Statistics of Income

2008 SOI Paper Series

Using Audit Data To Estimate Taxpayer Reporting Error in the Statistics of Income Division's Individual Tax Return Sample

by Kimberly Henry

INTERNAL REVENUE SERVICE

Using Audit Data To Estimate Taxpayer Reporting Error in the Statistics of Income Division's Individual Tax Return Sample

Kimberly Henry, Internal Revenue Service

The National Research Program (NRP) was implemented by the Internal Revenue Service (IRS) for individual tax returns in Tax Year 2001 to support tax research by selecting large random samples of tax returns to be audited (Brown and Mazur, 2003). The resulting NRP data are used here to estimate taxpayer reporting error in national-level totals of eight variables estimated from the Statistics of Income (SOI) Division's Form 1040 sample. Since SOI's individual sample data are based on preaudit information, estimates produced from it are affected by taxpayer misreporting. Both samples are large stratified Bernoulli samples, with different strata definitions and sampling rates. Only a small number of returns (433) were in both.

All eight deduction-related variables examined were overstated by taxpayers such that SOI's estimates of each variable's national-level total have a positive bias. To examine the extent of this, four alternative analyses are examined: the differences in estimated totals from both samples, two ratio-based adjustments to the SOI estimates, and post-stratified adjustments to the NRP estimates. The bias and variance of each method's estimated true total are used to evaluate the alternatives and determine the impact of the reporting error on national-level estimates. Error estimates using only NRP data are also provided to compare the estimates examined here to similar ones the IRS produces.

▶ Taxpayer Reporting Error

The taxpayer reporting error is defined as the difference between SOI's values edited for statistical purposes, which are based on taxpayers' originally reported values, and the corresponding values determined by NRP auditors. Thus, the audits are regarded as yielding the true values. This is different from other IRS taxpayer error studies (e.g., Bloomquist, 2004 and Plumley, 2005) that attempt to account for misreporting undetected by the NRP auditors; only taxpayer reporting error that was detected by the auditors is quantified here.

The "incentive" for taxpayers to alter their tax liabilities can lead to intentional misreporting, since lower reported amounts of income-related variables (particularly unreported income) and higher amounts of adjustment- and deduction-related variables contribute to lower amounts of tax owed. While it is legal for certain taxpayers to use itemized deductions to lower their amounts of income that is subject to tax, there are taxpayers who illegally (whether intentionally or not) inflate their reported deductions. Intentional and illegal misreporting of tax information is called tax evasion. However, unintentional misreporting may also occur due to a complex tax system, including the tax forms and laws, or inadvertent mistakes. This can happen particularly among less informed taxpayers (Slemrod and Bakija, 2004).

The most obvious effect of taxpayer misreporting is that taxpayers do not pay the amount of taxes they owe. In general, by understating income and overstating deductions, taxpayers pay less tax than they should. Measuring the amount of tax paid is relatively simple, but it is much more difficult to determine how much should have been paid. One periodic IRS estimate, the gross tax gap (the amount of true tax liability for a given tax year that is not paid voluntarily and on time (IRS, 2006)) was $345 billion for all types of 2001 tax.

▶ Description of the Data

An individual was required to file a Tax Year 2001 tax return based on gross income, marital status, age, and, to a lesser extent, dependency and blindness (Parisi, 2003). Gross income is all income received in

This paper was originally presented at the Second International Total Survey Error Workshop held in Raleigh, NC, on June 1-4, 2008.

the form of money, property, and investment services not expressly exempt from being taxed.

The data come from two separate IRS samples. The frame for both was the Calendar Year 2002 IRS Individual Master File (IMF). Both included Form 1040 (the basic individual income tax return), Form 1040A (a shortened version of Form 1040), and Form 1040EZ (the income tax return for single and joint filers with no dependents). Both samples included original filings, the first returns that are filed by U.S. citizens and residents to IRS and electronically keyed by IRS transcribers. Both samples excluded returns selected for operational audits prior to their sample selection processes and other filings, such as amended or duplicate returns. However, amended return information was taken into account in the audits.

Each sample included returns that the other regarded as out-of-scope. SOI's sample included certain "Non-Master File tax returns" that were not on the IMF due to limits on the number of digits allowed for monetary fields, certain returns filed in 2002 for tax years prior to 2001, and partial-year returns (e.g., ones filed quarterly, consolidating the partial-year information into one record). Civilian and military taxpayers in non-U.S. states, possessions, or territories were also excluded from NRP's sample and included in SOI's.

The SOI Sample Design

Stratification for SOI's sample used the following categories: (1) nontaxable returns with adjusted gross income/expanded income of $200,000 or more; (2) high combined total business receipts of $50,000,000 or more; and (3) presence/absence of special forms or schedules (Form 2555, Form 1116, Form 1040 Schedule C, and Form 1040 Schedule F). Stratum assignment was based on the order in which a return met one of these categories, e.g., if a return met (1) and (2), it fell into (1)'s strata. Within category (3), stratification used size of indexed total gross positive/negative income and an indicator of the return's "usefulness" for tax policy modeling purposes (Walker and Testa, 2003). Each return in the target population was assigned to a stratum based on these criteria.

The sample had two parts. Within each stratum, a .05-percent stratified simple random sample of 65,076 returns was selected (Weber, 2004). For other returns, a Bernoulli sample was also independently selected from each stratum, with sampling rates from 0.05 percent to 100 percent. SOI selected 191,975 returns from 130,571,421. Data capture and cleaning procedures resulted in a sample of 191,809 returns and an estimated population of 130,255,237.

The NRP Sample Design

A Bernoulli sample was also selected independently from each stratum for the NRP sample. The first level of NRP strata was the IRS division having jurisdiction for the returns, between the Wage and Investment (W&I) and Small Business-Self Employed (SBSE) Divisions. W&I was responsible for 1040 returns where most income was ordinary income (e.g., from taxpayers' salaries and wages), while SBSE was concerned with returns where the majority of taxpayer income was related to a business or farm (as reported on a Schedule C or F attached to the Form 1040). Further stratification was achieved using a combination of 1040 Form Type, size of Total Positive Income, Adjusted Gross Income, or Total Gross Receipts from a business/farm, and presence/absence of Schedules C and F. NRP selected 45,740 returns from a population of 125,811,411. Data capture and cleaning resulted in 44,768 returns from an estimated 125,790,458.

The sample and estimated population counts for particular taxpayer characteristics from both samples are given in Table 1. Despite large differences in sample counts, the estimated population counts are close.

Variables of Interest

Eight tax variables were chosen using four criteria: (1) the variables were reported by a relatively large number of taxpayers in both samples; (2) they were less susceptible than income and tax-related variables to being undetected by auditors, since the legal burden of proof is on the taxpayers to establish their accuracy; (3) they were of subject-matter interest, i.e., previous research had demonstrated they are misreported; and

Table 1. Number (#) of Sample and Estimated Number of Population Returns, by Characteristic and Sample

Characteristic	**SOI Sample**		**NRP Sample**	
	# Sample Returns [a]	*Estimated # Population Returns* [a]	*# Sample Returns*	*Estimated # Population Returns*
1040A returns	12,524	23,538,694	2,192	23,297,612
1040EZ returns	7,775	15,641,014	1,292	14,817,862
1040 returns	159,420	90,799,756	41,284	87,675,485
Schedule As	119,324	44,822,874	24,371	44,241,224
Electronically filed returns	32,012	46,848,690	11,037	46,916,186
Returns that used a paid preparer [b]	133,008	72,219,936	31,392	70,254,194
Total	179,719	129,979,464	44,768	125,790,458

a: Excludes internationsl returns.

b: Excludes returns with a paid preparer SSN/EIN provided (N/A in SOI sample), but associated preparer code was null.

(4) they were less affected by differences in the two samples' target populations.

Descriptive Tables

Table 2 shows the name, a brief description, and subject-matter interest for each variable. The number of errors and size of error rankings are from Bennett's (2005) initial assessment using the NRP data (his rankings excluded calculated variables, e.g., taxes). Table 3 shows the population counts and variable totals estimated from SOI's sample, before and after international returns were removed, and the resulting differences. "International" returns here were tax returns with a foreign address or a Form 2555 attached, indicating foreign income. SOI totals without international returns are used in all subsequent tables to avoid confounding the differences in Table 3 with the estimated taxpayer reporting error and make the samples more comparable. Despite this, the two samples' estimated population totals are still different: 129,773,275 from SOI's sample and 125,790,458 from NRP's, motivating the use of alternative adjustment methods. Table 4 shows the sample and estimated number of population returns with nonzero values (where NRP counts use auditor-determined values) for each variable, from both samples. The associated variable totals are examined later.

Despite differences between the numbers of sample returns in Table 5, the estimated population sizes are relatively close, with the exception of Cash Contributions (where SOI's estimate is larger by 4,816,401 returns). They are closest for State and local Taxes (where the NRP estimate is higher by 389,886 returns). This variable and Total Adjustments were the only ones where the NRP estimated number of returns is larger than SOI's; all others are smaller.

▶ Methodology Behind the Error Estimates

General Notation

For S_1 denoting SOI's sample and S_2 denoting NRP's sample, let

x_i be the taxpayer-reported value for a given variable on the IMF, for tax return $i \in S_1$ or $i \in S_2$;

y_i the same variable's value edited by SOI for return $i \in S_1$;

μ_i the (true) value determined by an auditor for return $i \in S_2$.

The x_i values y_i and are distinguished separately since they are not equal if there are processing errors (not from different IRS and SOI data editing rules, which is true for these variables) in the frame data. These errors in x_i are also corrected by auditors such that the difference between y_i and μ_i is assumed to be the taxpayer reporting error.

Table 2. Variable Name, Description, and Subject-Matter Interest, by Variable of Interest

Variable Name	*Location on 2001 Form(s)*	*Variable Description*	*Subject-Matter Interest* [a]
Cash Contributions	Line 15, Schedule A, Form 1040	Monetary contributions to certain organizations.	Highest number of errors; fifth highest in error amount ($13.1 billion).
Noncash Contributions	Line 16, Schedule A, Form 1040	Nonmonetary contributions to certain organizations.	Seventh highest number of errors.
Total Adjustments, Without SE Tax Adjustment	Lines 23-32 plus attachments, Form 1040	Various adjustment components (IRS 2003b) subtracted from AGI,[b] excluding that for Self-Employment (SE) taxes.	Underreporting SE taxes leads to incorrectly interpreting Total Adjustments as underreported; all other components are overstated. [c]
Total Taxes Deducted	Sum of Lines 5 to 8, Schedule A, Form 1040	Total of State and Local Taxes, Real Estate Taxes, and Personal Property/Other Taxes.	The total is included to examine the combined error effect from separate components.
State and Local Income Taxes	Line 5, Schedule A, Form 1040	Amount of deductible state and local taxes paid.	Error should be lowest; third-party information is required for this deduction.
Real Estate Taxes Paid	Line 6, Schedule A, Form 1040	Amount of deductible nonbusiness -related real estate taxes paid.	Fourth highest number of errors.
Other Taxes/Personal Property Taxes	Lines 7 and 8, Schedule A, Form 1040	Amount of deductible other nonbusiness-related taxes paid, including property taxes.	Eighth highest number of errors.
Exemptions	Lines 6, 38, Form 1040; Line 26 Form 1040A; Line 5, Worksheet F, Form 1040EZ	Total of all exemption amounts; a $2,900 deduction was allowed for each qualified exemption if AGI was less than $99,725.	Third highest number of errors.

a: Rankings exclude calculated items.

b: AGI = Adjusted Gross Income

c: Based on prior research in the IRS Office of Research

Table 3. Estimated Number of Population Returns and Estimated Variable Totals (in Thousands of Dollars), With and Without International (Int'l) Returns, and Resulting Differences

Variable	**Estimated Number of Population Returns**[a]			**Estimated Variable Total ($ 1,000's)**		
	Full SOI Sample Estimate	*Estimate Without Int'l Returns*	*Estimated Difference*	*Full SOI Sample Estimate*	*Estimate Without Int'l Returns*	*Estimated Difference*
Cash Contributions	37,855,184	37,792,234	62,950	104,747,174	104,439,939	307,234
Noncash Contributions	22,585,276	22,552,644	32,632	37,997,546	37,888,487	109,059
Total Adjustments, Without SE Tax Adjustment	13,612,165	13,559,691	52,474	42,437,809	42,052,057	385,752
Total Taxes Deducted	43,797,188	43,722,001	75,187	307,974,817	307,172,690	802,127
State and Local Taxes	37,037,062	36,988,695	48,367	196,430,907	195,868,643	562,264
Real Estate Taxes	38,716,754	38,655,137	61,617	101,853,670	101,660,730	192,940
Other Taxes/Personal Property Taxes	22,633,437	22,613,280	20,157	9,690,240	9,643,317	46,923
Exemptions	118,273,285	117,506,894	766,391	727,554,990	721,814,512	5,740,479

a: Number with nonzero variable amounts.

Table 4. Number of Sample Returns and Estimated Number of Population Returns with Nonzero Variable Amounts, SOI and NRP Samples, by Variable of Interest

Variable	*Number in SOI Sample* [a]	*SOI Population Size Estimate* [a]	*Number in NRP Sample* [b]	*NRP Population Size Estimate* [b]
Cash Contributions	103,385	37,792,234	19,400	32,975,833
Noncash Contributions	54,147	22,552,644	10,130	18,157,742
Total Adjustments, Without SE Tax Adjustment	40,914	13,559,691	16,593	17,679,580
Total Taxes Deducted	110,591	43,722,001	23,696	42,981,469
State and Local Taxes	96,382	36,988,695	19,441	36,186,830
Real Estate Taxes	103,045	38,655,137	21,433	37,378,581
Other Taxes/Personal Property Taxes	55,307	22,613,280	31,765	20,435,918
Exemptions	107,506	117,506,894	39,236	113,807,787

a: Number with nonzero SOI-edited variable amounts, excluding international returns.

b: Number with nonzero auditor-determined amounts.

The Difference Between the Two Samples' Estimates

The SOI sample-based total for variable of interest y is

$$\begin{aligned}\hat{Y} &= \sum_{i \in S_1} w_i y_i \\ &= \sum_h \frac{N_h}{n_h} \hat{y}_h,\end{aligned} \qquad (4.1)$$

where $w_i = \frac{N_h}{n_h}$ is the survey weight for each unit in stratum h, $h = 1, \ldots, H (= 216)$; N_h and n_h are the realized population and sample sizes in stratum h (i.e., conditioning on the numbers obtained at the completion of the Bernoulli sampling procedure); and $\hat{y}_h$ is the unweighted stratum h total of y.

Using the conditional strata sample and population sizes, estimators of totals and their variances reduce to those of simple random sampling within each stratum (Sarndal et al., 1992 and Valliant and Cassady, 1998). The variance estimate of (4.1) is thus:

$$var\left(\hat{Y}\right) = \sum_{h=1}^{H} \frac{N_h^2}{\left(\sum_h N_h\right)^2} \left(1 - \frac{n_h}{N_h}\right) s_h^2, \qquad (4.2)$$

where s_h^2 is the stratum sample variance.

For $l = 1, \ldots, L$ denoting the NRP sample strata, the estimated total of auditor-determined values is

$$\begin{aligned}\hat{M}_1 &= \sum_{i \in S_2} \tilde{w}_i \mu_i \\ &= \sum_{l=1}^{L} \frac{N_l}{n_l} \hat{\mu}_l\end{aligned}, \qquad (4.3)$$

where the weight $\tilde{w}_i$ is the ratio of realized population and sample sizes for all units in stratum l (N_l and n_l), and $\hat{\mu}_l$ is the stratum l total of auditor-determined values. Similar to (4.2), the variance of (4.3) is

$$var\left(\hat{M}_1\right) = \sum_{l=1}^{L} \frac{N_l^2}{\left(\sum_l N_l\right)^2} \left(1 - \frac{n_l}{N_l}\right) s_l^2 \qquad (4.4)$$

From (4.1) and (4.3), the aggregate-level estimate of the error in $\hat{Y}$ is

$$\hat{D}_1 = \hat{Y} - \hat{M}_1 \,. \qquad (4.5)$$

This approach is considered from the perspective of an external data user with access only to the two separate sample's estimates.

Combined Ratio-Adjusted Estimates

Estimator (4.5) does not account for any differences between the two samples, despite removing international tax returns from the SOI sample total in (4.1). As a result, alternative estimators are considered. First, a combined ratio adjustment to (4.1) for the national-level taxpayer reporting error produces:

$$\hat{M}_2 = \frac{\hat{M}_1}{\hat{X}}\hat{Y}, \quad = \hat{a}\hat{Y} \tag{4.6}$$

where $\hat{X} = \sum_{i \in S_2} \tilde{w}_i x_i$ is the total of original taxpayer reported values, estimated from the NRP. The adjustment factor $\hat{a}$ is a national-level ratio of the weighted total of auditor-determined values to the weighted total of the taxpayer-reported values (using the NRP sample weights). That is,

$\hat{a} > 1$ when taxpayers underreport a tax variable's amount;

$\hat{a} = 1$ indicates no change;

$\hat{a} < 1$ indicates taxpayers overstating it.

For the variables of interest, $\hat{a}$ ranged from .786 (for Other Taxes/Personal Property Taxes) to .996 (State and Local Income Taxes), indicating that taxpayers overstated these deductions.

Using a Taylor series approximation, the variance of $\hat{M}_2$ is

$$Var\left[l\left(\hat{M}_2 - M\right)\right] \approx \frac{1}{\hat{X}^2}\begin{bmatrix} \hat{Y}^2 Var\left(\hat{M}_1\right) + \frac{\hat{M}_1^2 \hat{Y}^2}{\hat{X}^2} Var\left(\hat{X}\right) \\ + \hat{M}_1^2 Var\left(\hat{Y}\right) + 2\frac{\hat{M}_1 \hat{Y}^2}{\hat{X}} Cov\left(\hat{M}_1, \hat{X}\right) \\ + 2\hat{M}_1 \hat{Y} Cov\left(\hat{M}_1, \hat{Y}\right) + 2\frac{\hat{M}_1^2 \hat{Y}}{\hat{X}} Cov\left(\hat{X}, \hat{Y}\right) \end{bmatrix}. \tag{4.7}$$

Using a linear substitute approximation (Wolter Sec. 6.5, Woodruff, 1971) to $\hat{a}\hat{Y}$ avoids calculating all the variance and covariance terms in (4.8). This leads to the following approximate variance estimate of $\hat{M}_2$:

$$var(\hat{M}_2) \approx \sum_{i \in S_1} var(w_i u_i) + \sum_{i \in S_2} var(\tilde{w}_i \tilde{u}_i), \tag{4.8}$$

where $u_i = \frac{\hat{M}_1 y_i}{\hat{X}}$ and $\tilde{u}_i = \frac{\hat{Y}}{\hat{X}}\left(\mu_i - \hat{M}_1 x_i\right)$. Using estimator (4.6) leads to the following error estimate:

$$\hat{D}_2 = \hat{Y} - \hat{M}_2 . \tag{4.9}$$

Separate Ratio Adjustments

This analysis method applies finer-level adjustments to the SOI data to account for the taxpayer reporting error. The setup is similar conceptually to cell-based adjustments used for survey nonresponse (e.g., Kalton and Kaspryzyk, 1986; Kalton and Maligalig, 1991; and Oh and Scheuren, 1983) and simply an extension of the ratio adjustment in (4.5). Here, the taxpayer misreporting ratio adjustment (as detected by auditors in the NRP sample) is applied to SOI's weighted strata totals. SOI's strata, which were assigned to each return in the NRP sample, have definitions that incorporate the taxpayer's size of income and particular attachments to the return, which is indirectly related to whether W&I or SBSE had jurisdiction over the tax returns. Thus, it is a reasonable assumption that taxpayers within the same stratum (as defined in SOI's sample) but residing in different samples have the same reporting behavior (as SOI's sample is representative of the tax filing population). Some of the 216 SOI strata across income categories were collapsed to ensure enough NRP returns within each one.

The SOI total in (4.1), written as the sum over the strata totals, is:

$$\hat{Y} = \sum_h \hat{Y}_h . \tag{4.10}$$

The estimate for the total of the variable y, adjusted for taxpayer reporting error, is the sum of adjusted SOI strata totals:

$$\hat{M}_3 = \sum_h \hat{a}_h \hat{Y}_h \tag{4.11}$$

where $\hat{a}_h = \sum_{i \in S_{2h}} \tilde{w}_i \mu_i \Big/ \sum_{i \in S_{2h}} \tilde{w}_i x_i$ is the adjustment factor for all units in stratum h.

The ratio $\hat{a}_h$ is calculated, from the NRP sample, as the ratio of the weighted stratum total of auditor-determined values to the weighted stratum total of originally reported taxpayers' values. It has the same interpretation as $\hat{a}$ in (4.6), just within each stratum defined for SOI's sample.

Estimator (4.11) is simply a stratified ratio estimator, despite $\hat{a}_h$ being calculated from a separate sample (this just determines its properties). Thus, the linearization and linear substitute variance estimates can be applied. Using a Taylor series approximation, the variance of $\hat{M}_3$ is

$$\sum_h Var\left[l\left(\hat{a}_h\hat{Y}_h - a_hY_h\right)\right], \tag{4.12}$$

where the components in (4.12) are identical to those in (4.7), just specified within each stratum h. A much simpler variance estimate, using a linear substitute approximation to $\hat{a}_h\hat{Y}_h$, leads to the following approximate variance estimate of $\hat{M}_3$:

$$\begin{aligned} var(\hat{M}_3) \approx &\sum_h var\left[\sum_{i\in S_{1h}} w_i u_i\right] \\ &+\sum_h var\left[\sum_{i\in S_{2h}} \tilde{w}_i\tilde{u}_i\right], \end{aligned} \tag{4.13}$$

where $u_i = \dfrac{\hat{M}_{1h}y_i}{\hat{X}_h}$ and $\tilde{u}_i = \dfrac{\hat{Y}_h}{\hat{X}_h}\left(\mu_i - \hat{M}_{1h}x_i\right)$.

From (4.9) and (4.10), the taxpayer reporting error estimate involves the following difference:

$$\hat{D}_3 = \hat{Y} - \hat{M}_3. \tag{4.14}$$

Poststratification Adjustments

A poststratification (PS) adjustment is also considered to overcome differences in the two target populations. The PS estimator of the true total of taxpayer values considered is

$$\hat{M}_4 = \frac{\sum_h N_h}{\sum_l N_l}\hat{M}_1. \tag{4.15}$$

Here, the ratio of the known SOI and NRP population totals (approximately 1.03) is applied to the NRP sample-based total of auditor-determined values. Since this ratio involves known population counts, the variance of (4.15) is simply

$$var(\hat{M}_4) = \left(\frac{\sum_h N_h}{\sum_h N_l}\right)^2 var(\hat{M}_1), \tag{4.16}$$

where $var(\hat{M}_1)$ is given in (4.4). The associated error estimate is

$$\hat{D}_4 = \hat{Y} - \hat{M}_4. \tag{4.17}$$

▶ Comparing Alternative Error Estimates

Ultimately, the desired measures are the taxpayer reporting error estimates, relative to M, the true total:

$$\%\,\text{Rel Err(wrt M)} = 100 \times \frac{\hat{Y} - M}{M}. \tag{4.18}$$

However, the true total M is unknown and estimated by four alternatives, given in (4.3), (4.6), (4.11), and (4.15). The taxpayer reporting error relative to the SOI estimate is:

$$\%\,\text{Rel Err(wrt } \hat{Y}) = 100 \times \frac{\hat{Y} - M}{\hat{Y}} \tag{4.19}$$

Two criteria were used to determine the best estimate of M: the bias and variance of each $\hat{M}$. The bias criterion is which alternative $\hat{M}$ is closest to the relative error (with respect to $\hat{Y}$) calculated using only the NRP data:

$$\hat{D} = \frac{\hat{X} - \hat{M}_1}{\hat{X}}. \tag{4.20}$$

These "benchmarks" are sample-based estimates from NRP's sample, but they do not have problems associated with the two different target populations and are more similar to compliance estimates typically produced by IRS. By definition, estimators (4.6) and (4.17) are algebraically equivalent.

Since each of the estimates produced from (4.3), (4.6), (4.11), and (4.15) is a sample-based estimate, the variance criterion was which $\hat{M}$ had the highest level of precision, measured by its coefficient of variation (CV), the ratio of the estimates' standard error to the estimate:

$$CV(\hat{M}) = 100 \times \frac{\sqrt{var(\hat{M})}}{\hat{M}}. \quad (4.21)$$

▶ Results

Selecting the "Best" True Total Estimates

Table 5 shows SOI's estimated totals and the estimated taxpayer error (relative to the SOI totals) using the differences in the sample estimates, ratio-based adjustments, and PS-based adjustments, and estimates using the NRP benchmarks given in (4.20). Using the NRP benchmarks and the bias criterion to compare the different methods, the combined and separate ratio estimates are exactly equal and very close to estimates produced using (4.20), respectively. All estimated relative taxpayer reporting errors are in the expected direction. That is, the taxpayer overstating in each deduction results in a positive bias, in each SOI sample-based total. Only the error estimated in Other Taxes/Personal Property Taxes using the sample differences was closer to the benchmark than that from the separate ratio estimate, but not by much. The estimates produced from the other two methods are generally not as preferable. For the sample differences, the relative taxpayer reporting error estimates in Table 5 for all variables are in the expected direction. However, the estimated errors

Table 5. Estimated SOI Estimated Totals (in Thousands of Dollars) and Taxpayer Reporting Error Estimates Relative to SOI Estimate, by Variable of Interest and Analysis Method

Variable	Taxpayer Reporting Errors, Relative to $\hat{Y}$					
	SOI Estimate [a]	*Benchmark* [b]	*Sample Diff's*	*Combine Ratio*	*Separate Ratio*	*PS Adj's*
Cash Charitable Contributions	104,439,939	14.0%	18.4%	14.0%	14.0%	15.9%
Noncash Charitable Contributions	37,888,487	11.3%	32.1%	11.3%	8.7%	30.1%
Total Adjustments, Without SE Tax Adjustment	42,052,057	5.8%	6.3%	5.8%	5.7%	3.5%
Total Taxes Deducted	307,172,690	1.8%	3.4%	1.8%	1.9%	0.5%
State and Local Taxes	195,868,643	0.4%	3.0%	0.4%	0.4%	0.1%
Real Estate Taxes	101,660,730	2.5%	2.6%	2.5%	2.6%	-0.4%
Other Taxes/Personal Property Taxes	9,643,317	21.4%	19.4%	21.4%	23.1%	17.0%
Exemptions	721,814,512	4.9%	7.5%	4.9%	4.8%	4.7%

a: SOI estimates do not include amounts from international returns.

b: Relative to $\hat{X}$.

Table 6. Estimated Coefficients of Variation of True Total Estimates, by Variable of Interest and Analysis Method

Variable	Coefficients of Variation for Alternative True Total Estimates			
	$CV(\hat{M}_1)$	$CV(\hat{M}_2)$	$CV(\hat{M}_3)$	$CV(\hat{M}_4)$
Cash Charitable Contributions	2.15%	3.01%	1.16%	2.15%
Noncash Charitable Contributions	6.04%	8.53%	3.08%	6.04%
Total Adjustments, Without SE Tax Adjustment	2.05%	3.18%	1.48%	2.05%
Total Taxes Deducted	1.19%	1.71%	0.40%	1.19%
State and Local Taxes	1.71%	2.46%	0.48%	1.71%
Real Estate Taxes	0.98%	1.47%	0.60%	0.98%
Other Taxes/Personal Property Taxes	2.86%	4.19%	2.10%	2.86%
Exemptions	0.47%	0.70%	0.30%	0.47%

Table 7. SOI Estimated Totals, Their CVs, and Taxpayer Estimates Relative to SOI and True Total Estimates, by Variable of Interest and Analysis Method

Variable	*Taxpayer Reporting Errors, Using $\hat{M}_3$*		
	SOI Estimate [a] *(CV)*	*Error Relative to $\hat{Y}$*	*Error Relative to $\hat{M}_3$*
Cash Charitable Contributions	104,439,939 (0.9%)	14.0%	16.2%
Noncash Charitable Contributions	37,888,487 (2.6%)	8.7%	9.5%
Total Adjustments, Without SE Tax Adjustment	42,052,057 (1.2%)	5.7%	6.1%
Total Taxes Deducted	307,172,690 (0.4%)	1.9%	1.9%
State and Local Taxes	195,868,643 (0.5%)	0.4%	0.4%
Real Estate Taxes	101,660,730 (0.5%)	2.6%	2.6%
Other Taxes/Personal Property Taxes	9,643,317 (1.5%)	23.1%	30.0%
Exemptions	721,814,512 (0.2%)	4.8%	5.0%

a: SOI estimates do not include amounts from international returns.

in general are much too large. The estimated relative errors when using the national PS-adjusted estimates are improvements over the sample differences, but they are not as close to the benchmarks as the estimates produced from the two ratio methods. The estimated relative errors for Cash and Noncash Contributions, State and Local Taxes, Other Taxes, Total Taxes, and Exemptions are closer to the benchmarks than the sample differences, but not close enough. Also, the errors for the other variables are further away—the error estimates are too small for Total Adjustments and negative for Real Estate Taxes, which implies that the PS correction is too large for these variables. This method appears not to work for all variables and thus is not optimal.

For the variance criterion of the true total estimates, Table 6 shows the estimated CVs of the alternative M estimates. Since the CVs of the separate ratio estimates are equal to or less than those of the combined ratio estimates for all variables, the separate ratio-based estimates of M are determined to be "best." Lastly, Table 7 contains the SOI total estimated taxpayer reporting errors, relative to the SOI estimate and $\hat{M}_3$, the adjusted preferred true total estimated using the separate ratio estimator in (4.11). The CVs of the SOI estimates are also provided, which is the relative size of sampling error associated with each SOI estimate.

Evaluating the Size of Errors

Examining the CV of each SOI estimate, the size of relative taxpayer reporting error is generally much greater than the amount of relative sampling error for every variable except State and Local Income Taxes. All relative errors are positive, which implies that taxpayers overstate these deductions to the extent that SOI's national-level totals are too large. And the amount by which they are too large, relative to both the SOI estimate and the estimated true total, is larger than the associated amount of relative sampling error for seven of the eight estimates. The largest relative differences are for the Other Taxes and Cash Charitable Contributions variables. For these variables, the estimated taxpayer reporting errors, relative to the SOI estimated totals, are 23.1 percent and 14.0 percent, respectively. The same amount of total taxpayer reporting error is estimated to be 30.0 percent and 16.2 percent of the estimated true totals, respectively.

Conclusions, Limitations, and Future Research

General Remarks

This research attempts to combine information from two samples, where one sample's values are preferred to the others, to produce estimates of error in the original sample. Despite the NRP sample being less than one-third the size of SOI's, it was large enough to produce generally reasonable error estimates at the national level.

Two general conclusions can be drawn. Firstly, all estimated errors in SOI's totals in Table 7 were positive, indicating that taxpayers are overstating all these deductions. The largest errors, relative to the size of the estimates, are for the Cash and Noncash Contributions and Other Taxes variables, despite the analysis method used. These variables thus have higher amounts of taxpayer reporting error than the other deductions examined. One possible explanation is that the error in Other Taxes is mostly due to inadvertent mistakes, as this is a more complicated itemized deduction for taxpayers to report. The errors in Cash Charitable Contributions, however, are probably more related to tax evasion, as cash is a common financial transaction that is fabricated (or in this case, possible nonexistent transactions).

State and Local Taxes and Real Estate Taxes had the lowest relative taxpayer reporting errors. Of these, the estimated relative taxpayer reporting error in State and Local Taxes was lower than the associated estimated amount of relative sampling error in the SOI sample-based total. For both these variables, State and local governments provide taxpayers with written statements of the associated deductible amounts. It is thus easier for taxpayers to report the correct amounts and harder for them to inflate these numbers, as there is an existing paper trail.

The results also verify empirically that all eight variables are misreported by taxpayers to a magnitude that is most often higher than the sampling error associated with each SOI sample estimate (which is very small for these variables). While the cause of this is not determined (whether misreporting was intentional or not), the result is the same: overstating deduction items leads to taxpayers subtracting amounts that are too large from their incomes, resulting in a lower amount of tax reported. It is assumed that the misreporting arises from a combination of these variables being more difficult for taxpayers to report and possibilities for evasion.

Secondly, the two different methods produced different error estimates. Some variables, such as Cash Contributions, Total Adjustments, and Exemptions seemed less sensitive (more robust) to which method was used; both methods produced relative errors close to the NRP benchmarks. Noncash Contributions, however, had different errors for the separate methods. Across the methods, the ratio methods were more consistent to the NRP benchmarks, and the separate ratio estimated true totals had the lowest estimated CVs. From this, they are preferable error estimates over using the difference in the two samples' estimates or PS adjustments to the NRP estimates.

Research Limitations

One limitation in this analysis is the assumption that the NRP auditors detected the true values. This may be reasonable for deductions, where the burden of proof is designated to the taxpayer, but not income- or tax-related variables. At the time of this research, IRS's Office of Research had not produced compliance estimates at the variable level. It will be useful to compare these results to theirs, when available.

Another data assumption is that the SOI estimates are without error. Scali et al. (2005) showed that the editing error in several 1040 tax variables was not significant, but they did not examine Schedule A variables. Also, processing error in both samples is assumed negligible.

Each analysis method considered also has associated explicit and implicit assumptions. Most are reasonable, under the circumstances, and each is discussed separately.

Assumptions for the sample differences are that the two samples and their target populations are comparable.

This is more reasonable for itemized deductions, where it is less likely that the two samples differed in their target populations, but this would need to be examined for income- and tax-related variables. Also, differences of two different sponsoring IRS offices, target populations, sample designs, data collection methods, and sample objectives produced error estimates that are too large when comparing the two samples' estimates that must be accounted for with more sophisticated analysis.

The combined and separated ratio-based adjustments assume that the taxpayer reporting behavior is homogeneous within each sample and stratum and between the two samples. That is, implicitly assumed is that the reporting behavior detected in the NRP returns is the same as that in the SOI returns in the entire sample or within the same stratum. More sophisticated methods already developed in the survey methodology field for forming nonresponse cells may provide better ways to create adjustment cells instead of just using the SOI sample strata definitions.

Lastly, the poststratification adjustments assume that the same adjustment works for all variables. These results show that this does not appear to hold.

Future Considerations

Several extensions are worth considering. Firstly, given that SOI's sample is selected annually with a large overlap of returns between different years' samples, it would be interesting to examine whether taxpayers associated with the 433 returns in both samples changed their reporting behavior due to the audits. Omitted results showed that these taxpayers are not representative of the general tax population, but using longitudinal analysis methods (e.g., Fitzmaurice et al., 2004) on multiple years of taxpayer reported values reported in SOI samples before and after the 2001 audits could examine reporting over time.

Secondly, despite removing the international returns, the two samples' estimated population totals are still quite different. A more sophisticated approach such as raking (e.g., Oh and Scheuren, 1987) may produce more stable results. Thirdly, alternative ways to form adjustment cells for the separate ratio estimator deserve attention, particularly since this method produced favorable results. Collapsing SOI strata further would also allow for incorporating multiple taxpayer characteristics.

The next Form 1040 NRP sample will be implemented for Tax Years 2006-2008. As this sample includes international returns, it would be beneficial for SOI to coordinate with NRP to ensure that similar (or more sophisticated) analyses can be conducted in the future.

▶ References

Bennett, C. (2005), "Preliminary Estimates of NRP," *Internal Revenue Service Research Bulletin*.

Binder, D. A. (1996), "Linearization Methods for Single Phase and Two-Phase Samples: A Cookbook Approach," *Survey Methodology*, 22, Number 1, pp. 19-22.

Bloomquist, K. (2004), "Multi-Agent Based Simulation of the Deterrent Effects of Taxpayer Audits," Proceedings of the November National Tax Association Conference.

Fitzmaurice, G.; N. Laird; and J. Ware, *Applied Longitudinal Analysis*, John Wiley and Sons, New York.

Internal Revenue Service (2002), *IRS Sets New Audit Priorities*, IRS News Release FS-2002-12, Washington DC.

Internal Revenue Service (2003), *Internal Revenue Service Data Book*, Fiscal Year 2002, Washington, DC.

Internal Revenue Service (2006), http://www.irs.gov/pub/irs-utl/tax_gap_update_070212.pdf.

Kalton, G. and D. Kasprzyk (1986), "The Treatment of Missing Survey Data," *Survey Methodology*, 12, pp. 1-16.

Kalton, G. and D. S. Maligalig (1991), "A Comparison of Methods of Weighting Adjustment for Nonresponse," Proceedings of the U. S. Bureau of the Census Annual Research Conference, pp. 409-824.

Oh, H. L. and F. J. Scheuren (1983), "Weighting adjustment for unit nonresponse," in W. Madow; I. Olkin; and D. Rubin, D. (editors), *Incomplete Data in Sample Surveys*, Volume 2, Academic Press, New York, pp. 143-184.

Oh, H.L. and F. J. Scheuren (1987), "Modified raking ratio estimation," *Survey Methodology*, 13, pp. 209-219.

McMahon, P. B. (2004), "Assessing Industry Codes on the IRS Business Master File," Proceedings of the Section on Survey Research Methods Section, American Statistical Association.

Plumley, A. H. (2005), Preliminary Update of the Tax Year 2001 Individual Income Tax Underreporting Gap Estimates," *Internal Revenue Service Research Bulletin*.

Scali, J.; V. Testa; M. Kahr; and M. Strudler (2005), "Measuring Nonsampling Error in the Statistics of Income Individual Tax Return Study," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Särndal, C.E.; B. Swensson; and J. Wretman (1992), *Model Assisted Survey Sampling*, Springer-Verlag, New York.

Slemrod, J. and Bakija (2004), *Taxing Ourselves: A Citizen's Guide to the Debate Over Taxes*, The MIT Press, Cambridge.

Valliant, R. and R. Casady (1998), "The Variance of An Estimator of Change from the Sample of Partnership Tax Returns," *Statistical Issues Related to SOI Family Cross-Sectional and Longitudinal Studies*, Contract Number TIRN0-96-D-0030 Task 5, Washington DC.

Walker, B. and V. Testa (2003), *Statistics of Income—2001, Individual Income Tax Returns*, Internal Revenue Service, Publication 1304, pp. 23-27.

Weber, M., (2004), "The Statistics of Income 1979-2002 Continuous Work History Sample Individual Tax Return Panel," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Wolter, K., (1985), *Introduction to Variance Estimation*, Springer-Verlag, New York.

Woodruff, R. S. (1971), "A simple method for approximating the variance of a complicated estimate," *Journal of the American Statistical Association*, 66, pp. 411-414.