indeed, we might improve nonprofit governance by reducing what's at stake. In large part regulators are so timid (at least publicly) because they don't want to discourage volunteers acting in good faith. As a result they don't send a sufficient signal (at least publicly) of the problems they encounter on nonprofit boards.[81]

But lack of transparency in their regulation of charities makes it impossible to assess the effectiveness of regulators in improving charity governance—or even whether they are acting at all. Few cases involving nonprofit fiduciary issues have reached the courts. Reform rather than punishment is generally the goal of the charity regulator, and charities as well prefer a chance to improve their behavior while avoiding embarrassment and personal liability. Most settlements are kept confidential. Finally, state attorneys general can act—or not act—out of parochial and political motives.[82]

Regulators have limited (financial and political) resources.[83] In that case, we might expect attorneys general to publicize their enforcement actions in order to benefit from the leveraging effect—miscreants in a

similar position would recognize themselves in the press release, and voluntarily straighten out.[84] Indeed, attorneys general do trumpet cases in which they catch someone violating the law. In other cases, where there's no real "bad guy"—but rather well-meaning fiduciaries caught in governance failures—states usefully could issue aggregate annual reports on the types of enforcement activities they undertook and outcomes achieved. See, for example, Pennsylvania's database of consent agreements and adjudications relating to charities, solicitors and fundraising counsel.[85] Regrettably, though, even the limited official reporting of enforcement activity tends to have a frustratingly short shelf-life. Press releases often vanish from attorney general websites when a new attorney general comes into office, thus undercutting the educational and deterrent value of publicizing enforcement actions; the Massachusetts attorney general's website no longer carries the very useful "Final Judgment Database" of legal actions, with links to the specific cases.

Private-sector solutions, while promising, have their own limitations. Notably, in 2008, the Charities Law Project at Columbia Law School began developing a website to assist attorneys general in fulfilling their responsibilities over charitable assets.[86] Although a separate intranet just for attorneys general might be created, so far most of the posted material is available to the public. The clearinghouse contains links to state and IRS websites (and specifically to state best practice guides) and summaries of law review articles. No enforcement materials have been posted yet, but a few recent settlements from around the country are available through links to materials for a panel on remedies presented at the March 2008 conference.[87] As of January 2011, the most recent conference shown on the project's website was held in April 2010; the page containing summaries of "AGs in the News" is current through November 2010.

2. IRS Determination Letters Denying or Revoking Exemption

As a threshold matter, despite the Service's fearsome reputation, it is as resource-constrained as the states. In 2010, the IRS Exempt Organizations Division employed only 942 people: 366 in Rulings and Agreements, 549 in Examinations, 14 in Customer Education and Outreach, and 13 in the EO Director's office. While total employment had gratifyingly grown from 837 in 2008 and 910 in 2009, the EO Division must oversee 1.8 million registered tax-exempt entities, including almost 1.2 million registered charities.[88] Thus, the development of published guidance (as well as examinations) suffers, putting pressure on practitioners to grasp at any type of informal guidance they can find.

Throughout the tax-practice world, practitioners and their clients have long benefited from the public availability of (redacted) versions of private letter rulings, audit memoranda, and other taxpayer-specific agency positions.[89] Marion Fremont-Smith explains how this type of informal transparency can improve tax administration in general: "Members of the bar were also able to identify issues needing study or revision, and call these to the attention of the Service as a group and not as partisans of individual clients."[90]

The Service, however, long refused to release redacted determination letters relating to denial or revocation of tax exemption. In a milestone decision issued in 2003, however, the District of Columbia Circuit held "that the portions of Treasury regulations sections 301.6110-1(a) and 301.6104(a)-1(i) that include denials and revocations 'within the ambit of section 6104' and prevent their disclosure violate section 6110's plain language."[91]

In annual revenue procedures, the Service sets forth the process for issuing determination letters and rulings on exempt status, both in response to applications for recognition of exemption and in cases of revocation or modification of determination letters or rulings. Section 8 of the revenue procedure describes the rules for disclosure. Notably, "[u]pon issuance of the final adverse determination letter or ruling to an organization, both the proposed adverse determination letter or ruling and the final adverse determination letter or ruling will be released under section 6110" . . . "after the deletion of names, addresses, and any other information that might identify the taxpayer", as set forth in Code section 6110(c).[92] Importantly, section 6104 applies only to material furnished by the organization or issued by the IRS,[93] and not to settlement agreements (termed "closing agreements") between the IRS and the organization unless the organization consents.[94]

These redacted denial and revocation letters began to appear in 2004. An early redacted denial letter was issued to a recreation center in which the Service found an inbred governance structure not likely to ensure public benefit; specifically, the IRS wrote: "Since all three members of your original board were related and

receiving compensation, we asked you to expand your board of directors by three to four non-related members of the community. [You added three new members.]" However, the IRS continued: "[a] full copy of your approved bylaws have not been received by the Service. The limited information provided indicates that the * * * may appoint and remove the directors. The * * * appear to be the three related directors."[95]

Incidentally, when faced with the prospect of a denial, why doesn't the applicant simply withdraw the application (this would not be a disclosable event)? Evidently, the denial letters are for groups that want judicial review, and the determination letter is the ticket to court. Alternatively, the IRS might back down and flag the file for examination after a period of operations.

With the continued issuance of denial and revocation letters, we have seen a flood of up to a dozen a week, adding up to hundreds a year.[96] An adverse ruling generally falls into one (or more) of three categories: private benefit, "commerciality," with, most recently, the return of the ground that the charity failed to conduct a charitable program "commensurate-in-scope" with its resources. The Service has denied exemption to nonprofits engaged in a variety of activities including adoption, insurance, financial services, religious publishing, conference centers, low-income housing, and retreats for caretakers—generally on the basis of their resemblance to similar for-profit businesses. Examples of recent determination letters with governance implications include the following, as summarized in the 2008 ACT report:

> PLR 200736031 (Dec. 7, 2006) (noting that married couple were sole officers and directors, there was no conflict of interest policy and couple did not recuse themselves when causing organization to contract for management services with for-profit company of which husband was sole shareholder); PLR 200535029 (June 9, 2005) ("Finally, despite the expansion of your governing board from three (3) to five (5) members, and the enactment of a conflict of interest policy, we still have some concern that your actual operations will be controlled and directed by B and his daughter C. We acknowledge that there is no evidence of any inurement to the benefit of these individuals, but then there has been no financial activity on your part to date.["]); PLR 200514021 (Jan. 13, 2005) ("There seems to be great likelihood of inurement to these individuals in that they all serve on the Board of Directors, and have a vote on compensation arrangements, leasing arrangements, and other financial matters that would affect the organization's financial interests as well as their own. This situation gives rise to an inherent conflict of interests that would potentially, adversely impact the financial well being of the organization. Thus, you have failed to show that B, C, D and E, through their positions on the Board, would not benefit from inurement . . . . ["]); PLR 200510031 (Nov. 15, 2004) ("There is not even one outside, disinterested board member to speak for the community. We must conclude that you violate the second fundamental rule for exempt organizations, and operate for private, not public benefit.["]) [97]

Unfortunately, the IRS website makes these exempt-organization determination letters available only as part of its general release of all determination letters.[98] Given how many of these determination letters we now have, and how cumbersome the process is of reviewing them, the Service—or another institution, with either public or private funding—could usefully collect and sort these documents.[99] The easiest way to find specific issues in these letters is to search a commercial electronic database, such as LEXIS or Westlaw.

Even when one can find a particular determination letter, the redactions[100] are simple elisions. As with all private rulings and memoranda, the redactors make no effort to give a sense of the substance underlying the facts.[101] Thus, we get such baffling indications as "$j" or "$ * * * " rather than, say, orders of magnitude, percentages, or relationships that would give a sense of the materiality of the problem; one recent revocation letter dealing with a complex structure referred to all names, places, and banks accounts by an undifferentiated "XX."[102]

The steady stream of denial and revocation letters has allowed the Service informally to stake out positions on basic substantive issues such as whether a particular activity is eligible for exemption.[103] For example, it is understood that the Service demands a minimum of three unrelated board members, although, because such a requirement does not appear in the statute or regulations, the Service cannot deny exemption on this basis alone. The 2008 ACT report comments: "We were not able to find guidance as to how the IRS takes governance

issues into account in the determination process, except in limited instances in the health care and low-income housing joint venture areas. We certainly appreciate that governance can bear on the operational test, among other issues. Our personal experience and research for this report suggest, however, that the IRS may require specific governance practices on an ad hoc and inconsistent basis." The Report cites two illustrations: "[D]etermination specialists may require organizations seeking exemption to have independent boards or at least some independent board members. Similarly, despite the fact that the Form 1023 specifically states that a conflict of interest policy is recommended but not required, our experience and interviews suggest that determination specialists often require adoption of such a policy, and occasionally require adoption of the sample form of policy included with the Form 1023 instructions."[104] Note, as the 2008 ACT report, adds: "There typically is no public record where taxpayers agree to make the changes required, strongly urged, or recommended by the IRS in the determination process and receive an exemption; or where an application is withdrawn."[105]

The ACT concludes that while "we have only anecdotal evidence regarding governance issues in the determination process . . . [,] the 'when' and 'what' [seem] unclear and not uniformly applied. We are concerned about the IRS having this level of discretion in cajoling or requiring specific governance process, particularly in the determination phase, where there usually is no track record evidencing operational failures."[106] Now, six years on, the IRS should use this substantial database of published denial and revocation letters to develop formal guidance. As with the revenue ruling on housing down-payment assistance organizations,[107] and in light of congressional endorsement of the Service's position on credit-counseling agencies,[108] the sector is entitled to revenue rulings or even regulations setting forth the agency's positions on organizational and operational issues, including nonprofit governance, that jeopardize exempt status. Such guidance would allow the Service to provide examples that show specific or relative dollar amounts and other facts masked by the redaction process.

3. Information Sharing: Disclosure from IRS to State Attorneys General

Amendments to Code section 6104 in the Pension Protection Act of 2006 (PPA) broadened the IRS's authority to provide certain information to state charity regulators, especially regarding exemption applications and denials.[109] The PPA extends to those state charity officials the section 6103(a) obligation to protect the confidentiality of the taxpayer information it receives. In March 2011, the Service proposed regulations under amended section 6104(c). The preamble emphasizes: "All disclosures authorized under section 6104(c) may be made only if the state receiving the information is following applicable disclosure, recordkeeping and safeguard procedures."[110] The National Association of State Charity Officials (NASCO) has commented, though, that in part because of the "cumbersome nature of the safeguard requirements and the resources needed to adhere to them," just three states (California, Hawaii, and New York) have reached information-sharing agreements with the IRS.[111] Indeed, NASCO asserted, the situation is now worse: "the PPA actually decreased disclosure of information to the states since the non-participating states no longer receive the pre-PPA notifications of final denials, revocations and notices of tax deficiencies."

E. Congressional Oversight

In a class by itself, and generally beyond the scope of this essay, was the devotion by Senator Charles Grassley—while he served as Chair and Ranking Member of the Senate Finance Committee—to publicizing abuses in the charitable sector. His most systematic effort began with a 2004 hearing and staff white paper on nonprofit governance,[112] followed by Senator Grassley's invitation to the Independent Sector to convene a blue-ribbon Panel on the Nonprofit Sector, which produced three influential reports.[113] Senator Grassley also issued a series of "love letters" to specific nonprofit organizations inquiring about their practices. This latter group included the American Red Cross, American University, the Nature Conservancy, and the Smithsonian Institution.[114] Industry-wide inquiries, often joined by Finance Committee chair Max Baucus, asked extensive questions about nonprofit hospitals' charity-care practices, higher educations' endowment spending, and, most recently, a group of televangelists of the "Prosperity Gospel" bent. These investigations had greater legitimacy when they covered nonprofit subsectors (rather than individual nonprofits) and the oversight of the IRS's performance in administering the laws. Indeed, Senator Grassley deserves much of the credit for the extensive exempt-organization reforms in the Pension Protection Act of 2006.[115] However, the IRS, as part of the executive

branch, has the enforcement responsibility and expertise to prosecute individual cases; moreover, as described above, the IRS must function under confidentially constraints that Senator Grassley seemingly felt unencumbered by. Perhaps not surprisingly, the first sign of public resistance to providing the information "requested" came from some of the televangelists.[116]

IV. Voluntary Disclosure by the Organization and Disclosure by Private Parties

A. Voluntary Disclosure by the Organization Itself

Charities often make disclosures to various constituencies without the compulsion of law. Prospective donors and grantmakers might condition funds on the production of satisfactory financial or other information. For example, before making grants to charities, many community foundations insist on being advised of such information as the names and relationships of board members and officers, the compensation of officers and relevant relationships, the identities of beneficiaries, audit data, and basic performance metrics. Government contracting rules, too, might demand reporting and audited financial statements. Beyond statutory requirements, the bylaws of membership organizations might require certain disclosures to the members. As discussed in Part II, charities have no excuse for refusing to provide basic information to members of the governing board, who should not be compelled to bring litigation to obtain that information on a timely basis.

While, as mentioned in Part III, the affairs of a nonprofit, non-governmental entity are private, and generally not subject to public disclosure, many of the reported troubles that have befallen charities in recent years could have been avoided had there been routine, timely and consistent public disclosure of basic information. Some of this information is already available through the regulatory and tax filings described in above, but usually only much after the fact (even when timely filed) and in a form that can be difficult for laymen to parse. The Panel on the Nonprofit Sector's Principles for Good Governance and Ethical Practices recommends: "A charitable organization should make information about its operations, including its governance, finances, programs, and activities, widely available to the public. Charitable organizations also should consider making information available on the methods they use to evaluate the outcomes of their work and sharing the results of those evaluations."[117] Charities should consider making clear in their bylaws or policies that transparency with the public is to be the norm, and deviations from that norm ought to require board consideration. The fact that transparency is the norm itself would deter many of the abuses made public.

For the benefit of the general public, nonprofits commonly post annual reports to their websites, but it is not so common to see Forms 990 and financial statements. For a laudable example of transparency, see the Ford Foundation's site,[118] which provides its articles of incorporation; bylaws; committee charters and membership; standards of independence; trustee code of ethics; staff code of conduct and ethics; procedures for approving affiliated grants; procedures for the receipt, retention and treatment of complaints regarding accounting, internal accounting controls and auditing matters; and annual reports and financial statements.

In a crisis, whether as a matter of damage control or sincerely to get ahead of the story, nonprofits should make timely disclosure. Spinning is a problem, though. For example, prior to the 2008 settlement, the dueling websites of the litigants over the Robertson gift to Princeton University to fund the Woodrow Wilson School represented an attempt to influence the court of public opinion.[119] Other recent scandals include the Smithsonian Institution[120] and the J. Paul Getty Foundation (discussed in Part IV.C).[121]

B. Media

Spurred by the perceived fund raising abuses by charities in response to the attacks on September 11, 2001, mainstream as well as specialty media interest in nonprofit governance has exploded. For those trying to keep up, important resources include the Chronicle of Philanthropy's daily posting of summaries (with links) of news stories published around the country,[122] as well as such legal nonprofit blogs as Don Kramer's Nonprofit Issues,[123] a group of legal academics' Nonprofit Law Prof Blog,[124] and Jack Siegel's CharityGovernance blog.[125]

Reporters often dwell on "fraud and abuse" in the nonprofit sector.[126] We run the risk, however, of overreaction to anecdotal information—since we don't know the denominator, is the fact that we're seeing more

stories an indication of increasing problems, or of increasing observation? In general, the increased availability of information on nonprofit operations increases the public expectation for more transparency.

C. Peer Regulators and Charity "Watchdogs"

Peer regulation in the nonprofit sector comes in two flavors—the third-party watchdogs and the trade associations. The watchdogs are donor-focused, and they typically provide assessments (sometimes using a star system or letter grades) regardless of whether the charity knows about the review or supplies information. However, the BBB Wise Giving Alliance—which assesses whether a given charity meets or does not meet its Standards for Charity Accountability—relies on information from the charity and states cases in which the organization failed to respond.[127]

By contrast, membership in the trade associations is voluntary, with the organizational member submitting both to the groups' standards[128] and to any disciplinary process for violation. Most groups are not as open as the Evangelical Council for Financial Accountability, which posts a chart of former members, indicating the reason—voluntary resignation or termination.[129] For example, Brian Gallagher, head of the United Way of America, said at the July 22, 2004 Senate Finance Committee roundtable that the UWA has decertified 30 UW's around the country in the previous two years. This information should have been more widely known—I couldn't even find it on the UWA's website. Peer organizations generally seem loathe to publicly discipline noncompliant members. While still an anomaly, compare the Council on Foundation's brief suspension of the J. Paul Getty Trust's membership, ending when the Trust adopted "reforms including new training and evaluation tools for board members, strengthened conflict-of-interest provisions, increased board oversight of real estate deals, and increased transparency of staff compensation and performance reviews."[130]

Finally, there is the behavior of nonprofit groups speaking out—or, more likely not—about specific misbehaving organizations or unacceptable practices as they occur. Isn't protection of the sector's reputation a duty of nonprofits themselves? The Independent Sector's Panel on the Nonprofit Sector energetic response to Senator Grassley's 2004 staff white paper culminated in a report containing 33 principles of self-regulation.[131] Some members of the working group, however, were disappointed that the principles are precatory only, and that the Nonprofit Panel could not achieve consensus around adopting a mechanism for certification and discipline. Deciding how to bell the cat is never easy.

Conclusion

The Internal Revenue Service does not have the resources to verify all tax exemptions on a routine basis. Rather, the IRS conducts a relatively small number of examinations (including targeted correspondence audits) of specific charities, either as part of a system of examining Forms 990 or pursuant to a particular compliance initiative (such as on political campaign activity, hospitals, and institutions of higher education).[132]

In 2009, the IRS's Exempt Organizations Division released a Governance Check Sheet[133] and a Governance Project Guide Sheet[134] for Completing the Project Check Sheet to be used by agents in examining Code section 501(c)(3) exempt organizations. The public can access these guidelines from a new webpage that explains: "A check sheet will be used by IRS' Exempt Organizations Examination agents to capture data about governance practices and the related internal controls of organizations being examined. The data will be included in a long-term study to gain a better understanding of the intersection between governance practices and tax compliance."[135] The webpage links to the Check Sheet and Guide Sheet and to other governance materials on the website,[136] notably an article entitled "Governance of Charitable Organizations and Related Topics"[137] included in the Life Cycle on-line educational tool for charities.

The IRS's recent focus on exempt-organization governance has attracted thoughtful commentary on both sides of the issue. Thomas Silk supports this endeavor of the IRS:

> It is not far-fetched to imagine a national scandal featuring a prominent charity in violation of standards of charitable governance but incorporated in a state with inadequate charitable enforcement. In the congressional hearings that might follow, the IRS would surely be in a far more defensible position if it had already gone forward to educate the charitable sector about

> the importance of good governance practices. Later legislation introduced by a supportive Congress may easily resolve any jurisdictional ambiguities about governance of charitable organizations and enforcement.[138]

On the other hand, Bonnie Brier (lead author of the 2008 ACT report quoted above) recently expressed skepticism that the described governance practices actually lead to good governance, and worries that charities will adopt them just to satisfy the IRS regardless of whether they are appropriate for the organization.[139] Marcus Owens, former top exempt organization official at the IRS, questions the IRS's authority to include governance questions on the Form 990. Senator Grassley responded to such objections by proposing legislation to provide statutory authority for the IRS to assert an interest in charity governance as an indicator of compliance with the federal tax-exemption regime.

I generally disagree with those critical of a role for the IRS in charity governance, at least to the extent these criticisms apply to the governance questions on the redesigned Form 990. Indeed, as described above, I submitted comments to the IRS on the 2007 draft of the redesigned Form 990, proposing for inclusion a series of questions on organizational structure and governance practices[140]—many of which were added in the final version. At that time, I had in mind the usefulness of the Form 990 to the governing board itself and to state regulators, to donors, to the media, and, yes, to researchers, even aside from what uses the IRS might make of the data. While recognizing the values of privacy discussed above, on balance I believe, these interests do not outweigh the benefits from transparency of the organization's governance structure to these outside constituencies. If a particular "best" practice is inappropriate in a particular case, the charity can and should provide an explanation on the Form 990. Thoughtful additional disclosure is an opportunity for the organization to demonstrate—if it can—how its structure and policies appropriately safeguard charitable assets.

Endnotes

1 Professor Law, Chicago-Kent College of Law, Illinois Institute of Technology. This essay grew out of my Norman A. Sugarman Lecture, "Governing in a Fishbowl: The Effects of Sunlight on Nonprofit Accountability," Mandel Center for Nonprofit Organizations, Case Western Reserve University (Cleveland, November 11, 2004). I appreciate comments received from Laura Brown Chisolm, Paul Feinberg, and students and other attendees. I am grateful for opportunities to present drafts at the Internal Revenue Service's 2011 Research Conference; the Chicago Bar Association Committee for Trade and Professional Associations Law (Chicago, January 11, 2011); the 38th Annual Conference of the Association for Research on Nonprofit Organizations and Voluntary Action (Alexandria, Virginia, Nov. 18, 2010); the Advanced Topics in Taxation Colloquium Series, Northwestern University School of Law (Chicago, April 19, 2010); and from Elizabeth Boris and discussant Joseph Cordes at the 2011 IRS-TPC Research Conference (Urban Institute, Washington, D.C., June 22, 2011).

This article cites portions of draft chapters I prepared for the American Law Institute as reporter of its project on Principles of the Law of Nonprofit Organizations. While Tentative Draft No. 1 (2007 and 2008) and Tentative Draft No. 3 (2011) (through 660) were approved by Council and the membership, they are subject to further revision and final approval.

The views in this paper do not necessarily represent those of the Internal Revenue Service. I thank Alan Plumley for permission to publish this article, in substantially identical form, in the 12 Fla. L. Rev. 183 (2012).

2 Evelyn Brody, *Charities in Tax Reform: Threats to Subsidies Overt and Covert*, 66 Tenn. L. Rev. 687, 694 (1999).

3 See Pub. L. 100-203, Sec. 10702(a), adding subsection (e) to Internal Revenue Code § 6104. In 1998 Congress replaced subsections (d) and (e) with the current subsection (d).

4 As I wrote in 1996:

The wealth of data demanded by the IRS inspired the following exchange between a member of the American Bar Association and the IRS's special assistant for exempt organization matters:

Mr. Gallagher: Howard, what does the IRS do with all this stuff?

> *Mr. Schoenfeld*: It's not so much what the IRS does with all this stuff. It's also what the public does with all this stuff. That's an equal part, I think, of what the question should be.

But what, then, is the public to do with all this stuff?

Evelyn Brody, *Institutional Dissonance in the Nonprofit Sector*, 41 Vill. L. Rev. 433, 501 (1996) (citing the exchange in *Edited Transcript of the Morning Sessions of the August ABA EO Committee Meeting in New Orleans, Panel III, The Role of the IRS in Promoting Public Accountability of Exempt Organizations*, 10 Exempt Org. Tax Rev. 805 (Oct. 1994).

[5] The closing agreements with Jimmy Swaggart Ministries, Pat Robertson's Christian Broadcasting Network, and Jerry Falwell's Old Time Gospel Hour not only required the payment of taxes, but also required the organizations to make changes in corporate governance and to publicize the general terms of the closing agreements. See *Public Statement, Jimmy Swaggart Ministries, Baton Rouge, Louisiana* (Dec. 17, 1991), 5 Exempt Org. Tax Rev. 205 (Feb. 1992); *Statement of Jerry Falwell Regarding Closing Agreement* (Feb. 17, 1993), 7 Exempt Org. Tax Rev. 876 (May 1993). See also Closing Agreement on Final Determination Covering Specific Matters (Oct. 1, 1993), available in LEXIS, Fedtax Library, Tax Notes Today File, as 97 TNT 251-24 (Dec. 1, 1997) (purporting to be between the IRS and the Church of Scientology, but never acknowledged by either party).

[6] Evelyn Brody, *A Taxing Time for the Bishop Estate: What Is the I.R.S. Role in Charity Governance?*, 21 U. Haw. L. Rev. 537 (1999) (Bishop Estate Symposium Issue). For the August 18, 1999 closing agreement between the Internal Revenue Service and the Kamehameha Schools / Bishop Estate (KSBE), go to www.ksbe.edu/newsroom/filings/final_v021029.pdf. This Closing Agreement required—in addition to a payment from KSBE to the IRS of $9 million plus interest (for a total of about $14 million)—significant governance reforms, as well as the Internet posting of the final Closing Agreement. The Closing Agreement did not cover any personal tax liability of the trustees.

[7] Treas. Reg. § 53.4958-6.

[8] See discussion in Part III.E, below. Senator Grassley's many press releases and accompanying documents are available at http://grassley.senate.gov/news/press_releases.cfm.

[9] The 2007 draft Form 990 and related schedules, the draft instructions, and the comments on these drafts, along with educational material, are available at www.irs.gov/charities/charitable/article/0,,id=185892,00.html. See also the Urban Institute's Center on Nonprofits and Philanthropy and Harvard University's Hauser Center for Nonprofit Organizations, 17th Emerging Issues in Philanthropy Seminar, "IRS Form 990 Redesign" (Washington, D.C., Sept. 10, 2007), at which over 60 attendees, including representatives from the Internal Revenue Service and congressional staff, as well as from sector organizations, practitioners, and scholars, discussed the draft Form 990.

[10] For example, a survey conducted by the charity watchdog BBB-Wise Giving Alliance suggested that the public does not accept fund raising costs over 15 percent, an unrealistically low number. *See* Grant Williams, *Watchdog Group Proposes Changes in Evaluating Charity Operations*, Chron. of Philanthropy, Jan. 24, 2002.

[11] Report of the Independent Expert Group on Expenses 9 (Feb. 2010), available at www.ncvo-vol.org.uk/sites/default/files/ExpensesReportfinal.pdf. The study was prompted by a scandal that erupted in the United Kingdom over expense reimbursements claimed by members of Parliament. *Id*. at 16-17.

[12] *Id*. at 50-51.

[13] See generally Evelyn Brody, *Accountability and Public Trust*, in The State of Nonprofit America 479 (Lester M. Salamon, ed. 2004).

[14] A.L.I., Principles of the Law of Nonprofit Organizations § 320(b)(8) (Tentative Draft No. 1, 2007 & 2008).

[15] See *id*., § 340. The most common exceptions are for some personnel issues.

[16] *Id*. § 320, comment g(6), suggests a list of documents that every board member should receive, including the current (and dated) versions of the charity's trust instrument or articles of incorporation; bylaws;

board policies applicable to board members (e.g., conflicts of interest, travel and expense reimbursement, and confidentiality, and any general ethical policy); a directory (with contact information) of board members and officers; charters of any board committees and committee assignments; an organizational chart and contact information for senior staff; the current budget and recent financial statements, including the outside's auditor's management letter; recent Forms 990; minutes of recent board meetings and, if applicable, of executive committee meetings; the charity's mission or vision statement, if prepared; and a schedule of dates and locations of upcoming meetings of the board and of the membership (if any).

[17] See exploration of this issue in Evelyn Brody & John Tyler, *How Public Is Private Philanthropy?: Separating Reality from Myth* (Philanthropy Roundtable Monograph 2009), available at www.philanthropyroundtable.org/files/Public_Private%20Monograph_high%20res_Final.pdf, and updated as *Respecting Foundation and Charity Autonomy: How Public Is Private Philanthropy?*, 85 Chi.-Kent L. Rev. 571 (2010).

[18] Francie Ostrower, Nonprofit Governance in the United States: Findings on Performance and Accountability from the First National Representative Study (Urban Institute 2007), available at www.urban.org/UploadedPDF/411479_Nonprofit_Governance.pdf.

[19] Importantly, for the subset of charities dubbed "private foundations" by federal tax law are prohibited from entering into transactions with insiders—other than the payment of reasonable compensation for services rendered. I.R.C. § 4941 (Taxes on Self-Dealing). For a full discussion of interested transactions, see § 330 of the 2007 ALI draft Nonprofit Principle, *supra* note 14.

[20] Ostrower, *supra* note 18, at 8 (footnote omitted). That study found: "among nonprofits engaged in financial transactions, most obtained goods at market value (74 percent), but a majority (51 percent) did report that they obtained goods below market cost. Under 2 percent reported paying above market cost. Keep in mind, too, that these are self-reports, and thus, if anything, the figures are likely to underreport transactions resulting in obtaining goods at above market value or at market value costs and overreport transactions resulting in obtaining goods below market cost." *Id*. (footnote omitted).

[21] See, e.g., Michael W. Peregrine, Ralph E. DeJong & Timothy J. Cotter, *Transparency: What the EO Board Needs to Know about Executive Compensation*, 46 Exempt Org. Tax Rev. 23 (Oct 2004).

[22] In contrast to questions asking whether the organization has written policies addressing conflicts of interest, whistleblowers, and document retention and destruction, and about participation joint ventures, the question relating to the process for determining the compensation of top management, officers, and key employees is rooted in statutory and regulatory requirements. See Code § 4958 (excess benefit transactions engaged in by 501(c)(3) and (c)(4) organizations). Treas. Reg. § 53.4958-6 sets forth a rebuttable presumption that a compensation arrangement or other transaction is reasonable if it is (1) approved in advance by an independent body acting for the organization (2) that obtained and relied on appropriate comparability data, and (3) that the body adequately documented its determination.

[23] Form 990 (2008), Core Form, Part VI (Governance, Management, and Disclosure), line 10. It is unfortunate that this question does not allow for the alternative of review prior to filing by a board committee, as recommended in comments submitted on the draft redesign. Regrettably, many time-pressed charities will likely prefer to file under an extension than answer "no."

[24] A charity might be expected to have to ensure that it preserves its tax exemption, but a charity may relinquish tax exemption "so long as the charitable organization's fiduciaries can demonstrate that they made a good faith determination that loss of exemption was in the best interests of the organization." Marion R. Fremont-Smith, "Relinquishing Tax Exemption: State and Federal Constraints," presented at the Nonprofit Forum, New York City (Oct. 16, 1991).

[25] Compare the new requirement that Forms 990-T, on which an exempt organization reports its unrelated business taxable income, are now subject to public disclosure. See Part III, below.

[26] This topic was the subject of an Emerging Issues in Philanthropy Seminar, sponsored jointly by the Urban Institute's Center on Nonprofits and Philanthropy and by Harvard University's Hauser Center for Nonprofit Organizations (Cambridge, Mass., Nov. 30, 2000).

[27] See Eugene Steuerle, *When Nonprofits Conduct Exempt Activities as Taxable Enterprises*. Emerging Issues in Philanthropy Seminar No. 4 (The Urban Institute and The Hauser Center for Nonprofit Organizations), available at www.urban.org/url.cfm?ID=310254.

[28] Fremont-Smith identifies New York, California, Massachusetts, Ohio, Illinois, Minnesota, and New Hampshire. Marion R. Fremont-Smith, Governing Nonprofit Organizations: Federal and State Law and Regulation 316-17 (Cambridge, Mass.: Belknap Press of the Harvard University Press 2004).

[29] *Id*. at 317 (identifying California, Mississippi, Minnesota and Oregon).

[30] The webpage for this project is: www.uniformlaws.org/Committee.aspx?title=Protection%20of%20Charitable%20Assets%20Act.

[31] Riley v. National Federation of the Blind of North Carolina, Inc., 487 U.S. 781 (1988); Maryland v. Joseph H. Munson Co., 467 U.S. 947 (1984); Village of Schaumburg v. Citizens for a Better Environment, 444 U.S. 620 (1980).

[32] *Madigan v. Telemarketing Associates, Inc.*, 538 U.S. 600 (2003).

[33] Version v4.01 (May 2010) supports 37 jurisdictions (36 states and the District of Columbia), and includes supplemental forms required by 13 jurisdictions. Go to www.multistatefiling.org. This charitable-solicitation registration form resulted from a joint project of the National Association of State Charities Officials, the National Association of Attorneys General, and the Multi-State Filer Program, a consortium of nonprofits.

[34] Churches (and their integrated auxiliaries) and small public charities (normally, $5,000 or less in gross receipts) are exempt from having to apply for recognition of tax exemption under Code § 501(c)(3), and churches and most small public charities (normally, after a phase-in period for tax years ending in 2010, $50,000 or less in gross receipts) do not have to file the annual Form 990 or Form 990-EZ. The requirement to file an "e-postcard"—Form 990-N can be filed only online—applies to small charities, but not to churches. The new legislation additionally requires notification to the IRS when an exempt organization terminates its existence. See I.R.C. §§ 6033, 6652, and 7428, as amended by the Pension Protection Act of 2006 § 1223.

[35] The number of § 501(c)(3) exempt organizations appearing in the IRS Business Master File grew 350 percent from 1975 to 2001, and, as of 2004, stood at 1,010,365. See Joint Committee on Taxation, Historical Development and Present Law of the Federal Tax Exemption for Charities and Other Tax-Exempt Organizations (JCX-29-05, April 19, 2005) (citing to IRS Statistics of Income Division reports and the Business Master File). The 2009 IRS Data Book reports almost 1.2 million 501(c)(3) organizations. See IRS Data Book, 2009, Pub. 55B, Washington, D.C. (March 2010), at www.irs.gov/pub/irs-soi/09databk.pdf, at tables 24 and 25. (All private foundations, regardless of revenue level, must file, and the Pension Protection Act of 2006 requires supporting organizations and organizations with controlled entities to file Form 990 even if their gross receipts are less than $25,000.)

[36] To publicize the new filing requirement for small charities, the IRS identified and contacted 640,000 potential e-Postcard filers in its database; based on survey results and historical filing patterns, it expected 166,000 e-Postcard filers. *EO 2008 Annual Report and 2009 Work Plan* 12 (Nov. 2008), available at www.irs.gov/pub/irs-tege/finalannualrptworkplan11_25_08.pdf. Evidently, filings came in from organizations too small to have had to file an exemption application (and thus do not appear on the Business Master File). A 2010 national study found that the largest categories of nonfilers were human service organizations (29 percent), public and societal benefit organizations (22 percent), and education organizations (15 percent); volunteer-run organizations, often with changing addresses, predominated. Amy Blackwood and Katie L. Roeger, National Center for Charitable Statistics, *Here Today, Gone Tomorrow: A Look at Organizations that May Have Their Tax-Exempt Status Revoked* 2-3 (July 8, 2010), available at http://www.urban.org/publications/412135.html. For information identifying those organizations that automatically lost their exemption for failure to file—and providing a one-time opportunity for retroactive reinstatement—go to www.irs.gov/charities/article/0,,id=239696,00.html.

[37] Kirsten A. Grønbjerg, Kellie McGiverin-Bohan, Kristen Dmytryk, and Jason Simons, *IRS Exempt Status Initiative: Indiana Nonprofits and Compliance with the Pension Protection Act of 2006*, at p. 18, Indiana Nonprofits: Scope & Community Dimensions Briefing 2011: No. 1 (July 1, 2011 (revised July 29, 2011), available at http://www.indiana.edu/~nonprof/results/database/IRSRevocation.html. This report found that suffering the highest rates of revocation were cemeteries, advocacy organizations, and nonprofit business associations, while "[f]raternal organizations, veterans groups and other organizations with close connections to national groups were most successful in avoiding having their tax-exempt status revoked, suggesting that communications networks helped such groups comply with the law." *Id*. at 3.

[38] The exemption application, Form 1023, asked about organizational form and changes in organizational form should have been reported on the Form 990, but this process was unreliable. Note that the IRS did not redesign the simplified Form 990-EZ or the private foundation form, 990-PF.

[39] See Comment Letter from Evelyn Brody to IRS (Sept. 14, 2007), available at pages 24-28 of www.irs.gov/pub/irs-tege/redesignedform990commentsgeneral_9_14_07_i.pdf.

[40] See generally Elaine Waterhouse Wilson, *More Than You Ever Wanted to Know (or Tell!): Heightened Compensation Disclosure on the New Form 990*, 60 Exempt Org. Tax Rev. 273 (June 2008).

[41] Brody comment, *supra* note 39.

[42] Advisory Committee on Tax Exempt and Government Entities (ACT), *The Appropriate Role of the Internal Revenue Service with Respect to Tax-exempt Organization Good Governance Issues*, June 11, 2008, at 3, page 89 of the PDF available at www.irs.gov/pub/irs-tege/tege_act_rpt7.pdf.

[43] *Id*. at 29 (PDF at 115). The report cites to Dana Brakman Reiser, *There Ought to Be a Law: The Disclosure Focus of Recent Legislative Proposals for Nonprofit Reform*, 80 Chi.-Kent L. Rev. 559 (2005).

[44] For tax years beginning in 2008, an exempt organization with annual revenue of more than $25,000 and less than $1 million and assets of less than $2.5 million could file the simpler Form 990-EZ. For the 2009 year, the cutoff dropped to less than $500,000 of revenue and less than $1.25 million of assets. For 2010 and later, the lower end of the revenue breakpoint rises to more than $50,000 and the upper end drops to less than $200,000 (see note 36, above, for the Form 990-N "e-postcard") and less than $500,000 in assets. See Internal Revenue Service, Overview of Form 990 Redesign For Tax Year 2008 (Dec. 20, 2007), at 2, available at www.irs.gov/pub/irs-tege/overview__form__990__redesign.pdf.

[45] Rev. Proc. 2011-15 (Jan. 13, 2011), available at www.irs.gov/pub/irs-drop/rp-11-15.pdf.

[46] Note Urban Institute studies showing that a high percentage of Forms 990 are filled out by professionals. Fremont-Smith, *supra* note 28, at 457-58. Thus this is not a question of amateurs not knowing what they're doing.

[47] For example, in 2004, the Pennsylvania secretary of state filed suit against nonprofits and their officers for 1,200 false Forms 990's, which were also filed with the state. The 2007 settlements (available from the database in note 85, below) called for four national charities to pay $150,000 and to stop fund raising in Pennsylvania; the charities acknowledged that they did not report, among other things, H.R. Wilkinson as a key employee; related-party transactions; and relationships among officers, employees, directors or members.

[48] Advisory Committee on Tax Exempt and Government Entities (ACT), *Exempt Organizations: Group Exemptions—Creating a Higher Degree of Transparency, Accountability, and Responsibility*, June 15, 2011, starting at page 291 of the PDF available at www.irs.gov/pub/irs-tege/tege_act_rpt10.pdf.

[49] See generally Code § 6103.

[50] See generally Code § 6104.

[51] See Part III.E, below.

[52] See Brody & Tyler, *supra* note 17, at 597 nn.61 & 62 and accompanying text.

[53] Under federal election laws, because of the enhanced public interest in open and fair elections, generally all but the smallest donors and amounts contributed to federal political campaigns must be identified (some states have similar "clean government" rules); this result leads some strategists to

advise conducting issue-related advocacy through Internal Revenue Code § 501(c)(4) organizations. Issues relating to political activity and election law and regulation, including tax-law rules and filing requirements, are generally beyond the scope of this essay.

54 N.J. Stat. § 45:17A-24(f).

55 Go to www.guidestar.org. See also the Urban Institute's National Center for Charitable Statistics' website for filed Forms 990, along with statistical analysis, at http://nccs.urban.org.

56 Copies of Scanned EO Returns Available, www.irs.gov/charities/article/0,,id=150268,00.html.

57 See Cheryl Chasin, Debra Kawecki & David Jones, *Form 990*, Chapter G of FY2002 IRS EO Continuing Professional Education Text, available at www.irs.gov/pub/irs-tege/eotopicg02.pdf

58 For "EO Tax Law Training Articles" from FY1979 through FY2004, go to www.irs.gov/charities/article/0,,id=161088,00.html. These training materials are also available through a topical index, at www.irs.gov/pub/irs-tege/cpeindexbytopic.pdf.

59 The non-disclosure of the identity and contributions of donors to exempt organizations has made 501(c)(4) organizations, which can engage in political speech so long it is not their primary activity, a tempting vehicle for avoiding the disclosure requirements of federal election law. See the brief discussion in note 67, below, and accompanying text.

60 Joint Committee on Taxation, *General Explanation of Tax Legislation Enacted in the 109th Congress*. JCS-1-07, Jan. 17, 2007. Available at www.house.gov/jct/s-1-07.pdf.

61 The current version, revised in 2006, is available at www.irs.gov/pub/irs-pdf/f1023.pdf. It would be great if GuideStar could collect and post these once exemption is granted. The Forms 1023 (especially the ones filed electronically, when the IRS makes this process available) would provide an interesting database for study.

62 ACT, *supra* note 42, at 32-33 (pages 18-19 of the PDF) (footnotes omitted).

63 "Re-Engineering Form 1023 to Identify Problem Organizations Before Exemption Is Granted: Watch out for the "Penalties of Perjury" Statement" (November 3, 2004), at http://charitygovernance.blogs.com/charity_governance/2004/11/reengineering_f.html.

64 See United States v. Mubayyid, 658 F. 3d 35 (1st Cir. 2011), upholding convictions for filing false Forms 990 within the limitations period against defendant Mubayyid.

65 In an oral comment at the Senate Finance Committee Staff Roundtable held in Washington, D.C., on July 22, 2004 (which this author attended), attorney Douglas Mancino recommended that exempt organizations be required to report compensation on a more current basis, citing as precedent the quarterly filings required by the SEC of public companies.

66 IRS, *e-File for Charities and Nonprofits*, at www.irs.gov/efile/article/0,,id=108211,00.html.

67 *Id*.

68 See also GuideStar's service: "Gov@GuideStar offers a suite of tools designed specifically for government users of GuideStar data. These research and reporting tools enable government decision makers to perform critical tasks with greater ease and confidence."

69 2010 IRS Data Book tbl. 2 (released March 11, 2011), available at www.irs.gov/taxstats/article/0,,id=102174,00.html.

70 IRS Exempt Organizations FY 2010 Annual Report and FY 2011 Work Plan, Dec. 15, 2010, at 13, available at www.irs.gov/pub/irs-tege/fy2011_eo_workplan.pdf.

71 See Citizens United v. FEC, 130 S. Ct. 876 (2010); the proposed DISCLOSE Act (2010); and papers presented at "Nonprofit Speech in the 21st Century: Time for a Change?," the Annual Conference of the National Center on Philanthropy and the Law, New York University School of Law (New York City, Oct. 28-29, 2010).

72 Compare the constitutional limits on mandated speech, as set forth in the *Riley* trilogy discussed above. See generally Evelyn Brody, *Entrance, Voice, Exit: The Constitutional Bounds of the Right of Association*, 35 U.C.-Davis L. Rev. 821 (2002).

[73] U.S. Congress, Joint Committee on Taxation. 2000. *Study of Present-Law Taxpayer Confidentiality and Disclosure Provisions as Required by Section 3802 of the Internal Revenue Service Restructuring and Reform Act of 1998, Volume II: Study of Disclosure Provisions Relating to Exempt Organizations*. JCS-1-00, Jan. 28. Available at www.house.gov/jct/s-1-00vol2.pdf. Pages 89-90 of the report provide a background to Form 990 public disclosure.

[74] See Grant Williams, *Tax Report Shakes Up Charities*, Chron. of Philanthropy, March 9, 2000.

[75] Independent Sector Comments on Joint Committee on Taxation Study on Disclosure by Tax-exempt Organizations 3 (no date), available at www.independentsector.org/programs/gr/Comments.PDF.

[76] Separately, the Independent Sector "[took] issue with the JCT Report's characterization of tax exemption and the charitable deduction as government subsidies and the Report's view that the receipt of those subsidies creates a strong presumption in favor of increased disclosure." The Independent Sector pointed to "years of serious academic debate over whether the charitable exemption and deduction are appropriately viewed as special benefits or as structural necessities of a properly calculated income tax."

[77] *Id*. at 8. A few years later, the IRS included in its 2003 Exempt Organization Continuing Professional Education text a helpful set of Q&A's on how to fill out (and therefore read) the Form 990. Go to www.irs.gov/pub/irs-tege/eotopich03.pdf.

[78] Oral comment of Betsy Adler, then-chair of the Exempt Organizations Committee of the American Bar Association Tax Section, Senate Finance Committee Staff Roundtable (Washington, D.C., July 22, 2004) (author's notes).

[79] Evelyn Brody, *Agents Without Principals: The Economic Convergence of the Nonprofit and For-Profit Organizational Forms*, 40 N.Y.L. Sch. L. Rev. 457 (1996).

[80] Chasin, et al., *supra* note 56.

[81] For draft principles relating to enforcement, see § 610 (state attorneys general) and § 620 (Internal Revenue Service as charity regulator) in American Law Institute, Principles of the Law of Nonprofit Organizations (Tentative Draft No. 3, 2011) (approved through section 660).

[82] See Evelyn Brody, *Whose Public? Parochialism and Paternalism in State Charity Law Enforcement*, 79 Ind. L.J. 937 (2004).

[83] *Id*. Garry Jenkins conducted a survey, to which all but one of the states responded, finding that 74 percent of the states had one or fewer full-time-equivalent attorneys devoted to charitable oversight, and that 17 states assigned no attorneys to that function. Garry W. Jenkins, *Incorporation Choice, Uniformity, and the Reform of Nonprofit State Law*, 41 Ga. L. Rev. 1113, 1128-29 (2007). Legal staffs exceeding 2.5 FTE's are found in California (12), Connecticut (5), Illinois (7), Indiana (4), Massachusetts (6), Minnesota (5), New York (20.5), Ohio (10), Pennsylvania (12), and Texas (6). *Id*. at table 1.

[84] Recall Dr. Strangelove's complaint: "Deterrence is the art of producing, in the mind of the enemy, the fear to attack. The whole point of the Doomsday Machine is lost if you keep it a secret! Why didn't you tell the world, eh?!" See www.youtube.com/watch?v=cmCKJi3CKGE (The Doomsday Machine in *Dr. Strangelove*), at 3:50.

[85] Available at www.portal.state.pa.us/portal/server.pt/gateway/PTARGS_0_0_84867_0_0_43/http;/pubcontent.state.pa.us/publishedcontent/publish/cop_general_government_operations/dos/am/content/charities/consent_agreements_and_adjudications.html?qid=83559484&rank=10.

[86] Go to www.law.columbia.edu/center_program/ag/policy/CharitiesProj/.

[87] Go to www.law.columbia.edu/center_program/ag/policy/CharitiesProj/events/conference/ConferenceMar08.

[88] IRS Exempt Organizations FY 2010 Annual Report, *supra* note 69, at 6. See also the Government Accountability Office report for a 2005 House Ways and Committee Hearing, *Tax-Exempt Sector: Governance, Transparency, and Oversight Are Critical for Maintaining Public Trust* 17 (GAO-05-561T 2005) ("From fiscal year 2000 through 2004, IRS staffing for overseeing tax-exempt entities stayed relatively flat as measured by the number of FTE staff assigned to oversee tax-exempt entities.").

Despite its mind-boggling potential workload, TE/GE's enforcement activities reach only a small fraction. *EO Tax Journal* editor Paul Streckfus commented of the data on compliance reported on page 2 of EO's FY 2010 Annual Report: "The graph tells us that in FY 2009—of 16,960 returns examined—6,773 pertained to compliance checks and 10,187 pertained to traditional examinations." He adds: "[O]f those 10,187 returns examined, only 3,445 were Forms 990 and 990-EZ. The rest were mostly employment tax returns (4,582) and 990-Ts (962). . . . [M]ost audits involve more than one year, so an audit of one organization may involve multiple 990s. My best guess was that this translated to 1,723 organizations being subject to a traditional audit in FY 2009. . . . Regardless, we are talking a .002 audit rate, not 2%, but .2 %, pretty close to infinitesimal, especially when you exclude targeted audits [of colleges and hospitals]. . . ." *EO Tax Journal 2010-185* (Dec. 16, 2010).

89 It took a series of Freedom of Information Act lawsuits by Tax Analysts, publisher of *Tax Notes* magazine and the *Exempt Organization Tax Review*, to compel the IRS to release these items.

90 Fremont-Smith, *supra* note 28, at xiv.

91 Tax Analysts v. Internal Revenue Service, 350 F.3d 100, 104-05 (D.C. Cir. 2003). The court described the legislative history: "Congress passed the Tax Reform Act [of 1976] to protect taxpayer privacy while requiring the IRS to disclose written determinations. Our holding advances that purpose: the IRS must disclose determinations denying or revoking tax exemptions, but do so in redacted form, thus protecting the privacy of the organizations involved. The Treasury regulations, in contrast, keep denials and revocations completely secret, preventing the very monitoring of the IRS that the Tax Reform Act was designed to facilitate." *Id*. at 104.

92 Rev. Proc. 2011-9, 2011-2 I.R.B. 283 (Jan. 10, 2011), at § 8.02.

93 Because a closing agreement is a "bilateral agreement signed by both IRS and taxpayer," it was "not issued by" the IRS and thus was not subject to the clause of Code § 6104(a)(1)(a) making disclosable information issued by IRS with respect to organization's application for tax-exempt status. Tax Analysts v. Internal Revenue Service, 2004 U.S. Dist. LEXIS 28032, 93 AFTR 2d 1250, n.2 (D.D.C. 2004), *citing to prior decision at* 53 F. Supp. 2d 449, 452-53 (D.D.C. 1999). This litigation ended when the D.C. Circuit upheld the district court's refusal to compel the Service to disclose the closing agreement referred to in a press release issued by the Christian Broadcasting network.

94 The IRS Chief Counsel's office notified its attorneys of the procedures to follow "when advising Internal Revenue Service employees concerning a determination that publicizing a closing agreement between a taxpayer and the Internal Revenue Service advances tax administration." When the parties agree that "public disclosure of a closing agreement (or any of its terms)" is warranted, in general, it would be through an IRS news release, or a jointly authored statement, which would be released at the time the closing agreement is executed." Chief Counsel Notice CC-2008-014 (April 14, 2008).

95 Determination Letter 20044033E (April 5, 2004) (released Oct. 18, 2004).

96 Author's estimate. The FY 2010 Annual Report makes no mention of either revocation or denial numbers, nor of closing agreements. In earlier years, the IRS finalized 78 closing agreements with § 501(c) organizations in fiscal year 1999; 72 in fiscal year 1998; and 65 in fiscal year 1997. See Joint Comm. on Taxation, *Study of Present-Law Taxpayer Confidentiality and Disclosure Provisions as Required by Section 3802 of the Internal Revenue Service Restructuring and Reform Act of 1998, Volume II: Study of Disclosure Provisions Relating to Tax-Exempt Organizations* 38 n.97 (JCS-1-00, Jan. 28, 2000), available at www.house.gov/jct/s-1-00vol2.pdf (citing the IRS Exempt Organization Return Inventory and Classification System). In fiscal year 1999, the IRS revoked the exempt status of 97 organizations, of which 20 were exempt under § 501(c)(3); in fiscal year 1998, the IRS revoked the exemption of 97 organizations, 38 of which were described in § 501(c)(3); and in fiscal year 1997, the IRS revoked the exemption of 89 organizations, 38 described in § 501(c)(3). *Id*. at 27 n.56 (citing the IRS Audit Information Management System, Tables 41 and 42).

97 ACT, *supra* note 42, at 34, n.116 (page 120 of PDF).

98 Go to www.irs.gov/app/picklist/list/WrittenDeterminations.html. While the website makes it possible to sort determination letters by something called the UIL number, the letters are coded in obscure and

unhelpful ways. For example, UIL 501.06-02 begins helpfully, under Code section 501, but "06-02" means "Conduct of Business for Profit." This category is to be distinguished from "501.06-02 Conduct of Business for Profit." And what to make of "501.03-30 Organizational and Operational Tests" and "Profit v Not for Profit"? Moreover, categorical assignments do not seem to be made with great care. For example, Determination Letter 200634046 (Aug. 25, 2006), which involves a nonprofit corporation that lost its exemption on grounds of private inurement, is filed under "501.03-04 Unincorporated Associations." Of course, no single category is going to be helpful when the reasons for revocation are manifold.

[99] Leading practitioner and author Bruce Hopkins maintains a collection of citations, organized by the reason for denial or revocation (or for continued exemption), at www.nonprofitlawcenter.com/resources.php.

[100] In the case of private rulings, the redactors, in the first instance, are the requesting taxpayers themselves.

[101] In comments on the Joint Committee's 2000 disclosure study, the Independent Sector "strongly oppose[d]," among other JCT recommendations, those that would require the Service to make unredacted disclosure of written determinations and related file documents, closing agreements and audit results, exemption applications at the time of filing, and Forms 990-T (for its unrelated business taxable income) and 1120 (of any affiliated organizations of tax-exempt organizations). The Independent Sector supported, assuming technical refinement, giving greater flexibility for IRS information sharing with state charity regulators, a proposal enacted in the Pension Protection Act, as described below.

[102] Determination Letter 2010-52-022 (release date Oct. 5, 2010).

[103] For example, in a 2008 letter denying recognition of tax-exempt status under § 501(c)(3), the IRS set out "12 specific conditions" for recognizing an LLC under the organizational test of section 501(c)(3); while the letter cited no authority for these conditions, they appear in McCray & Ward L. Thomas, *Limited Liability Companies as Exempt Organizations—Update*, IRS Exempt Organizations Continuing Professional Education Text for FY 2001, at www.irs.gov/eo.

[104] ACT Report, *supra* note 42, at 3 (page 89 of the PDF).

[105] *Id*. at 33 (page 119 of the PDF).

[106] *Id*. at 3 (page 89 of the PDF).

[107] Rev. Rul. 2006-27; 2006-1 C.B. 915.

[108] I.R.C. § 501(q), added by the Pension Protection Act of 2006.

[109] See also Rev. Proc. 2011-9, *supra* note 91, at § 8.03 (citation omitted; emphasis added): "The Service may notify the appropriate State officials of a refusal to recognize an organization as tax-exempt under § 501(c)(3). The notice to the State officials may include a copy of a proposed or final adverse determination letter or ruling the Service issued to the organization. In addition, *upon request by the appropriate State official*, the Service may make available for inspection and copying the exemption application and other information relating to the Service's determination on exempt status." Separately, the IRS may disclose to appropriate state officials "the name, address, and identification number of any organization that has applied for recognition of exemption under section 501(c)(3)." *Id*. at 8.04. In calendar year 2009, the Service made 334 disclosures to state officials under § 6104(c). Internal Revenue Service, *Disclosure Report for Public Inspection Pursuant to Internal Revenue Code Section 6103(p)(3)(C)*, at 3, published by the Staff of the Joint Committee on Taxation (JCX-25-10) (April 15, 2010), available at www.jct.gov/publications.html?func=startdown&id=3680.

[110] See Notice of Prop. Rulemaking, *Disclosure of Information to State Officials Regarding Tax-Exempt Organizations*, REG-140108-08, 76 Fed. Reg. 13932-13937 (Mar. 15, 2011).

[111] Moreover, the comment letter states: "Those states that have entered into such agreements have limited their receipt of information to paper documents to avoid the substantial burdens of maintaining safeguards required for the maintenance of electronic data, since an audit of the statewide data center would be required. It is truly regrettable that [appropriate state officers] find themselves having to forego the efficiencies and other benefits of electronic information technology, especially as they strive to modernize their own systems."

[112] Staff of the U.S. Senate Finance Committee, *Discussion Draft for Hearing on Charity Oversight and Reform: Keeping Bad Things from Happening to Good Charities* (June 22, 2004), available at http://finance.senate.gov/imo/media/doc/062204stfdis.pdf.

[113] Available at www.nonprofitpanel.org. See also Part IV, below.

[114] These letters and, often, the responses, can be found in the press releases pages at http://grassley.senate.gov.

[115] See generally Transcript of Remarks of Dean Zerbe, CLE Program on Representing and Managing Tax-Exempt Organizations, Georgetown University Law Center, April 24, 2008, in 13 EO Tax J. 38, 39 (July/August 2008) (setting forth reflections on the congressional oversight process and goals by a former key tax aid to Senator Grassley).

[116] See "Review of Media-Based Ministries," Memorandum to Senator Grassley from Theresa Pattara and Sean Barnett 16-32 (Jan. 6, 2011), *available through* http://grassley.senate.gov/news/Article.cfm?customel_dataPageID_1502=30359. Attorney Marcus Owens, on behalf of one of the target churches, had written to Senators Baucus and Grassley on November 27, 2007: "If a [Senate] subpoena were issued, the Church and its members could be afforded certain confidentiality protections, which, like the privacy protections of section 6103, would reduce the likelihood of any public discourse regarding its religious beliefs." This letter is available on LEXIS in the Fedtax Library, Tax Notes Today file, as *Attorney Urges Grassley to Defer to IRS on Ministry Inquiry*, 2007 TNT 235-29, Dec. 6, 2007.

[117] Panel on the Nonprofit Sector, *Principles for Good Governance and Ethical Practice: A Guide for Charities and Foundations*, Principle #7 (Oct 2007), available at www.nonprofitpanel.org/report/principles/Principles_Guide.pdf. See generally John Tyler's forthcoming monograph, "Philanthropic 'Transparency': The Good, the Bad and the Useful" (Philanthropy Roundtable, Mar. 4, 2011 draft).

[118] Go to www.fordfoundation.org/about/governance. Annual reports are available at www.fordfoundation.org/about/annualreports, and its latest financial statement is available at www.fordfoundation.org/about/financials.

[119] Only Princeton's webpage survives. See www.princeton.edu/robertson.

[120] See the governance material posted at www.si.edu/about/regents.

[121] See, for example, the Getty's governance page at www.getty.edu/about/governance/. The posted material includes the Getty's mission statement, trust indenture, bylaws, board of trustees, board committees, trust officers and program directors, policies, financial information, annual and other reports, and the California attorney general's 2006 investigative report and the 2008 closure of the state's monitoring process.

[122] Available at http://philanthropy.com/free/update/update.htm.

[123] Available at www.nonprofitissues.com.

[124] Available at http://lawprofessors.typepad.com/nonprofit/.

[125] Available at www.charitygovernance.com.

[126] See the collections in Marion R. Fremont-Smith and Andras Kosaras, *Wrongdoing by Officers and Directors of Charities: a Survey of Press Reports 1995-2002*, 42 Exempt Org. Tax Rev. 25 (Oct. 2003); Marion R. Fremont-Smith, *Pillaging of Charitable Assets: Embezzlement and Fraud*, 46 Exempt Org. Tax Rev. 333 (Dec. 2004).

[127] See the Implementation Guide available at www.give.org. As explained in the preface to the Standards: "The overarching principle of the BBB Wise Giving Alliance Standards for Charity Accountability is full disclosure to donors and potential donors at the time of solicitation and thereafter. However, where indicated, the standards recommend ethical practices beyond the act of disclosure in order to ensure public confidence and encourage giving. As voluntary standards, they also go beyond the requirements of local, state and federal laws and regulations." Note that I have served on the board of the BBB Wise Giving Alliance since 2006.

[128] See, e.g., the Evangelical Council for Financial Accountability, "Seven Standards of Responsible Stewardship," www.ecfa.org/Standard 5, titled Transparency, reads: "Every member shall provide a copy of its current financial statements upon written request and provide other disclosures as the law may require. The financial statements required to comply with Standard 3 must be disclosed under this Standard. A member must provide a report, upon written request, including financial information on any specific project for which has sought or is seeking gifts."

[129] The most common reason for termination was failure to submit renewal information. Go to www.ecfa.org/FormerMembers.aspx.

[130] Council on Foundation, *Council Restores Membership of Getty Trust*, CFSource, Volume III, Issue 10 (May 2006), available at www.cof.org/Council/newsletter.cfm?ItemNumber=4285&navItemNumber=2499. Note that the press release is no longer posted on the Council's website!

[131] Panel on the Nonprofit Sector, *supra* note 117.

[132] See also Commissioner Shulman's Nov. 28, 2008 talk to Independent Sector: "We're . . . taking other proactive action like starting to check up on young exempt organizations to ensure that after a few years in operation they are in fact fulfilling an exempt purpose." Speech available at www.irs.gov/newsroom/article/0,,id=188567,00.html.

[133] Go to www.irs.gov/pub/irs-tege/governance_check_sheet.pdf.

[134] Go to www.irs.gov/pub/irs-tege/governance_guide_sheet.pdf. Anecdotally, at least some agents have begun their examinations by asking the organizations to fill out the governance check sheet.

[135] Go to www.irs.gov/charities/article/0,,id=216068,00.html.

[136] Go to www.irs.gov/charities/article/0,,id=178221,00.html.

[137] Go to www.irs.gov/pub/irs-tege/governance_practices.pdf

[138] Thomas Silk, *Good Governance Practices for 501(c)(3) Organizations: Should the IRS Become Further Involved?*, 57 Exempt Org. Tax Rev. 183, 183 (Aug. 2007), and, for different audiences, 107 J. Tax'n 45 (June 2007) and 10 Int'l J. Not-for Profit Law (Dec. 2007).

[139] Bonnie Brier, "The New Governance Project of the Exempt Organizations Division of the Internal Revenue Service," paper presented at the Nonprofit Forum (New York City, Feb. 24, 2010).

[140] Brody, note 39 *supra*.

Appendix

SUMMARY AND GOVERNANCE PAGES OF CORE FORM OF REDESIGNED FORM 990 (2007 DRAFT AND 2008 FINAL VERSIONS)

Form **990**

Return of Organization Exempt From Income Tax

Under section 501(c), 527, or 4947(a)(1) of the Internal Revenue Code (except black lung benefit trust or private foundation)

▶ The organization may have to use a copy of this return to satisfy state reporting requirements.

Department of the Treasury Internal Revenue Service (77)

OMB No. 1545-0047

20XX

Open to Public Inspection

A For the 20XX calendar year, or tax year beginning , 20XX, and ending , 20

B Check if applicable:
- ☐ Address change
- ☐ Name change
- ☐ Initial return
- ☐ Termination
- ☐ Amended return
- ☐ Application pending

Please use IRS label or print or type. See Specific Instructions.

C Name of organization

Number and street (or P.O. box if mail is not delivered to street address) | Room/suite

City or town, state or country, and ZIP + 4

D Employer identification number

E Telephone number ( )

F Name and address of Principal Officer:

G Website: ▶

H Enter amount of gross receipts $

I Accounting method: ☐ Cash ☐ Accrual ☐ Other ▶

J Books In care of Located at

K Organization type (check only one) ▶ ☐ 501(c) ( ) ◀ (insert no.) ☐ 4947(a)(1) or ☐ 527

Telephone number ( )

L Year of Formation:

M State of legal domicile ▶

Part I Summary

Activities & Governance

1 Briefly describe the organization's mission:

2 List the organization's three most significant activities and the activity codes (Part IX):
a Code b Code c Code

3 Enter the number of members of the governing body (Part III, line 1a) 3

4 Enter the number of independent members of the governing body (Part III, line 1b) 4

5 Enter the total number of employees (Part VIII, line 9a) 5

6 Enter the number of individuals receiving compensation in excess of $100,000 (Part II, line 2) 6

7 Enter the highest compensation amount reported on Part II, Section A (sum of columns D and E) 7

8a Enter officer, director, trustee, and other key employee compensation (Part V, line 5, column (B)) 8a

b Divide line 8a by line 17 %

9a Enter total gross unrelated business revenue from Part IV, line 14, column (C) 9a

b Enter net unrelated business taxable income from Form 990-T, line 34 b

10 Check this box ☐ if the organization discontinued its operations or disposed of more than 25% of its assets and attach Schedule N.

		Amount	% of Total
Revenues	11 Contributions and grants (Part IV, line 1g, column (A))		
	12 Program service revenue (Part IV, line 2g, column (A))		
	13 Membership dues and assessments (Part IV, line 3, column (A))		
	14 Investment income (Part IV lines 4, 5, 6, 8, 10d)		
	15 Other revenue (Part IV, lines, 3, 7, 9d, 11c, 12c, and 13e, column (A))		
	16 Total revenue add lines 11 through 15 (must equal Part IV, line 14, column (A))		100%
Expenses	17 Program service expense (Part V, line 24, column (B))		
	18 Management and general expenses (Part V, line 24, column (C))		
	19a Fundraising expenses (Part V, line 24, column (D))		
	19b Percentage of contributions (divide line 19a by line 11) %		
	20 Total expenses (must equal Part V, line 24, column (A))		100%
	21 Net income (line 16 minus line 20)		

		Beginning of Year	End of Year
Net Assets or Fund Balance	22 Total assets (Part VI, line 17)		
	23 Total liabilities (Part VI, line 27)		
	24a Net assets or fund balances line 22 minus line 23		
	24b Total expenses (line 20) as percentage of net assets (line 24a) %		

Gaming & Fundraising		(I) Gross Revenue	(II) Expenses	(III) Net to organization	(IV) Divide column (II) by column (I)
25	Gaming	(Schedule G, Part II, line 1 column (d))	(Schedule G, Part III, line 7)	(Schedule G, Part II, line 8)	%
26	Fundraising (other than gaming)	(Schedule G, Part I, line 1b column (ii) total)	(Schedule G, Part I, line 1b column (iv) total)	(Schedule G, Part I, line 1b column (v) total)	%

Form 990 (20XX) Page 4

Part III Statements Regarding Governance, Management, and Financial Reporting

			Yes	No
1a	Enter the number of members of the governing body 1a			
b	Enter the number of independent members of the governing body 1b			
2	Did the organization make any significant changes to its organizing or governing documents? If "Yes", briefly describe these changes.	2		
3a	Does the organization have a written conflict of interest policy?	3a		
b	If "Yes," how many transactions did the organization review under this policy and related procedures during the year? 3b			
4	Does the organization have a written whistleblower policy?	4		
5	Does the organization have a written document retention and destruction policy?	5		
6	Does the organization contemporaneously document the meetings of the governing body and related committees through the preparation of minutes or other similar documentation?	6		
7a	Does the organization have local chapters, branches or affiliates?	7a		
b	If yes, does the organization have written policies and procedures governing the activities of such chapters, affiliates and branches to ensure their operations are consistent with the organization's?	7b		
8	Does an officer, director, trustee, employee or volunteer prepare the organization's financial statements? Indicate whether an independent accountant provides any of the following services: Compilation ☐ Review ☐ Audit ☐	8		
9	Does the organization have an audit committee?	9		
10	Did the organization's governing body review this Form 990 before it was filed?	10		

11 How do you make the following available to the public? Check all that apply.

Organizing/Governing Documents	☐ n/a	☐ website	☐ other website	☐ office	☐ other	_____
Conflict of Interest Policy	☐ n/a	☐ website	☐ other website	☐ office	☐ other	_____
Form 990	☐ n/a	☐ website	☐ other website	☐ office	☐ other	_____
Form 990-T	☐ n/a	☐ website	☐ other website	☐ office	☐ other	_____
Financial Statements	☐ n/a	☐ website	☐ other website	☐ office	☐ other	_____
Audit Report	☐ n/a	☐ website	☐ other website	☐ office	☐ other	_____

12 List the states with which a copy of this return is filed:

Form **990**

Department of the Treasury
Internal Revenue Service

Return of Organization Exempt From Income Tax

Under section 501(c), 527, or 4947(a)(1) of the Internal Revenue Code (except black lung benefit trust or private foundation)

► The organization may have to use a copy of this return to satisfy state reporting requirements.

OMB No. 1545-0047

2008

Open to Public Inspection

A For the 2008 calendar year, or tax year beginning , 2008, and ending , 20

B Check if applicable:
- ☐ Address change
- ☐ Name change
- ☐ Initial return
- ☐ Termination
- ☐ Amended return
- ☐ Application pending

Please use IRS label or print or type. See Specific Instructions.

C Name of organization

Doing Business As

Number and street (or P.O. box if mail is not delivered to street address) | Room/suite

City or town, state or country, and ZIP + 4

D Employer identification number

E Telephone number ( )

G Gross receipts $

F Name and address of principal officer:

H(a) Is this a group return for affiliates? ☐ Yes ☐ No

H(b) Are all affiliates included? ☐ Yes ☐ No
If "No," attach a list. (see instructions)

H(c) Group exemption number ►

I Tax-exempt status: ☐ 501(c) ( )◄ (insert no.) ☐ 4947(a)(1) or ☐ 527

J Website: ►

K Type of organization: ☐ Corporation ☐ Trust ☐ Association ☐ Other ►

L Year of formation:

M State of legal domicile:

Part I Summary

Activities & Governance

1 Briefly describe the organization's mission or most significant activities: ----------

2 Check this box ► ☐ if the organization discontinued its operations or disposed of more than 25% of its assets.

3	Number of voting members of the governing body (Part VI, line 1a)	3	
4	Number of independent voting members of the governing body (Part VI, line 1b)	4	
5	Total number of employees (Part V, line 2a)	5	
6	Total number of volunteers (estimate if necessary)	6	
7a	Total gross unrelated business revenue from Part VIII, line 12, column (C)	7a	
b	Net unrelated business taxable income from Form 990-T, line 34	7b	

		Prior Year	Current Year
Revenue	8 Contributions and grants (Part VIII, line 1h)		
	9 Program service revenue (Part VIII, line 2g)		
	10 Investment income (Part VIII, column (A), lines 3, 4, and 7d)		
	11 Other revenue (Part VIII, column (A), lines 5, 6d, 8c, 9c, 10c, and 11e)		
	12 Total revenue—add lines 8 through 11 (must equal Part VIII, column (A), line 12)		
Expenses	13 Grants and similar amounts paid (Part IX, column (A), lines 1–3)		
	14 Benefits paid to or for members (Part IX, column (A), line 4)		
	15 Salaries, other compensation, employee benefits (Part IX, column (A), lines 5–10)		
	16a Professional fundraising fees (Part IX, column (A), line 11e)		
	b Total fundraising expenses (Part IX, column (D), line 25) ► ----------		
	17 Other expenses (Part IX, column (A), lines 11a–11d, 11f–24f)		
	18 Total expenses. Add lines 13–17 (must equal Part IX, column (A), line 25)		
	19 Revenue less expenses. Subtract line 18 from line 12		

		Beginning of Year	End of Year
Net Assets or Fund Balances	20 Total assets (Part X, line 16)		
	21 Total liabilities (Part X, line 26)		
	22 Net assets or fund balances. Subtract line 21 from line 20		

Part II Signature Block

Under penalties of perjury, I declare that I have examined this return, including accompanying schedules and statements, and to the best of my knowledge and belief, it is true, correct, and complete. Declaration of preparer (other than officer) is based on all information of which preparer has any knowledge.

Sign Here

► Signature of officer — Date

► Type or print name and title

Paid Preparer's Use Only

Preparer's signature | Date | Check if self-employed ► ☐ | Preparer's identifying number (see instructions)

Firm's name (or yours if self-employed), address, and ZIP + 4 | EIN ► | Phone no. ► ( )

May the IRS discuss this return with the preparer shown above? (see instructions) ☐ Yes ☐ No

Part VI **Governance, Management, and Disclosure** *(Sections A, B, and C request information about policies not required by the Internal Revenue Code.)*

Section A. Governing Body and Management

			Yes	No
	For each "Yes" response to lines 2–7b below, and for a "No" response to lines 8 or 9b below, describe the circumstances, processes, or changes in Schedule O. See instructions.			
1a	Enter the number of voting members of the governing body	1a		
b	Enter the number of voting members that are independent	1b		
2	Did any officer, director, trustee, or key employee have a family relationship or a business relationship with any other officer, director, trustee, or key employee?	2		
3	Did the organization delegate control over management duties customarily performed by or under the direct supervision of officers, directors or trustees, or key employees to a management company or other person?	3		
4	Did the organization make any significant changes to its organizational documents since the prior Form 990 was filed?	4		
5	Did the organization become aware during the year of a material diversion of the organization's assets?	5		
6	Does the organization have members or stockholders?	6		
7a	Does the organization have members, stockholders, or other persons who may elect one or more members of the governing body?	7a		
b	Are any decisions of the governing body subject to approval by members, stockholders, or other persons?	7b		
8	Did the organization contemporaneously document the meetings held or written actions undertaken during the year by the following:			
a	The governing body?	8a		
b	Each committee with authority to act on behalf of the governing body?	8b		
9a	Does the organization have local chapters, branches, or affiliates?	9a		
b	If "Yes," does the organization have written policies and procedures governing the activities of such chapters, affiliates, and branches to ensure their operations are consistent with those of the organization?	9b		
10	Was a copy of the Form 990 provided to the organization's governing body before it was filed? All organizations must describe in Schedule O the process, if any, the organization uses to review the Form 990	10		
11	Is there any officer, director or trustee, or key employee listed in Part VII, Section A, who cannot be reached at the organization's mailing address? *If "Yes," provide the names and addresses in Schedule O*	11		

Section B. Policies

			Yes	No
12a	Does the organization have a written conflict of interest policy? *If "No," go to line 13*	12a		
b	Are officers, directors or trustees, and key employees required to disclose annually interests that could give rise to conflicts?	12b		
c	Does the organization regularly and consistently monitor and enforce compliance with the policy? *If "Yes," describe in Schedule O how this is done*	12c		
13	Does the organization have a written whistleblower policy?	13		
14	Does the organization have a written document retention and destruction policy?	14		
15	Did the process for determining compensation of the following persons include a review and approval by independent persons, comparability data, and contemporaneous substantiation of the deliberation and decision:			
a	The organization's CEO, Executive Director, or top management official?	15a		
b	Other officers or key employees of the organization?	15b		
	Describe the process in Schedule O. (see instructions)			
16a	Did the organization invest in, contribute assets to, or participate in a joint venture or similar arrangement with a taxable entity during the year?	16a		
b	If "Yes," has the organization adopted a written policy or procedure requiring the organization to evaluate its participation in joint venture arrangements under applicable federal tax law, and taken steps to safeguard the organization's exempt status with respect to such arrangements?	16b		

Section C. Disclosure

17 List the states with which a copy of this Form 990 is required to be filed ▶

18 Section 6104 requires an organization to make its Forms 1023 (or 1024 if applicable), 990, and 990-T (501(c)(3)s only) available for public inspection. Indicate how you make these available. Check all that apply.

☐ Own website ☐ Another's website ☐ Upon request

19 Describe in Schedule O whether (and if so, how), the organization makes its governing documents, conflict of interest policy, and financial statements available to the public.

20 State the name, physical address, and telephone number of the person who possesses the books and records of the organization: ▶

5

Appendix

Conference Program

New Perspectives on Tax Administration: An IRS-TPC Research Conference
Urban Institute, 2100 M Street, N.W., Washington, DC • June 22, 2011

Program

8:30 – 9:00 Check-in, Continental Breakfast

9:00 – 9:05 Welcome and Introductions
Eric Toder (Urban Institute) and Rosemary Marcuss (IRS:RAS)

9:05 – 9:15 Opening Remarks
Douglas Shulman (IRS Commissioner)

9:20 – 10:50 Session 1: Impact of Service on Compliance

Moderator: *Melissa Vigil (IRS: RAS)*

- Taxpayer Information Services: Reporting Effects of Information Quality
  Michael Jones (Bridgewater State University), Michael McKee (Appalachian State University), and Christian A. Vossler (University of Tennessee)
- Multi City Study of the Effect of Assistance on Compliance
  Tiffanie N. Bruch, David C. Cico, and Saima S. Mehmood (IRS: W&I Research& Analysis)
- Predicting Aggregate Taxpayer Compliance Behavior
  Brian Erard (B. Erard & Associates), Alan Plumley (IRS: RAS), and Derek Snaidauf (IBM Business Consulting)

Discussant: *Marsha Blumenthal (University of St. Thomas)*

10:50 – 11:00 Break

11:00 – 12:00 Session 2: Individual Compliance Behavior

Moderator: *Elaine Maag (Urban Institute)*

- On the External Validity of Tax Compliance Experiments
  James Alm (Tulane Univ.), Kim M. Bloomquist (IRS: RAS), and Michael McKee (Appalachian State Univ.)
- Reconsidering the Deterrence Paradigm of Tax Compliance
  Mark D. Phillips (Univ. of Chicago, Ph.D. student)

Discussant: *Charles Christian (Arizona State University, retired)*

12:00 – 1:00 Lunch and Keynote Speaker

David Walker (Founder, President and CEO of the Comeback America Initiative)

1:00 – 2:30 Session 3: Estimating the Tax Gap

Moderator*: Eric Toder (Urban Institute)*

- The Practicality of a Top Down Approach to Estimating the Direct Tax Gap
  Marcus Rubin (HMRC, UK)
- The Individual Income Reporting Gap: What We See and What We Don't
  Brian Erard (B. Erard & Associates) and Jonathan Feinstein (Yale University)
- The Federal Tax Position of Persons Who Were Not Reported on Filed Tax Returns in 2005
  Joshua Lawrence, Michael Udell, and Tiffany Young (Ernst & Young LLP)

Discussants*: Katherine Baer and Emily Lin (Office of Tax Analysis, U.S. Department of the Treasury)*

2:30 – 2:45 Break

2:45 – 4:15 Session 4: New Disclosure and Regulation Issues

Moderator*: Barry Johnson (IRS: RAS)*

- Use of Assets in Large and Mid-Size Corporations: An Econometric Analysis of the Manufacturing Sector Using Data From Form M-3
  George Contos and John Guyton (IRS: RAS), Jean LaVelle and Deborah Myers (IRS, LB&I)
- Sunshine and Shadows on Charity Governance: Public Disclosure as an IRS Regulatory Tool
  Evelyn Brody (Illinois Institute of Technology)
- Who Needs Credit at Tax Time and Why: A Look at Refund Anticipation Loans and Refund Anticipation Checks
  Rachel Brash, Jessica F. Compton, Nancy Pindus, C. Eugene Steuerle, and Brett Theodos, (Urban Institute), and Karen Masken (IRS:RAS)

Discussant*: Joe Cordes (George Washington University)*

4:15 – 4:30 Wrap-up

Janice Hedemann (Conference Chair, IRS:RAS)