A Stitch in Time:
Knitting the Results of the Tax Year 2003/2004 NRP
S Corporation Reporting Compliance Study

Drew Johns
Internal Revenue Service

September 2009

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily reflect the views or the official positions of the U.S. Department of the Treasury or the Internal Revenue Service. All results have been reviewed to ensure that no confidential information is disclosed.

Paper was originally presented at the 2009 IRS Research Conference:
https://www.irs.gov/statistics/soi-tax-stats-2009-irs-research-conference.

Slides are available at:
https://www.irs.gov/pub/irs-soi/09resconawardscorp.pdf.

Introduction

For Tax Year (TY) 2001, IRS estimated that flow-through income (partnerships, S corporations, and estates and trusts) on individual income tax returns was underreported (net of overreporting) by 18% and contributed $22 billion to the tax gap. In order to gain a better understanding of the compliance of S corporations, the IRS National Research Program (NRP) randomly selected for audit 1,180 S corporation TY2003 income tax returns and 3,700 TY2004 tax returns. This paper pools the two tax years of the S corporation reporting compliance study into a single dataset. The data and sample weights are adjusted so that the pooled data can be reasonably interpreted as reflecting the population of TY2004 S corporation returns. The pooled data are then used to compare the compliance of small, closely held S corporations to sole proprietors that file Schedule C with their Form 1040.

Pooling the data in this way serves several purposes. First, pooling of the data for the two tax years improves the precision of the compliance estimates. Second, adjusting the TY2003 data to be more representative of TY2004 simplifies the interpretation and presentation of the results of the study. Finally, for the National Research Program reporting compliance study of individual income tax returns, IRS initiated annual compliance studies of around 13,000 individual income tax returns beginning with TY2006. These annual studies of individual income tax returns are much smaller than the TY2001 individual reporting compliance study of about 45,000 income tax returns. To obtain a similar precision as the larger TY2001 study and sufficient number of returns to estimate the individual income tax underreporting gap, IRS will need to pool the data from multiple years of the smaller studies.[1] The TY2003/2004 NRP S corporation reporting compliance study presents an opportunity to explore methodologies for pooling data from

multiple tax years in preparation for analyzing future individual income tax reporting compliance data.

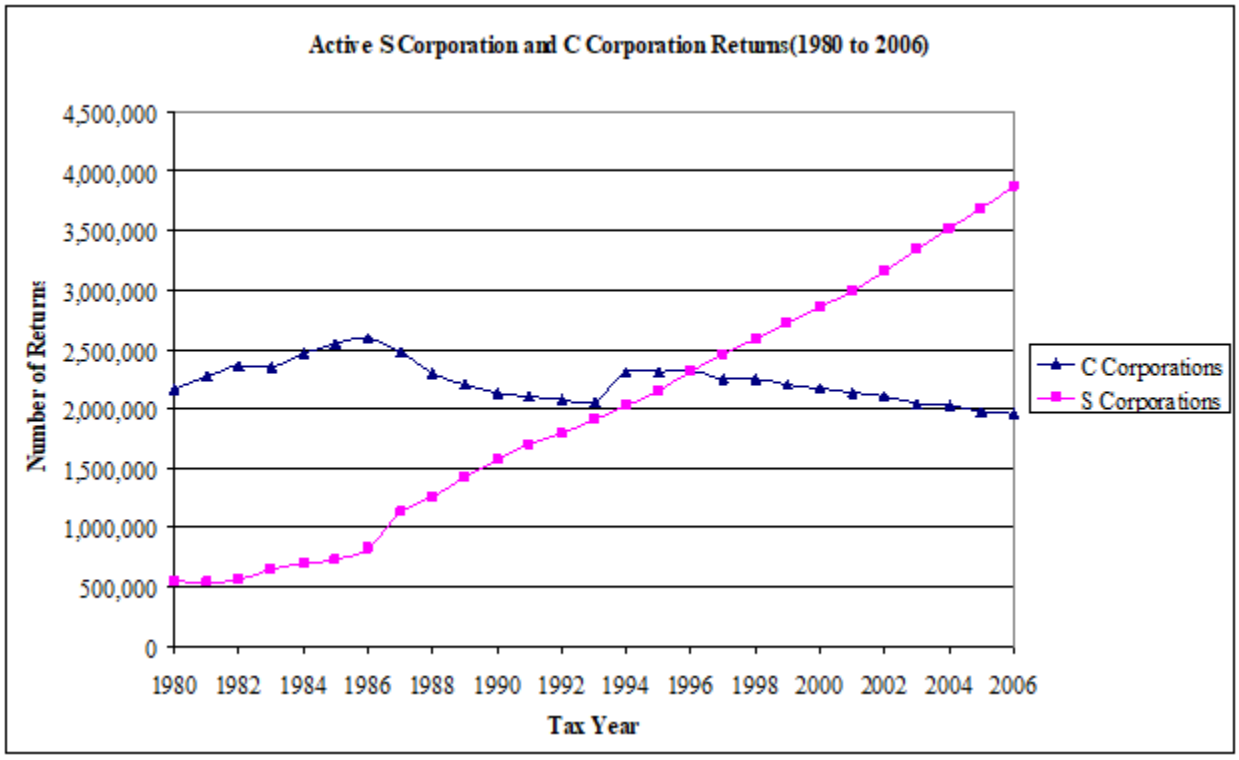
Background on S Corporations

S corporations must file an income tax return (Form 1120S), but as flow-through entities, their income is generally not taxed at the corporate level. Instead, the corporation's income flows down on a pro rata basis to its shareholders who are then taxed at their respective rates, regardless of whether the S corporation actually distributes the income to its shareholders. For the most part, the rules governing S corporations avoid the double taxation of profits for the shareholder by allowing the corporation to distribute earnings tax-free.[2] However, shareholder officers are required to take a reasonable salary for the services they provide the S corporation. The S corporation is able to deduct the officer's compensation and the shareholder is required to report that compensation as wages and salary on their Form 1040. Unlike distributions, officer's compensation is subject to employment taxes. Corporations also must meet several requirements in order to qualify as an S corporation. For example, the corporation must be a domestic entity with fewer than 75 shareholders (changed to 100 for tax years after 2004). [3] About 90% of S corporations have only 1 or 2 shareholders.

For Tax Years 2001 to 2004, IRS Statistics of Income estimates that approximately 80% of S corporation total net income reported on Schedule K of Form 1120S in a given year was from ordinary income or loss from trade or business activities.[4] Long-term capital gains made up roughly 10% of total net income with the actual values ranging from 7 to 13%. Rental real estate

income accounted for 3 to 4% of total net income while interest accounted for another 3 to 7%, with interest income declining about one percentage point per year over the period.

The avoidance of double taxation (as in the case of C corporations) makes the S corporation an attractive entity for taxpayers looking to choose an entity form that minimizes their overall tax burden. A significant amount of research in the mid-1990s looked at the effect of taxation on the choice of business form, particularly on the impact of the Tax Reform Act of 1986 (TRA86). [5] The population of S corporations received a boost due to the TRA86 which reduced individual income tax rates relative to corporate tax rates. Figure 1 uses IRS Statistics of Income data to show a diverging trend in the filing of S corporation and C corporation returns beginning in 1986. There were approximately 826,000 S corporation returns filed for TY1986 compared with 3.9 million returns filed for TY2006, an increase of nearly 370%. Over the same period, the number of C corporation returns filed fell 25% from 2.6 million 2.0 million.



Rolling Samples

The concept of combining smaller samples over time to increase the precision of estimates and to allow for improved estimation over small domains was largely developed by Leslie Kish and the methodology is referred to as a rolling sample design.[6] When samples are costly to administer and there is a desire for timely and detailed information, a rolling sample design is an option. The U.S. Census Bureau uses a rolling sample design with the American Community Survey (ACS). For some geographical areas, one year of ACS data is sufficient to provide estimates at the desired level of precision. Smaller counties or states will require three or even five years of data to obtain the desired level of precision.

Two analytical challenges of rolling samples are combining the data across years and then interpreting estimates based on the combined data. The U.S. Census Bureau considered several options for combining the data (Tersine and Asiala, 2007). The approaches can be summarized into two general categories:

(1) averaging (either simple or weighted) annual estimates or

(2) pooling the data into a single dataset and developing case level weights.

The first approach is somewhat appealing in its simplicity. If an analyst has estimates from two independent random samples of approximate equal size, then the two estimates are summed and then divided by two to obtain a simple average for the period. However, in the situation of the S corporation study, the two samples are of unequal size with approximately 25% of the pooled sample for TY2003 and the remaining 75% for TY2004. Rather than weighting the two separate

estimates equally in this scenario, one could apply weights equal (or approximately equal) to each tax years proportion of the pooled sample. As an example, the net misreporting percentage (NMP) is defined as the total weighted net misreported amount of income (or expenses) divided by the weighted sum of the absolute values of the amounts that should have been reported. Based on the unpooled results from the TY2003/2004 NRP S corporation reporting compliance study, the NMP for ordinary business income (loss) is estimated to be 13.9% for TY2003 and 15.0% for TY2004. A simple average with equal weighting provides an estimated NMP of 14.5% for the two years. Alternatively, a weighted average of the two years (.25*13.9% + .75*15.0%) provides an estimated NMP of 14.7% for the period.[7]

The Census Bureau uses the second approach for the ACS. With the ACS, the data are pooled across multiple years (e.g., three years) and then weighted by controlling to the average population (by geographic region) across those years[8]. The Census Bureau also controls for inflation in their period estimates through the national level Consumer Price Index (CPI) from the Bureau of Labor Statistics (Tersine and Asiala, 2007).

When estimates across multiple years are averaged (as in the first approach) or multiple years of data are pooled (as in the second approach) there remains the challenge of interpreting those results. The Census Bureau elected to interpret the multi-period estimates based on pooling the ACS data as period estimates, essentially representing the average for the years of the data that were pooled. An alternative would be to use either a model or set of assumptions to adjust the data from earlier years to reflect the latest year's population and interpret the pooled data as reflecting the latest year. The data could be "aged" or "knit" in this way by either controlling on

known population totals from the latest year of the study or applying observed growth rates and adjustment factors to earlier data. One advantage of this approach is the increased ease of interpreting the data to reflect a single year. In the case of the TY2003/2004 NRP S corporation reporting compliance study, the pooled data could be interpreted as reflecting TY2004. One drawback of this approach is the introduction of some unknown amount of additional uncertainty in the estimates through the model or assumptions used to adjust the data.

Data and Methodology

NRP originally selected 1,453 S corporation returns out of almost 3.3 million returns from TY2003 and 4,401 returns out of almost 3.5 million returns from TY2004 for a total of 5,854 returns. The sample for each tax year was stratified by audit code and reported assets and was designed to be representative of the S corporation population for its respective tax year.[9] After accounting for returns that were excluded from the study, the TY2003/2004 NRP S corporation reporting compliance study consisted of 1,180 TY2003 returns and 3,700 TY2004 returns. [10] The sampling frames for the TY2003 and TY2004 samples were slightly different and a limited number of audits were not completed by the time of this analysis. Therefore, each sample was limited to returns that were filed in the calendar year following the year the return was due and any remaining open cases were excluded from the analysis. This additional removal of cases resulted in a final sample for analysis of 1,144 returns for TY2003 and 3,671 returns for TY2004, for a combined total sample of 4,815 returns. Both samples of the study are relatively small and IRS believed that two years of the study could be combined in order to improve the precision of the noncompliance estimates from the study. Therefore, the TY2003/2004 S corporation reporting compliance study shares characteristics similar to a rolling sample.

In this paper, the NRP data are pooled into a single dataset instead of averaging the annual estimates. However, the approach in this paper differs from the Census Bureau's methodology for pooling the ACS data. Adjustment factors are applied to the data and weights, instead of reweighting the pooled data to population targets. These adjustments to the data allow the pooled data reasonably to reflect the TY2004 S corporation population rather than interpreting the results as an average over the two years of the study. To pool the TY2003 and TY2004 S corporation samples into a single dataset, four adjustment factors are applied to the data: a Population Growth Factor (PGF), a Sample Proportion Factor (SPF), an Income Growth Factor (IGF), and an Expense Growth Factor (EGF). The PGF captures growth in the number of S corporation returns filed. The SPF ensures that the weights of the pooled dataset sum to the number of returns for a single year rather than two years. Separate IGF and EGF factors allow for the possibility that income and expenses may grow at different rates.[11] S corporation return population data from the IRS Compliance Data Warehouse (CDW) were used to estimate the PGF, IGF, and EGF.

Population Growth Factor

A population growth factor equal to the ratio of the population of S corporation returns in TY2004 to the population of S corporation returns in TY2003 (grouped by level of reported assets) is applied to the sampling weights for TY2003 returns. The PGF is defined as:

$$PGF = \begin{cases} 1.0528 & \text{Reported Assets} < \$200{,}000 \\ 1.0582 & \$200{,}000 <= \text{Reported Assets} < \$10{,}000{,}000 \\ 1.0966 & \text{Reported Assets} >= \$10{,}000{,}000 \end{cases}$$

Sample Proportion Factor

A sample proportion factor equal to each tax year's proportion of sample returns in the pooled sample is applied to sampling weights for both TY2003 and TY2004 returns. Table 1 shows the sample sizes and weighted population by level of reported assets for each year of the study and for the pooled data after applying the first two adjustment factors. The SPF is defined as:

$$SPF = \begin{Bmatrix} 1{,}144/4{,}815 = 0.2376 & \text{Tax Year 2003 Returns} \\ 3{,}671/4{,}815 = 0.7624 & \text{Tax Year 2004 Returns} \end{Bmatrix}$$

Table 1: Sample Sizes and Weighted Population Estimates by Activity Code for the NRP TY2003/2004 S Corporation Reporting Compliance Study

	TY2003		TY2004		Pooled Data	
Level of Reported Assets	**Sample Count**	**Weighted Population Estimate**	**Sample Count**	**Weighted Population Estimate**	**Sample Count**	**Weighted Population Estimate**
Assets under $200K	531	2,479,680	1,703	2,609,695	2,234	2,609,912
$200K <= Assets < $10M	546	785,774	1,763	833,313	2,309	832,883
Assets >= $10M	67	28,254	205	30,575	272	30,672
	1,144	**3,293,708**	**3,671**	**3,473,583**	**4,815**	**3,473,468**

Income Growth Factor

An income growth factor equal to the ratio of the population average "total income (loss)" (Line 6 of Form 1120s) for TY2004 to the population average "total income (loss)" for TY2003 by level of reported assets was applied to all TY2003 income line items. The IGF was also applied to the "returns and allowance" and "cost of goods sold" lines on the Form 1120S. These two lines are components of "total income (loss)" and it assumed that they grow at the same average

rate as “total income (loss)”. Separate IGFs by level of reported assets allows for differences in income growth between smaller and larger S corporations. The IGF is a nominal growth factor that implicitly captures growth in real economic activity as well as growth in prices, both of which are necessary to knit two tax years together. The IGF uses growth in average “total income (loss)” rather than the total because the growth in the total would include population growth that the PGF already captures. The IGF is defined as:

$$IGF = \begin{Bmatrix} 1.0337 & \text{Reported Assets} < \$200{,}000 \\ 1.0346 & \$200{,}000 <= \text{Reported Assets} < \$10{,}000{,}000 \\ 1.0444 & \text{Reported Assets} >= \$10{,}000{,}000 \end{Bmatrix}$$

Expense Growth Factor

An expense growth factor equal to the ratio of the population average “total deductions” (Line 20 of Form 1120s) for TY2004 to the population average “total deductions” for TY2003 by level of reported assets was applied to all TY2003 expense and deduction line items, except for the “returns and allowances” and “cost of goods sold” which received the IGF. The EGF values compared to the IGF values show that expenses of the larger S corporations grew at a slower rate than their income while income and expenses of the smallest S corporations grew at a similar rate. The EGF is defined as:

$$EGF = \begin{Bmatrix} 1.0329 & \text{Reported Assets} < \$200{,}000 \\ 1.0199 & \$200{,}000 <= \text{Reported Assets} < \$10{,}000{,}000 \\ 1.0264 & \text{Reported Assets} >= \$10{,}000{,}000 \end{Bmatrix}$$

Summary of S Corporation Entity Level Compliance Results [12]

The **net misreported amount** (NMA) is defined for income items as the amount of underreported income net of overreported income (it is positive for net underreported income), and is defined for offset items (e.g., expenses) as the amount of overreported offsets net of underreported offsets (it is positive for net overreported offsets to income). The **net misreporting percentage** (NMP) is defined as the ratio of the NMA to the sum of the absolute values of the amounts that should have been reported. The **error rate** is defined as the ratio of the number of returns with an error to the number of returns that should have reported something on that line. The deduction from ordinary business income of officers' compensation at the entity level is typically offset as wage income at the individual taxpayer level, resulting in no effect on tax at the shareholder level. Therefore the error estimates discussed in this section exclude errors to officers' compensation.[13] Some of the tables in this section include the **relative margin of error** for the NMA, NMP, and error rate. The relative margin of error of an estimate is defined as the margin of error of the estimate (defined as one half of the width of a 95% confidence interval) expressed as a percentage of the estimate. In the case of the NMP and error rate, the relative margin of error is a percentage of a percentage. Table 2 shows that the estimate for the TY2004 NMP for ordinary business income (loss) is 14.6% with a 17% relative margin of error. This means that a 95% confidence interval around the NMP for ordinary business income (loss) would range from 12.1% to 17.1%.[14]

Based on the pooled results of the TY2003/2004 National Research Program S corporation reporting compliance study, Table 2 shows compliance estimates for select separately stated items on the Schedule K of the Form 1120S for TY2004. S corporations underreported an

estimated $56 billion of total income for TY2004. [15] This amount represents a net misreporting percentage (NMP) of 14%. Underreported ordinary business income accounts for 92% of TY2004 total underreported income. In TY2004, ordinary business income was misreported by almost $52B, resulting in an NMP of nearly 15%. About 68% of S corporations that should have reported ordinary business income made an error that affected ordinary business income. This error rate includes errors of any size, including small errors. After ordinary business income, the two largest contributors to underreported income by S corporations were net rental real estate income or losses and the section 179 deduction.[16] About $1.5 billion in net rental real estate income was underreported in TY2004, resulting in an NMP of 13%. About $1 billion of the section 179 deduction was overreported, resulting in an NMP of nearly 7%.

Table 2: Estimates of Tax Year 2004 S Corporation Reporting Compliance: Schedule K - Shareholders' Shares of Income and Deductions -- Amounts are in $ Millions and Counts are in Thousands[1]

Schedule K Income/Deduction	Returns that Reported	Amount Reported	Returns That Should Have Reported	Amount Should Have Reported	Returns with Errors	NMA	NMP	Error Rate	Relative Margin of Error NMA	Relative Margin of Error NMP	Relative Margin of Error Error Rate
Income (loss):											
Ordinary business income (loss)	3,221	224,359	3,180	264,083	2,196	39,390	11.1%	69.1%	13%	18%	3%
Ordinary business income (loss) excluding Officers' Comp. NMA [2]	3,221	224,359	3,180	264,083	2,171	51,524	14.6%	68.3%	10%	17%	3%
Net rental real estate income (loss)	255	6,522	256	8,092	91	1,454	13.3%	35.4%	47%	38%	16%
Interest income	851	7,214	876	7,858	87	641	8.2%	9.9%	67%	65%	19%
Net short-term capital gain (loss)	53	542	58	755	7	187	14.1%	12.4%	138%	132%	75%
Net long-term capital gain (loss)	98	11,237	104	11,633	15	414	2.5%	14.8%	115%	116%	47%
Net section 1231 gain (loss)	198	28,709	192	29,145	35	436	1.3%	18.1%	290%	300%	31%
Other income (loss)	37	890	47	947	17	46	3.9%	35.5%	212%	225%	49%
Total Sch. K Income (loss)	**3,384**	**281,909**	**3,346**	**324,807**	**2,288**	**42,652**	**10.3%**	**68.4%**	**12%**	**18%**	**3%**
Total Sch. K Income (loss) excluding Officers' Comp. NMA [2]	**3,384**	**281,909**	**3,346**	**324,807**	**2,269**	**54,706**	**13.2%**	**67.8%**	**9%**	**16%**	**3%**
Deductions:											
Section 179 deduction	839	16,893	829	15,917	172	1,060	6.7%	20.7%	42%	43%	16%
Contributions	1,026	3,490	1,045	3,558	79	42	1.2%	7.6%	251%	254%	23%
Total Sch. K Deductions	**1,523**	**27,401**	**1,528**	**26,439**	**257**	**-1,180**	**-4.5%**	**16.8%**	**39%**	**42%**	**13%**
M-1 Income/Loss Reconciliation	**3,373**	**254,508**	**3,339**	**298,368**	**2,297**	**43,832**	**11.2%**	**68.8%**	**12%**	**18%**	**3%**
M-1 Income/Loss Reconciliation excluding Officers' Comp. NMA [2]	**3,373**	**254,508**	**3,339**	**298,368**	**2,281**	**55,886**	**14.3%**	**68.3%**	**9%**	**17%**	**3%**

(1) TY04 estimates based on pooling the TY03 and TY04 results from the NRP Tax Year 2003/04 S corporation reporting compliance study. Estimates are based on the raw findings from the NRP study and are not adjusted for noncompliance that was not detected during the audit. Some items are not shown. Relative margins of error are based on 95% level of confidence.

(2) Errors to officer's compensation are excluded from the calculation of Net Misreported Amount.

Table 3 shows compliance estimates for select items related to ordinary business income reported on the front page of the Form 1120S. Underreporting of net receipts (receipts less returns and allowances) accounted for nearly 34% of total net misreported ordinary business for TY2004. However, less than half a percent of net receipts were underreported. About 27% of S corporations that should have reported net receipts misreported their net receipts. Ordinary business income other than net receipts accounted for nearly 6% of total net misreported ordinary business income. The net overreporting of the cost of goods sold accounted for almost 10% of

underreported ordinary business income, although there was significant variation in the misreporting of the cost of goods sold.[17]

Table 3: Estimates of Tax Year 2004 S Corporation Reporting Compliance: Form 1120S - Ordinary Income -- Amounts are in $ Millions and Counts are in Thousands[(1)]

Form 1120S Line Item	Returns that Reported	Amount Reported	Returns That Should Have Reported	Amount Should Have Reported	Returns with Errors	NMA	NMP	Error Rate	Relative Margin of Error NMA	NMP	Error Rate
Income:											
Balance (Net Receipts)	**2,885**	**4,241,719**	**2,891**	**4,258,850**	**792**	**17,408**	**0.4%**	**27.4%**	**32%**	**33%**	**7%**
Cost of Goods Sold	1,780	2,722,362	1,775	2,717,101	659	4,989	0.2%	37.1%	87%	88%	7%
Gross Profit	**2,890**	**1,519,357**	**2,898**	**1,541,748**	**1,127**	**22,397**	**1.4%**	**38.9%**	**17%**	**19%**	**6%**
Other Income (Loss)	796	115,172	814	117,814	118	2,737	2.2%	14.5%	65%	70%	17%
Net Gain (Loss) from F4797	189	7,649	196	7,993	43	344	3.6%	21.8%	81%	91%	30%
Total Income (Loss)	**2,991**	**1,642,178**	**2,996**	**1,667,556**	**1,209**	**25,479**	**1.5%**	**40.4%**	**17%**	**18%**	**5%**
Deductions:											
Compensation of Officers	1,743	165,846	1,939	178,882	433	-12,134	-6.8%	22.3%	15%	16%	10%
Other Deductions	3,124	474,624	3,077	456,652	1,753	17,235	3.7%	57.0%	17%	19%	4%
Rents	1,822	85,373	1,768	83,856	253	1,491	1.8%	14.3%	34%	36%	15%
Net Depreciation	2,097	62,692	2,012	60,240	493	2,183	3.6%	24.5%	31%	36%	9%
Interest	1,562	36,620	1,519	35,685	267	885	2.5%	17.6%	33%	36%	12%
Repairs and Maintenance	1,767	24,826	1,720	23,439	263	1,356	5.8%	15.3%	32%	33%	14%
Advertising	1,809	40,274	1,790	39,680	196	518	1.3%	11.0%	48%	53%	17%
Bad Debts	290	7,150	271	6,677	34	473	6.8%	12.5%	55%	59%	37%
Taxes and Licenses	2,791	71,791	2,763	71,186	291	559	0.8%	10.5%	46%	45%	13%
Salaries and Wages	1,481	407,382	1,482	406,863	114	447	0.1%	7.7%	118%	118%	19%
Employee Benefit Programs	561	24,471	533	23,920	86	522	2.2%	16.1%	32%	38%	20%
Pension, Profit-Sharing, etc., Plans	442	16,341	436	16,037	19	302	1.9%	4.4%	94%	96%	45%
Total Deductions	**3,200**	**1,417,819**	**3,165**	**1,403,473**	**2,025**	**13,911**	**1.0%**	**64.0%**	**25%**	**25%**	**3%**
Total Deductions excluding Officers' Comp. NMA[(2)]	**3,200**	**1,417,819**	**3,165**	**1,403,473**	**1,983**	**26,045**	**1.9%**	**62.7%**	**12%**	**14%**	**3%**
Ordinary Income (loss)	**3,221**	**224,359**	**3,180**	**264,083**	**2,196**	**39,390**	**11.1%**	**69.1%**	**13%**	**18%**	**3%**
Ordinary Income (loss) excluding Officers' Comp. NMA[(2)]	**3,221**	**224,359**	**3,180**	**264,083**	**2,171**	**51,524**	**14.6%**	**68.3%**	**10%**	**17%**	**3%**

(1) TY04 estimates based on pooling the TY03 and TY04 results from the NRP Tax Year 2003/04 S corporation reporting compliance study. Estimates are based on the raw findings from the NRP study and are not adjusted for noncompliance that was not detected during the audit. Some items are not shown. Relative margins of error are based on 95% level of confidence.

(2) Errors to officer's compensation are excluded from the calculation of Net Misreported Amount.

Deductions from ordinary business income accounted for nearly 51% of total misreported ordinary business income in TY2004. S corporations misreported just under 2% of the total amount of deductions that should have been reported. About 29% of S corporation deductions are made on the "salaries and wages" line. Taxpayers misreported about one tenth of one

percent of "salaries and wages" in TY2004. Businesses are required to issue W-2s to employees and the Social Security Administration (SSA) who then processes and shares the W-2 information with the IRS. Businesses also must file employment tax returns and report wages and salaries to the IRS. The IRS has the ability to compare wages reported on the employee's tax return and the employer's employment tax returns to the wages deducted on the S corporation income tax return. As errors would be detected, S corporations have an incentive to deduct and report salaries and wages accurately. This result is consistent with the accuracy of deductions by sole proprietors who file Schedule C with their Form 1040. The Schedule C expense with the lowest detected net misreporting percentage from the NRP TY2001 individual reporting compliance study was the "wages" line with an NMP of less than 2%.

The S corporation expenses with the highest net misreporting percentages were the "repairs and maintenance" and "bad debts" lines. These two expense lines were misreported by nearly 6% and 7%, respectively, of the amounts that should have been reported in TY2004. These two expense lines also had the highest net misreporting percentages on Schedule C for sole proprietors in the TY2001 NRP individual reporting compliance study, 21% for "repairs and maintenance" and 71% for "bad debts".

"Other deductions" reported on Line 19 of the Form 1120S accounted for nearly 33% of total deductions that should have been reported for TY2004. About 66% of TY2004 overreported deductions from ordinary business income were from "other deductions" reported on Line 19. Overall, "other deductions" accounted for about 33% of TY2004 total net misreported ordinary business income. "Other deductions" had an NMP of nearly 4% and was the most frequent type

of error with an error rate of 57%. Net depreciation had the next highest error rate among expense lines with an error rate of nearly 25%.

Table 4 shows the misreporting of "other deductions" by category for TY2004.[18] The five "other deductions" categories with the highest net misreporting percentages were "automobile and truck" (16% in TY2004), "travel" (13%), "meals and entertainment" (11%), "miscellaneous expenses" (9%), and "tools and factory supplies" (8%). The four categories with the largest net misreported amount also were misreported with the highest frequency: "automobile and truck", "travel", "insurance", and "other expenses".

Table 4: Estimates of Tax Year 2004 S Corporation Reporting Compliance: Other Deductions Reported on Line 19 of Form 1120S -- Amounts are in $ Millions and Counts are in Thousands[(1)]

Form 1120S "Other Deductions"	Amount Reported	NMA	NMP	Error Rate
Automobile and truck	22,375	3,088	16%	41%
Travel	14,984	1,683	13%	28%
Insurance	57,425	1,847	3%	27%
Other expenses	91,728	2,946	3%	25%
Legal and professional fees	24,940	1,208	5%	12%
Printing and office supplies	24,766	1,325	6%	16%
Tools and factory supplies	6,530	505	8%	17%
Utilities	23,280	349	2%	14%
Telephone	17,027	434	3%	16%
Outside expenses	16,911	234	1%	15%
Meals and entertainment	4,826	472	11%	20%
Misc. expenses	5,064	446	9%	14%
Contract Labor	21,795	475	2%	15%
Equipment and rental	3,736	259	7%	13%
Computer expenses	5,141	112	2%	7%
Amortization	5,309	251	5%	9%
Dues and subscriptions	3,600	135	4%	8%
Postage and courier	6,488	30	0%	4%
Shipping and freight	23,675	85	0%	10%
Public relations	913	32	4%	14%

(1) TY04 estimates based on pooling the TY03 and TY04 results from the NRP Tax Year 2003/04 S corporation reporting compliance study. Some items are not shown.

Officers' Compensation

Although entity level officers' compensation errors typically offset at the shareholder level, underreported officers' compensation contributes to the employment tax underreporting gap. Unlike ordinary business income, officers' compensation is subject to employment taxes and the

underreporting of officers' compensation typically results in underreporting of employment taxes. As the S corporation study results show, officers' compensation is the only net underreported deduction to ordinary business income and it is underreported by a relatively high percentage. Officers' compensation was underreported by a net $12 billion in TY2004, nearly 7% of the amounts that should have been reported.

Compliance by Level of Reported Assets

Table 5 shows the misreporting of ordinary business income and other select items by level of reported end-of-year assets. While S corporations with assets under $200,000 accounted for only 24% of reported ordinary business income for TY2004, these smaller S corporations accounted for 53% of the total net misreported ordinary business income. The largest S corporations with assets over $10 million accounted for 39% of reported ordinary business income, but only 9% respectively of net misreported ordinary business income. The smallest S corporations had an NMP of 27% for ordinary business income TY2004, compared to an NMP of 4% for the largest S corporations.

Table 5: S Corporation Reporting Compliance by Level of Reported Assets for Tax Year 2004:
-- Amounts are in $ Billions and Counts are in Thousands[1]

Level of Assets	Returns that Reported	Amount Reported	NMA	Share of Amount Reported	Share of NMA	NMP	Error Rate	Relative Margin of Error	
								NMA	NMP
Assets under $200K									
Net Receipts	2,186	710	11	17%	62%	2%	31%	40%	40%
Total Income	2,253	458	13	28%	52%	3%	42%	22%	23%
Officer's Compensation	1,248	68	-9	41%	73%	-11%	25%	18%	18%
Total Deductions[2]	2,438	405	14	29%	53%	3%	64%	18%	20%
Ordinary Income[2]	2,447	53	27	24%	53%	27%	70%	12%	16%
$200K <= Assets < $10M									
Net Receipts	673	1,986	5	47%	29%	0%	18%	53%	53%
Total Income	709	736	10	45%	38%	1%	35%	26%	26%
Officer's Compensation	473	73	-3	44%	25%	-4%	15%	26%	28%
Total Deductions[2]	732	653	10	46%	38%	2%	59%	17%	18%
Ordinary Income[2]	744	83	20	37%	38%	14%	64%	16%	16%
Assets >= $10M									
Net Receipts	26	1,546	2	36%	9%	0%	11%	138%	140%
Total Income	30	448	3	27%	10%	1%	41%	65%	69%
Officer's Compensation	22	25	0	15%	2%	-1%	9%	107%	111%
Total Deductions[2]	30	360	2	25%	8%	1%	58%	47%	48%
Ordinary Income[2]	30	88	5	39%	9%	4%	69%	44%	56%

(1) TY04 estimates based on pooling the TY03 and TY04 results from the NRP Tax Year 2003/04 S corporation reporting compliance study. Estimates are based on the raw findings from the NRP study and are not adjusted for noncompliance that was not detected during the audit. Relative margins of error are based on 95% level of confidence.

(2) Errors to officer's compensation are excluded from the calculation of Net Misreported Amount.

The majority of misreported net receipts also occurred on returns from smaller S corporations. About 62% of TY2004 misreported net receipts were by S corporations with reported assets under $200,000; although these corporations account for only 17% of total reported net receipts. The largest S corporations, those with over $10 million in assets, accounted for 36% of reported net receipts, but only 9% of misreported net receipts.[19]

As with net receipts and total ordinary business income, total deductions from ordinary business income were misreported at a higher rate by smaller S corporations. The smallest S corporations reported 29% of all deductions for TY2004, but accounted for 53% of total misreported deductions.

The majority of misreported officers' compensation, 73% for TY2004, was misreported by small S corporations despite accounting for only 41% of total reported officers' compensation. Small S corporations underreported officers' compensation by 11% in TY2004. Only 2% of total net misreported officers' compensation was misreported by the largest S corporations with assets over $10 million, although these corporations accounted for 15% of the total reported officers' compensation.

S corporation and Sole Proprietor Misreported Net Profits (Loss)

S corporations may have compliance characteristics similar to unincorporated sole proprietors who file Schedule Cs with their Forms 1040. Most S corporations are closely held with only one or two shareholders and are typically small businesses. Smaller S corporations may include less formal businesses, with less formal recordkeeping, and relatively more cash transactions. The NRP S corporation study provides some support that compliance of the smallest, closely held S corporations is similar to the compliance of sole proprietors. Table 6 shows TY2004 compliance estimates for S corporations with assets less than $200,000 by number of reported shareholders. S corporations with only one or two shareholders misreported a larger percentage of their ordinary business income than S corporations with three or more shareholders. There

was no detected difference in the net misreporting percentage of smaller S corporations with one shareholder versus two shareholders.

Table 6: S Corporation Reporting Compliance by Number of Reported Shareholders for Tax Year 2004: Reported Assets less than $200,000
-- Amounts are in $ Billions and Counts are in Thousands[1]

Number of Shareholders	Returns that Reported	Amount Reported	NMA	Share of Amount Reported	Share of NMA	NMP	Error Rate	Relative Margin of Error NMA	Relative Margin of Error NMP
One									
Net Receipts	1,404	403	5	57%	49%	1%	32%	58%	58%
Total Income	1,445	255	8	56%	58%	3%	43%	28%	28%
Officer's Compensation	791	38	-6	56%	69%	-14%	29%	21%	21%
Total Deductions[2]	1,562	216	9	53%	67%	4%	66%	23%	25%
Ordinary Income[2]	1,562	39	17	73%	62%	28%	72%	14%	18%
Two									
Net Receipts	633	215	5	30%	44%	2%	28%	65%	62%
Total Income	651	140	4	31%	32%	3%	42%	45%	42%
Officer's Compensation	370	19	-2	28%	26%	-11%	21%	32%	32%
Total Deductions[2]	705	126	4	31%	29%	3%	62%	32%	33%
Ordinary Income[2]	711	14	8	26%	30%	28%	67%	25%	27%
Three or more									
Net Receipts	150	91	1	13%	7%	1%	28%	70%	74%
Total Income	158	63	1	14%	10%	2%	33%	70%	78%
Officer's Compensation	87	10	0	15%	5%	-4%	11%	145%	144%
Total Deductions[2]	171	63	1	16%	5%	1%	49%	38%	51%
Ordinary Income[2]	174	0	2	1%	7%	16%	55%	51%	76%

(1) TY04 estimates based on pooling the TY03 and TY04 results from the NRP Tax Year 2003/04 S corporation reporting compliance study. Estimates are based on the raw findings from the NRP study and are not adjusted for noncompliance that was not detected during the audit. Relative margins of error are based on 95% level of confidence.

(2) NMA does not include misreported officer's compensation.

Table 7 compares the smallest S corporations with only one or two shareholders to the compliance of sole proprietors, estimated using results from the NRP TY2001 individual reporting compliance study.[20] Since sole proprietors may file multiple Schedule Cs, a composite Schedule C was created by combining all Schedule Cs for each Form 1040. Table 7 further

splits returns that should have reported a net profit (referred to as "Per Exam" profit) versus returns that should have reported a net loss. Schedule C gross income (income before deductions) and total expenses were misreported at different rates than the equivalent S corporation components of net profit or loss (referred to as ordinary business income or loss on the Form 1120S). However, net profit or loss of returns that should have reported a profit was misreported at essentially the same rate, about a quarter of their net profits, by S corporations and sole proprietors. TY 2001 sole proprietors that earned a net profit had an NMP of 27% for their net profit, compared to TY2004 S corporations which had an NMP of 29%. These two types of small businesses have different net misreporting percentages for gross income and total expenses, but similar overall misreporting of net profits. This observation suggests that net profits may be a significant factor determining the total amount of noncompliance. In other words, the net misreporting percentage of gross income and total expenses may be determined to an extent by the profitability of the business.

Table 7: Net Misreporting Percentages of TY2004 S Corporations with fewer than $200,000 of reported assets and 1 or 2 shareholders and TY2001 Sole Proprietors (Schedule C)[(1)]

	"Per Exam" Profit		"Per Exam" Loss		Total	
	S corp. TY04	Sch. C TY01	S corp. TY04	Sch. C TY01	S corp. TY04	Sch. C TY01
Gross Income	3%	9%	1%	-7%	3%	8%
Total Expenses	4%	12%	2%	14%	4%	12%
Net Profit or Loss	29%	27%	20%	27%	28%	27%

(1) TY04 estimates based on pooling the TY03 and TY04 results from the NRP Tax Year 2003/04 S corporation reporting compliance study. TY01 estimates based on results from the NRP TY2001 individual filer study. Estimates are based on the raw findings from the NRP studies and are not adjusted for noncompliance that was not detected during the audit.

Conclusions

This paper presents a methodology for pooling the samples from the two years of the NRP S corporation reporting compliance study and reasonably interpreting the pooled data as representing a single tax year. S corporations are estimated to have underreported their total income by $56 billion in TY2004, representing about 14% of the total income that should have been reported for those years. Smaller S corporations with assets under $200,000 misreported a higher percentage of their income than larger corporations. The closely held smaller S corporations with only one or two shareholders misreported their net profits by 29%, a similar percentage as TY2001 sole proprietors that file Schedule C, who misreported their net profits by about 27%. In addition, the misreporting of officers' compensation, about $12 billion in TY2004, would result in underreported employment taxes.

The IRS has moved to rolling samples for the individual NRP reporting compliance study, meaning smaller samples are being selected for several continuous years. The latest individual study began in TY2006 and is scheduled to proceed for at least three years with about 13,000 returns per year. The strategy behind the more frequent but smaller samples is to receive more timely data and combine multiple years to have estimates of comparable quality as the single year larger studies. Since the S corporation study spans two years, this study provided experience working with multi-year studies and will inform the strategy for analyzing the individual rolling sample.

Endnotes

[1] See Brown and Johns (2007) for an overview of NRP underreporting studies including the decision to move towards rolling samples for the individual underreporting studies and some of the expected challenges.

[2] Distributions of income in excess of the shareholder's basis are generally treated as sales of business assets and subject to taxation as capital gains. In some cases where the S corporation was previously a C corporation and carried over accumulated earnings and profits, the distribution in excess of basis is treated as dividend income for the shareholder and subject to additional tax.

[3] Shareholders may be individuals (married couples count as a single shareholder) that are not nonresident aliens, certain estates and trusts, and certain exempt organization. S corporations may not be insurance companies, banks using the reserve method of accounting, or certain other specific types of corporations. S corporations are further required to have only one class of stock.

[4] S corporations report their ordinary business income or loss, including gross receipts and expenses, on Form 1120S. The ordinary income or loss from the front of the Form 1120S is carried over to the Schedule K and reported along with additional portfolio income (e.g. interest income, rental income, and capital gains) and portfolio deductions (e.g. corporate charitable deductions and the Section 179 depreciation deduction).

[5] Plesko (1994) discusses the taxation of S corporations compared to C corporations and partnerships and finds that the tax treatment of S corporations plays a large and significant role in the type of business entity chosen. Plesko (1995a) discusses how tax law changes in 1982 and 1986 increased the incentives to incorporate or convert to an S corporation. Plesko (1995b) concludes that the increase in the growth of S corporations following TRA86 was tax-motivated. Slemrod (1995) mentions that TRA86 lowered the top individual income tax rate to below the basic corporation income tax rate for the first time in U.S. history.

[6] See Alexander (2001) for an overview of the history of rolling samples and the influence of Leslie Kish.

[7] Analysts have suggested various weighting schemes for separate aggregate estimates of a rolling sample, including: weighting each year equally, weighting each year proportional to its sample size, or using some mechanism to provide additional weighing to the most recent sample. The unequal weighting of the two NMP estimates here reflects the relative sample sizes. Since the samples were of unequal sizes with the TY2004 sample just over three times the size of the TY2003 sample, extra weight is given to the TY2004 estimate.

[8] See Tersine and Asiala (2007) for the ACS multi-period estimates methodology. See Beaghen and Weidman (2007) for additional details on the statistical issues and interpretation of the ACS multi-period estimates. Tersine and Asiala (2007) note several benefits of the pooled approach (e.g. added flexibility in weighting schemes, improved accuracy because fewer post-strata need to be collapsed, and the ability to use the same production procedures for the pooled sample and the annual samples).

[9] Audit codes are assigned during processing of the return and reflect various characteristics that may indicate the presence of a compliance issue. S corporations with more than $10 million in reported assets and without any audit codes were further stratified by whether or not they were involved in credit intermediation or investment banking, according to their North American Industrial Classification System (NAICS) codes.

[10] About 70% of the exclusions were due to intentional oversampling by NRP to ensure the sample was of sufficient size and were never intended to be part of the final sample. About 14% of the cases were excluded because the corporation operated in a designated disaster area or the shareholders experienced some other hardship such as illness or death. About 8% of the cases were excluded because the return was already selected for examination and NRP was unable to obtain the return. The remaining cases were excluded because the taxpayer could not be located, the return could not be obtained by NRP, or the taxpayer refused to cooperate. Although it can be assumed that most exclusions were excluded at random, the nonrandom exclusions could bias the estimates. In order to adjust for

the bias, post-stratification weights were developed by IRS statisticians based on region, major industry code, and reported ordinary income.

[11] The Income Growth Factor (IGF) and Expense Growth Factor (EGF) would also reflect any changes in reporting compliance across the two years. For example, if the same average income was earned in both years but reporting compliance declined, then the IGF would reflect the change in reporting compliance and not a change in average earned income.

[12] Note that these analyses are based on the raw findings from the NRP study and are not adjusted for noncompliance that was not detected during the audit. Therefore they cannot be used to derive tax gap estimates. It is also important to point out that misreporting of income at the entity level may not necessarily result in misreporting of income and tax by the shareholder.

[13] The two errors may not always exactly offset one another at the shareholder level. For example, in some cases where the shareholder is unable to claim a loss because of basis, passive activity, or at risk limitations, the shift of income from Schedule E flow-through income to wages may not offset one another. The increased officers' compensation reduces the loss at the entity level but not at the shareholder level in this scenario.

[14] The 95% confidence interval in this example is calculated as 14.6% +/- 2.5 percentage points. The margin of error (2.5 percentage points) is calculated as 0.17*14.6 percentage points.

[15] Total income here is equivalent to the M1 Income/Loss Reconciliation line on the Schedule K. This is ordinary business income plus other portfolio income (e.g. Rental income, interest, and capital gains) net of Schedule K deductions (Charitable deductions, section 179 depreciation deduction, and other portfolio deductions). For returns not required to report this line, the equivalent amount was calculated.

[16] Section 179 allows certain businesses to deduct all or part of the cost of certain types of investment in the year in which the property is put to use. The section 179 deduction is used in place of depreciating the property and is limited.

[17] Whenever net receipts are underreported, the cost of goods sold also tends to be underreported. This leads to large estimated amounts of underreported cost of goods sold and overreported cost of goods sold, with a relatively smaller net misreported cost of goods sold in absolute terms. For example, cost of goods sold are estimated to be overreported by $13 billion and underreported by $8 billion with a net misreported amount of $5 billion.

[18] "Other Deductions" reported on Line 19 were assigned to one of 40 different categories, including automobile and truck expenses, meals and entertainment, utilities, printing and office supplies, insurance, commissions, etc. Many of the expense categories were only reported by a few returns and are therefore not shown in Table 4. Two of the expense categories are not very informative. "Miscellaneous" expenses was used whenever the corporation wrote the word miscellaneous and "other" expenses was used for any expenses that did not fit in one of the 40 categories. Also, some returns did not attach an itemized listing of expenses. Those expenses are not included in the categorization shown in Table 4; however they are included in the overall total shown in Table 3.

[19] Given the small sample sizes for the largest S corporations, even after pooling the data, caution should be used before drawing conclusions based upon the results for the S corporations with more than $10 million in assets.

[20] The NRP TY2001 sole proprietor compliance estimates for Schedule C are based only what examiners detected in the study. Historically, IRS tax gap estimates account for the fact that examiners are not able to detect all of the noncompliance during an audit, even an audit of the extent and quality of an NRP audit. Therefore the estimates of Schedule C noncompliance here are significantly less than previously published tax gap noncompliance estimates. The IRS does not have equivalent estimates for undetected S corporation income, therefore the "raw" unadjusted Schedule C compliance estimates are used to increase the comparability of the estimates from the two studies. The hypothesis is that examiners are able to detect noncompliance of the smallest, closely held S corporations at the same rate as sole proprietors with Schedule C income.

References

Alexander, Charles H. (2001) "Still Rolling: Leslie Kish's "Rolling Samples" and the American Community Survey." *Survey Methodology* 28, No. 1, 35–41.

Beaghen, Michael and Lynn Weidman (2007) "Statistical Issues and Interpretation of the American Community Survey's One-, Three-, and Five-Year Period Estimates." 2007 Joint Statistical Meetings, Section on Survey Research Method: 3034-3041.

Brown, Bob and Drew Johns (2007) "National Research Program—Methods and Plans." Paper presented at the 2007 IRS Research Conference, Washington, D.C. (June 2007)

Plesko, George A. (1994) "Corporate Taxation and the Financial Characteristics of Firms." *Public Finance Quarterly* 22, No. 3 (July 1994): 311-334.

Plesko, George (1995a) "Gimme Shelter? Closely Held Corporations Since Tax Reform." *National Tax Journal* 48, No. 3 (September 1995): 409-416.

Plesko, George (1995b) "The Role of Taxes in Organizational Choice: S Conversions After the Tax Reform Act of 1986." 1995 meeting of the American Accounting Association, Orlando, FL, (August 1995).

Slemrod, Joel (1995) "Income Creation or Income Shifting? Behavioral Responses to the Tax Reform Act of 1986." *The American Economic Review* 85, No. 2, Papers and Proceedings of the Hundredth and Seventh Annual Meeting of the American Economic Association Washington, DC, January 6-8, 1995 (May 1995): 175-180

Tersine, Anthony G. and Mark E. Asiala (2007) "Methodology for the Production of American Community Survey Multiyear Estimates." 2007 Joint Statistical Meetings, Section on Survey Research Method: 3013-3023.