The IRS
Research Bulletin

Proceedings of the 2007 IRS Research Conference

Research
Analysis
Statistics

National Research Program— Methods and Plans

Bob Brown and Drew Johns, IRS Office of Research

I n 2000, the Internal Revenue Service (IRS) established the National Research Program (NRP) office as part of its efforts to develop and monitor strategic measures of taxpayer compliance. NRP now provides compliance measures for different types of taxes and different groups of taxpayers, with the three major forms of compliance being:

- filing compliance—measured as the percentage of required returns timely filed;

- payment compliance—measured as the percentage of reported tax liability on timely-filed returns timely paid; and

- reporting compliance—measured as the accuracy of income, deductions, expenses, and other items reported on timely filed returns.

Each form of noncompliance contributes to the overall tax gap, which estimates the difference between what the IRS estimates is due to the Federal Government and what the Government actually collects. Using NRP data, the IRS estimated the Tax Year 2001 Gross Tax Gap to be $345 billion. It is expected that various enforcement activities and voluntary late payments will bring in $55 billion, leaving the Net Tax Gap at $290 billion. Misreporting is by far the largest component of the Tax Gap, accounting for nearly 83 percent of the TY 2001 Gross Tax Gap.

This paper focuses on the IRS's efforts to measure reporting compliance, examining the benefits and challenges to using reporting compliance data to support strategic and operational decisionmaking.

Background

Measuring taxpayer compliance for the IRS is a basic measure of the Service's effectiveness in administering the Tax Code. It is analogous to measuring the net profit for a private sector business.[1] Just as a company's management would have an incomplete picture of the organization's opera-

---

[1] "IRS's Comprehensive Approach to Compliance Measurement," *National Tax Journal*, Volume LVI, Number 3, September 2003, p. 689.

tions without computing its net profits, the IRS would have a distorted view of its operations without up-to-date measures of taxpayer compliance.

A decade ago, the Internal Revenue Service and outside observers recognized a major hole in understanding its operations because the data driving its compliance measures were increasingly outdated, due to a break of at least 10 years for all but a few measures of compliance.[2] One implication of this break was that IRS compliance and tax gap estimates were becoming less reliable because IRS researchers were basing these estimates on increasingly dated raw data.

The Evolution of IRS Reporting Compliance Studies[3]

The IRS established a Taxpayer Compliance Measurement Program (TCMP) in 1964, starting with an examination of individual income tax returns for Tax Year 1963 returns, and, for nearly 3 decades, the TCMP had been one of the IRS's most important compliance initiatives.[4] The most prominent feature of this program was the series of individual income tax compliance surveys, which consisted of thorough audits of representative samples of individual income tax returns approximately every 3 years. However, other groups of taxpayers were also subject to reporting compliance studies during that time.[5]

Over the years, the IRS conducted periodic TCMP studies on all major individual and business return types. The size of the samples differed, generally reflecting the size of the taxpayer populations covered by the studies. The TCMP studies of individual income tax returns used samples of approximately 50,000-55,000 returns for a given tax year.[6] Compliance studies for other groups of taxpayers generally required smaller samples. Each study shared certain basic characteristics. They all utilized randomly selected samples of returns designed to statistically reflect the overall filing population, and each return was subjected to line-by-line scrutiny, with taxpayers

---

[2] National Commission on Restructuring the IRS, final report, paragraph 83, June 1997.

[3] We based the first part of this section on a comprehensive review of the Taxpayer Compliance Measurement Program, most of which are described in internal IRS documents. Publicly available documents include "Income Tax Compliance Research," Internal Revenue Service, Department of the Treasury, 1988; "Taxpayer Compliance Measurement Program Handbook," Internal Revenue Service, Department of the Treasury, May 1989; "Compliance Data," Government Accounting Office, Report #GGD-96-89, April, 1996; "Status of IRS's Efforts To Develop Voluntary Compliance Measures," Government Accounting Office, Report #01-353, June, 2001; "New Compliance Research Effort Is On Track, but Important Work Remains," Government Accounting Office, Report #02-769, June, 2002; and "A Survey of Random Audit Programs in OECD Countries," OECD Forum on Tax Administration, November 2002.

[4] Ibid.

[5] Ibid.

[6] Ibid.

generally required to document the entries on each line item on the returns. The IRS attempted to use its more experienced examiners in conducting TCMP audits. TCMP examinations were also subject to more intensive quality reviews than operational audits (and the returns were corrected for math errors) to ensure that the auditors identified the correct amounts of tax due on the returns.

The purpose of this series of studies was twofold. First, the IRS expected the studies to provide an estimate of overall reporting compliance for various individual groups of taxpayers, as well as for the entire tax system. Second, and perhaps more importantly, the studies sought to develop improved ways to select returns for examination.[7] By developing data on the characteristics of returns that exhibited a degree of noncompliance, the IRS was able to focus resources on examinations that made sense from an agency viewpoint. Prior to the development of the Discriminant Function Score (based on these compliance studies), IRS experience was that about half of all audits resulted in "no tax change," which wasted both IRS and taxpayer resources. After these workload selection formulas were developed and widely utilized, IRS examinations based on these formulas showed around one-fifth of the examinations resulting in "no change."[8]

Besides developing workload selection formulas, these studies led to several policy changes. The observation that taxpayers more accurately reported items on their returns for which the IRS received reports from third parties relative to items for which third-party reporting was unavailable over time increased the pressure to require information reporting for a variety of items.[9] In addition, the observation that numerous taxpayers appeared to be misreporting claims for dependent exemptions led to a requirement that taxpayers report a taxpayer identification number (generally a Social Security number) for each dependent claimed (e.g., child). The result of this policy change was that several million dependents "disappeared" the year after this requirement was put into law.[10]

However, the IRS discontinued these reporting compliance efforts after the Tax Year 1988 TCMP study.[11] The IRS attempted to restart the program but could not develop an approach that met the objections of policymakers in

---

[7] "New Compliance Research Is on Track, but Important Work Remains," Government Accounting Office, Report #02-769, June 2002, p. 5.

[8] Ibid.

[9] Mortgage interest, proceeds from stock sales, and proceeds from the sale of homes are prominent examples of expanded information reporting that resulted from various TCMP studies.

[10] "Where Have All the Dependents Gone?," *Trend Analyses and Related Statistics—1990 Update*, Department of Treasury, Internal Revenue Service, Publication 1500, August 1990.

[11] "Compliance Data," Government Accounting Office, Report #GGD-96-89, April 1996, p. 1.

both the Executive and Legislative Branches. One cannot point to a single reason for the demise of TCMP; rather, it appears to have resulted from a confluence of a number of different factors. These factors included concerns about the burden placed on taxpayers, the amount of IRS resources devoted to TCMP, an unclear connection between the costs of TCMP and the resultant benefits (to the IRS and to the taxpaying public), and a general negative feeling toward the IRS on the part of Congress and a substantial portion of the American public.[12]

With each passing year, the IRS's risk associated with the aging data grew. Not only was the IRS making compliance decisions using data that were, in some cases, more than 20 years old, but the populations of many types of tax returns had also changed dramatically in number and composition since the last concerted effort to measure reporting compliance of those taxpayers.

To address the shortfall in information about taxpayer compliance that is required to run the IRS effectively, then-Commissioner Charles Rossotti established the predecessor to the National Research Program in April 2000.[13] In establishing the infrastructure for these measures, the IRS identified four primary goals:

- gather strategic information about taxpayer compliance behavior that will allow the IRS to better allocate its resources to enforcement and other activities;

- update workload selection formulas;

- collect data that will provide some insight into the causes of reporting errors (to aid in providing taxpayer service); and

- develop those data to update IRS estimates of the tax gap.

The IRS adopted a philosophy that seeks to balance taxpayer burden with the acquisition of high-quality compliance data from NRP reporting compliance studies. To achieve this balance, NRP made several departures from the TCMP methodology. The first departure was a greater reliance on data already available to the IRS, a process called "case-building" in the terminology of NRP.[14] A second departure was the classification of NRP returns. In TCMP studies, IRS examiners verified entries on every line item

---

[12] The Senate Finance Committee conducted a series of widely publicized hearings in the mid-1990s that spotlighted many of these negative feelings toward the IRS.

[13] "New Compliance Research Effort Is on Track, but Important Work Remains," Government Accounting Office Report #02-769, June 2002, p. 1.

[14] In building an NRP case, the IRS assembles information from a variety of internal and third-party data sources. This information allows the IRS to substantiate some or all of the information on a tax return.

on each return in the sample. In contrast, NRP employs a well-trained group of experts to analyze the sampled returns, and associated case building materials, and then list those items requiring taxpayer verification.

The ability of NRP to maintain the integrity and security of the data it gathers and analyzes is critical to its success. Throughout the NRP process, the IRS maintains taxpayer privacy by collecting only taxpayer data necessary for compliance measurement. It is critical not only from the perspective of data integrity, but the NRP must be able to ensure stakeholders that access to and use of these data are only carried out under established IRS guidelines and procedures for securing and using taxpayer data. Processes are in place to ensure accountability and appropriate access and use of data related to NRP.

The initial NRP reporting compliance study used a random sample of approximately 46,000 individual income taxpayers filing returns for Tax Year 2001. Using the case building and classification processes, NRP accepted the reported amounts on nearly 3,000 returns. NRP verified tax return information on an additional 1,800 returns via correspondence. One of the many positives to come out of this initial study was the positive reactions that resulted from the IRS's attention to taxpayer burden. Customer satisfaction survey data indicate that taxpayers participating in NRP audits had satisfaction scores significantly higher than taxpayers going through operational audits.

Despite the successes, the initial reporting compliance study revealed a number of weaknesses, the primary one being the burden such studies imposed on the IRS workforce. Between 1988 and 2003, the IRS workforce declined by approximately 30 percent. As a result, the addition of nearly 50,000 research audits severely taxed the operating divisions' ability to absorb the additional work. In planning for the second return type, NRP tweaked its methodology more by moving to a multiyear study.

In an effort to accurately assess the tax gap and the reporting compliance of flowthrough business entities, the IRS first designed a pilot reporting compliance study of both partnerships filing Form 1065 and corporations that file Form 1120S (more commonly known as S corporations). The results of the pilot study pointed the IRS to a more complete reporting compliance study focusing on just S corporations, beginning with entities filing returns for Tax Year 2003. The study of S corporations contributes to IRS's overall strategic goals and its research compliance measurement efforts, and marks the first time in over 2 decades that the IRS has reviewed the reporting compliance of flowthrough entities.

The new study will provide information on reporting compliance, data to estimate the tax gap, and aid to policymakers in their decisions regard-

ing tax policy, as well as assist IRS in developing priorities for taxpayer education programs to help reduce noncompliance. The S corporation study involves a sample of approximately 5,000 returns filed for either Tax Years 2003 and 2004. About 25 percent of the sample is Tax Year 2003 returns, while the large majority of returns come from Tax Year 2004. NRP delivered the 2003 returns to the field for examination in October 2005. Shipments of 2004 returns for examination started in April 2006.

This study follows the NRP methodology of having experienced IRS revenue agents review the cases and classify specific issues for field examiners to pursue during their audits. With no information documents to validate entries on S corporation tax returns, the examiners are looking at more issues than they did for the Form 1040 reporting compliance study. Still, the IRS will not likely ask for documentation for every line item on any of the selected S corporation returns.

As of the end of July 2007, the IRS had contacted every taxpayer in the sample, starting the audit process. IRS examiners had completed their audits in more than 4,500 cases. With the vast majority of cases near completion, NRP expects to close its study database in September 2007 to release the initial database for research purposes.

The IRS based its decision to study S corporations on two primary considerations. First, the population of S corporation returns had grown rapidly in the past 20 years, and, in Tax Year 2003, S corporation returns accounted for 61.9 percent of all corporate tax return filings. The nearly fivefold increase in the number of S corporations since the last compliance study of 1120S returns in 1984 makes this population an especially important focus for an NRP study. Using the data obtained from this NRP study, the IRS will be able to:

- determine various measures of compliance including the Net Misreporting Percentage for S corporation Ordinary income,

- determine the effectiveness of audit selection techniques, and

- provide data to refine the tax gap estimate.

Secondly, the IRS had conducted no compliance measurement audits for S corporations since the Taxpayer Compliance Measurement Program (TCMP) involving Tax Year 1984 returns. The IRS currently conducts few operational audits of S corporations; for returns filed in Calendar Year 2003, the IRS examined 6,402 S corporation returns. This is an audit rate of less than 0.2 percent of the total returns filed for Form 1120S.

To calculate the impact on the tax gap estimates, NRP examiners will not only evaluate the S corporations' tax returns, but also assess the related

tax returns of individuals and entities that are shareholders in the selected S corporations to determine whether the actual tax liability is correct. Compliance studies of flowthrough entities such as S corporations are complex primarily because they report only income and losses. Those who have ownership share in the profits and losses of the corporation and should pay the appropriate tax liability.

Consequently, for every S corporation, at least two tax returns should be examined to determine whether the correct tax has been paid—the Form 1120S of the business plus the individual returns of shareholders. In addition, the IRS will assess whether the S corporations qualify to file these types of returns.

In looking forward to the upcoming iteration of the individual reporting compliance study, NRP will take the multiyear sample approach one step further. The study will start in October 2007, using approximately 13,200 randomly-selected Tax Year 2006 returns. The study will use similar sample sizes in subsequent years. Combining audit results from the first 3 years' samples (roughly 39,600 returns) will provide compliance estimates of comparable statistical precision to those from the Tax Year 2001 NRP study with a sample size of more than 46,000 returns. Unlike previous reporting compliance studies, the new NRP approach will not stop after results from the TY 2009 audits are in; rather, NRP will continue to conduct audits of annual samples.

An advantage of using this method, which combines results over rolling 3-year periods, is that the IRS could then update compliance estimates and develop more efficient workload plans on an annual basis by adding each new year's compliance results to the 2 prior years of existing data. With previous studies, which started from scratch, tax returns were drawn from a single year, and involved examinations of more than 45,000 taxpayers; and, the IRS faced gaps of at least 3 years before updating its compliance estimates.

Benefits and Successes of the NRP Methodology

The IRS used data from the 2001 Form 1040 reporting compliance study, along with updated payment and filing compliance data, to develop a new point estimate of the tax gap. The Service is using the NRP data to drive research and other initiatives to narrow the tax gap in the future. IRS researchers have already examined NRP data to assess compliance characteristics of specific taxpayer segments, such as high-income individuals.

The IRS shares aggregate findings from NRP studies with other Government agencies, facilitating a greater understanding of taxpayer incomes. For example, the Service provided tabulations of Schedule C and C-EZ data to the Bureau of Economic Analysis for use in adjusting various national economic indices.

Each unnecessary or ineffective contact with a compliant taxpayer has a direct and negative impact on the public's perception of the effectiveness and fairness of the Federal tax system. Additionally, compliant taxpayers want to know that the IRS is capable of ensuring that everyone pays his or her fair share. The IRS uses a variety of techniques, such as document matching, correspondence, and audits, to verify that taxpayers accurately report their tax liabilities on their returns. Optimally, the IRS audits those returns most likely to have errors, and, although the Service uses various methods to identify errors, the most efficient method involves workload selection formulas used to select returns for examination.

Prior to the 2001 NRP reporting compliance study, the last such exercise occurred with Tax Year 1988 returns. The advanced age of the workload selection models made it increasingly difficult to identify, for audit purposes, returns with underreported tax liability. With the release of Tax Year 2001 NRP reporting compliance data, the IRS was able to update its workload selection formulas for the first time in about 15 years. New formulas went into service in January 2006, and the IRS made minor updates for implementation in 2007.

The IRS is committed to applying its limited resources where they are of most value in reducing noncompliance while ensuring fairness, observing taxpayer rights, and reducing the need to burden those who do comply. The service will use NRP results to more effectively manage its compliance programs and design prefiling activities that help taxpayers comply with the tax law. Continued updates of quality reporting data is vital to this strategy. The IRS is launching a new NRP reporting compliance study for Form 1040 filers that will use an innovative multiyear rolling sampling methodology. The advantage of using this approach, which combines results over rolling 3-year periods, is that the IRS can make annual updates to compliance estimates and develop more efficient workload plans on an annual basis, after the initial three annual studies. The new approach also uses smaller samples in any given year, reducing the burden of research on both the taxpayer population and IRS examiner resources.

The launching of NRP reporting compliance studies has had a number of indirect benefits to the IRS. These include refinements to examination tools and processes; adoption of new technologies, such as virtual servers

and Web-based support tools; and advances in the development and delivery of training.

Other important benefits of NRP, to both the IRS and taxpayers, include:

- taxpayers benefit from redesigned forms, improved communications, suggested law changes, and enforcement focused on noncompliant taxpayers. The Service will have improved ability to do this because NRP examiners determine likely reasons for observed noncompliance and NRP data will identify line items on tax returns that create the most difficulty for taxpayers trying to comply with tax law. NRP continues to drive IRS efforts to refine identification of the reasons for noncompliance. In Calendar Year 2006, NRP spearheaded a multifunctional initiative within the IRS to review and revise "reason codes" used to describe why taxpayers misreport incomes, deductions, taxes, and credits;

- taxpayers experience fewer burdens during an audit because NRP pioneered a comprehensive case-building approach that provides examiners with a wide range of information about the taxpayer's situation before an audit takes place. This allows the examiner to focus more on questionable items and potentially improve audit productivity;

- NRP also pioneered a number of other processes that improve audit productivity. NRP returns were the first to use the IRS's application for capturing examination results in a LAN environment, facilitating the sharing of case file information and reducing the time associated with reviewing and closing cases. NRP also automated the classification process and the population of tax return data into the report-generating system, eliminating thousands of hours associated with the creation of audit case files; and

- obtaining a measure of overall income tax compliance allows the IRS to measure its "bottom line" and make resource allocation decisions in ways that can improve its overall performance. At the 2004 IRS Research Conference, keynote speaker Senator Charles Grassley, then chairman of the Senate Finance Committee, pointed out the importance of NRP studies to guide Congress in its tax policy.

NRP Individual Underreporting Study Tax Years 2006, 2007, and 2008: Analytical and Methodological Challenges

NRP data provide the basis for a broad range of analyses, including workload selection formulas, burden and service analysis, detailed compliance measures, and estimating the individual underreporting portion of the overall gross tax gap. The move to smaller annual studies promises benefits such as updates to these analyses, but challenges lie ahead. This section highlights some of the more significant challenges, including dealing with a smaller sample size in early years and pooling data across multiple years, and presents some early thinking on how we may handle those challenges.

The Tax Year 2001 NRP Individual Underreporting Study consisted of a sizeable sample of approximately 46,000 returns, which provided enough observations for updating workload selection formulas and estimating the amount of income that was not detected by classifiers and examiners. The current multiyear individual underreporting study proposes selecting approximately 39,600 returns in equal thirds over Tax Years 2006, 2007, and 2008. The design of the sample considers the needs of workload selection and the overall requirements of estimating the raw Voluntary Reporting Rate (VRR) and the EITC compliance rate within 2 percentage points at a 95-percent level of confidence.[15] The requirements are met based on pooling 3 years of data into one larger sample, meaning that updates to these analyses may not be available in the first 2 years or will be based on a lower level of confidence.

Detection-Controlled Estimation

Estimating the individual income underreporting tax gap has typically consisted of three components: 1) estimates based on errors detected by examiners during random audits, 2) adjustments for unreported income that examiners were unable to detect and 3) average marginal tax rates applied to the components of income and offsets to income. Prior to estimating the Tax Year 2001 tax gap, estimates of the amount of income not detected during the random audits consisted of multipliers based on a comparison of

---

[15] The VRR is generally defined as the ratio of tax reported on timely filed returns to the amount of tax that should have been reported. The raw VRR is based on raw noncompliance detected during NRP audits and does not take into consideration noncompliance that classifiers and examiners were unable to detect and should not be confused with the overall tax-gap-based VRR. The EITC compliance rate is defined as the ratio of the amount of EITC that should have been claimed to the amount of EITC that was claimed.

1976 TCMP audit results where examiners did not have use of information reporting (IRP) documents with the income reported on those documents. The results of the comparison showed that, for every $1 detected without the use of IRP documents, another $2.28 went undetected. This resulted in the use of a 3.28 multiplier, with some variations depending on type of income.[16] For purposes of estimating the Tax Year 2001 tax gap, IRS wanted to update the methodology for estimating undetected income and contracted with Dr. Brian Erard and Professor Jonathon Feinstein to implement a methodology originally developed by Professor Feinstein known as Detection-Controlled Estimation (DCE).[17]

DCE is an econometric technique that when applied to tax compliance typically estimates two equations jointly, 1) a noncompliance equation which models the amount of detected and undetected underreported income and 2) a detection equation that models the detection rate. The intuition underlying the DCE methodology is the idea of modeling the differences in the abilities of examiners to detect income. Because examiners play a significant role in the methodology, ideally, there would be a relatively large group of examiners who each examined a large number of returns and line items on those returns. If, instead, there are only a small number of examiners who audited multiple returns, then there may be challenges with using the DCE methodology. In other words, the allocation of returns to examiners and the extent of the audits play an important role in determining the level of detail and reliability of the estimates that DCE can provide.

The decision to spread the sample over multiple years has implications for implementing the DCE methodology. For Tax Year 2006, approximately 13,200 returns will be selected. Following past NRP procedures, these returns will go through classification and essentially be assigned to one of three categories: 1) accepted as filed or with adjustments, 2) correspondence exam, or 3) face-to-face exam. The first implementation of the DCE methodology for estimating the Tax Year 2001 tax gap was limited to using the results of the face-to-face exams. Assuming the same scenario holds for future analyses, fewer than the 13,200 selected returns will actually be available to use with DCE. Ideally, these returns will be efficiently assigned to a cadre of examiners who will primarily handle NRP audits, thereby potentially increasing the reliability of DCE by concentrating the returns and lines examined. By moving toward annual studies, essentially institutionalizing the NRP individual studies, the likelihood of developing a cadre of examin-

---

[16] Internal Revenue Service (1996), *Federal Tax Compliance Research: Individual Income Tax Gap Estimates for 1985, 1988, and 1992*, Publication 1415 (Revised 4-96), Washington, DC.

[17] Feinstein, Jonathan S. (1991), "An Econometric Analysis of Income Tax Evasion and its Detection," *RAND Journal of Economics*, Volume 11, Number 1 pp.14 – 35.

ers may be increased. Even with a greater concentration of returns among examiners, there may not be enough observations to employ DCE during the first or second years of the study. However, if the annual studies are extended beyond Tax Year 2008, then annual updates with DCE methodology may be possible by pooling the previous 3 years of NRP results.

Workload Selection

In addition to estimating the tax gap, NRP data are essential for updating workload selection formulas used to target audits toward taxpayers who have a greater likelihood of owing taxes. By improving the targeting of audits, IRS increases the collection of underreported income taxes and reduces the burden on compliant taxpayers. Updating the workload selection formulas also imposes requirements on the NRP sample methodology. As with DCE estimation, the smaller number of returns selected in the first and second year may delay the updates to workload selection formulas until the third year, by which time pooling the data will provide sufficient observations. Continuing with annual studies past Tax Year 2008 would then allow annual updates to the workload selection formulas by pooling the previous 3 years of NRP results.

Other Analyses

Raw NRP data are also used to identify areas that may be more susceptible to taxpayer errors because of opportunities for noncompliance or because of a relatively higher degree of burden. These analyses typically rely on detailed raw NRP audit results based on errors detected by examiners at the line item level, including lines on supporting schedules. The results are used to develop and support legislative proposals such as increasing information reporting requirements and influence resource allocation, tax administration, and taxpayer service decisions. The design of the multiyear study provides a high degree of confidence in the overall raw VRR based on 3 years of pooled data and a reduced, but still relatively high, degree of confidence based on the smaller annual sample. Although the reliability of the detailed line item raw results is not a direct requirement of the design of the sample, the stratification of the sample does take into consideration the additional forms and schedules that taxpayers file with their tax returns, such as Schedule C or Form 2106. In general, measures of raw compliance for specific line items that are not reported on many returns may be less reliable based on annual results and require pooling the data over 3 years. It will be tempting with re-

sults across many line items to compare annual compliance rates across line items and years, but the validity of any comparison depends on the degree of reliability which is related to the size and design of the pooled sample.

Pooling the Data over Multiple Years

In order to provide detailed estimates, NRP data will need to be pooled across multiple years. The U.S. Census Bureau takes a similar approach with the American Community Survey (ACS).[18] The ACS is a new annual survey of demographic and housing characteristics by the Census Bureau that provides detailed estimates across geographic areas. The ACS provides annual estimates for geographic areas with 65,000 or more people, pools data over 3 years for areas with a minimum of 20,000 people, and pools 5 years of data for the smallest areas. This approach is analogous to the one likely to be used by users of NRP data. Some tax return line items may have sufficient coverage to provide annual updates, but others will require pooling the results.

The Census Bureau pools ACS data by iteratively adjusting the sample weights to meet target population data across various segments of the population. The population targets for the ACS 3-year estimates are based on averages over the pooled years. NRP data can initially be pooled in a similar manner, but NRP will most likely use the population totals for the final year of the samples that are pooled together as its targets. For example, the samples for Tax Years 2006, 2007, and 2008 can be pooled together with the stratum weights adjusted using the following process within each stratum:

> **Step 1:** Prior to pooling the data, assign each case a weight equal to the inverse of the probability of being selected, where the probability of being selected is the ratio of the number of returns sampled in the case's stratum to the stratum population. Except for further adjustments made by NRP, this is the typical weight that would initially be assigned to each case.

> **Step 2:** Multiply the weight assigned in Step 1 by the ratio of the sum of the weights within the given stratum in Tax Year 2008 to the sum of all weights for that stratum across Tax Years 2006, 2007, and 2008.

We are often interested in a point estimate of underreporting for a specific tax year. Each tax year, the individual income tax brackets, stan-

---

[18] Rivers, Kerri L.; Mark Mather; and Linda A. Jacobsen (2006), "American Community Survey: A Guide for Data Users," The Population Reference Bureau.

dard deductions, and exemption amounts creep upwards impacting effective average marginal tax rates. Alternatively, Congress may make changes to the actual tax rates. The above method would supply a pooled sample that hits the population targets for Tax Year 2008, but the actual data across tax years may need to be adjusted further to account for such issues as inflation, changes in tax law, economic changes like the distribution of income across various types of income, and other potential factors. In order to account for inflation, dollar amounts are adjusted in the ACS estimates to the most recent year by the Consumer Price Index (CPI). This approach could be used to adjust the NRP pooled data, but it does not really address the larger issue of data compatibility across tax years.

One way of handling the issue of compatibility of the results across tax years is to use the reported and corrected amounts of items like income, filing status, and number of exemptions with a detailed tax calculator that is calibrated to the tax year of analysis. The tax calculator could then be used to calculate the reported and corrected amounts of taxes, credits, deductions, and other variables whose maximum amounts and phaseout levels vary by Tax Year. The pooled data could be aged in a manner such that the reported amounts of items that feed into the calculator for Tax Years 2006 and 2007 are adjusted proportional to the reported amounts from 2008. The following steps would yield a pooled data set with the population estimates equal to the population targets for Tax Year 2008, the weighted reported amounts equal to the weighted estimates for Tax Year 2008, and the net misreporting percentages for each case and stratum maintained within each third of the pooled sample.

Step 1: Within a given tax year and stratum, calculate the total weighted reported amount for a given line item. If the weights have been previously adjusted so that the sum of the pooled weights equals the Tax Year 2008 population totals, then use the adjusted weights.

Step 2: On a case-by-case basis for Tax Year 2006, multiply the reported amount for Tax Year 2006 by the ratio of the total weighted reported amount calculated in Step 1 for Tax Year 2008 to the total weighted reported amount calculated in Step 1 for Tax Year 2006. Repeat for the Tax Year 2007 cases.

Step 3: On a case-by-case basis for Tax Years 2006 and 2007, multiply the amount that should have been reported and the net misreported amount by the same ratio used in Step 2.

Step 4: Use the tax calculator and the adjusted weights and adjusted income variables to estimate the remaining reported amounts and amounts that should have been reported for items that depend on the target year's tax laws.

The above discussion emphasizes the potential need for a microlevel pooled dataset that represents one specific tax year, which is the likely requirement for updating tax gap estimates. If all we are interested in are macrolevel estimates of noncompliance based on the raw data, then the results of each study could be averaged over 3 years. One possible method of averaging would be to calculate a weighted mean with the weights inversely proportional to the estimated variances.[19] This method implicitly assumes that the measures being averaged actually represent the same mean. Although the tax years differ and compliance may evolve over time, this assumption may be reasonable enough in order to average the results across years. Another option would be to take the simple weighted mean. A third option would be to pool 3 years' worth of data without necessarily adjusting the weights or the values if we were not concerned about estimating a specific year and only wanted the average over 3 years. Either way, the data could also be adjusted to control for inflation by using the CPI index. This method of adjusting for inflation was used by researchers in Statistics of Income (SOI) to compare selected income and tax return items across tax years.[20]

Regardless of the method of pooling the data, caution must be used when comparing results across different time periods. Even comparing results from 1 year to the next will require knowing the statistical reliability of the estimates, given reduced sample sizes for the annual studies. Comparisons of estimates from pooling 3 years of data will require particular scrutiny because time periods will overlap. For example, if the annual studies are extended to Tax Year 2009, then the 3-year estimates using Tax Years 2006, 2007, and 2008 will overlap with estimates based on pooling Tax Years 2007, 2008, and 2009. Comparing results from the two separate 3-year pools of data would essentially be comparing Tax Year 2009 to Tax Year 2006 because Tax Years 2007 and 2008 wash out. Although direct

[19] Cochran, William G. and Sarah Porter Carroll (1953), "A Sampling Investigation of the Efficiency of Weighting Inversely as the Estimated Variance," *Biometrics*, Volume 9, Number 4, pp. 447- 459.

[20] Strudler, Michael and Michael Parisi "Selected Income and Tax Items from Inflation-Indexed Individual Tax Returns, 1990-2001," *Statistics of Income Bulletin,* Volume 23, Number 4, pp. 200-212. Available at http://www.irs.gov/pub/irs-soi/01ininfl.pdf.

comparisons may be complicated, if the annual studies are institutionalized and maintained, then the pooled results will be useful for analyzing trends.

NRP S Corporation Study of Tax Years 2003 and 2004: Analytical and Methodological Challenges

NRP is currently in the process of a study of income and tax underreporting on S corporation income tax returns (Forms 1120-S) for Tax Years 2003 and 2004. Approximately 25 percent out of around 5,000 returns were selected during Processing Year 2004 for Tax Year 2003, with the remainder of the returns selected during Processing Year 2005 for Tax Year 2004. Audits of many of these cases are still ongoing. S corporations are unique in that the corporation must file an income tax return, but the majority of its income is exempt from taxation at the corporate level. Instead, the corporation's income flows down on a proportional basis to its shareholders who are then taxed at their individual rates. For Tax Years 2003 and 2004, in order to elect to be treated as an S corporation, a corporation must have been a domestic entity with no more than 75 shareholders who consist of individuals, married couples, estates, certain exempt organizations, and certain trusts. The law was changed for tax years after 2004 to allow up to 100 shareholders of an S corporation. In reality, the vast majority of S corporations have far fewer shareholders, with the average number of shareholders between one and two.

The division of the study over 2 tax years presents analytical challenges similar to those created by the new multiyear individual income tax underreporting study that began with Tax Year 2006. Each study is a composite of the tax years being sampled. The summary results of the Tax Year 2003 study could be averaged with the summary results of the Tax Year 2004 study with either a simple mean or a weighted mean where the weights are inversely proportional to their respective estimated variances. This methodology could provide an overall estimate of the net misreported amounts for various line items from Form 1120-S based on errors detected by examiners.

Flowthrough Income and Tax Gap Measurement

The income from an S corporation flows down to shareholders. Thus, to actually estimate the tax impact of the underreporting of income and overreporting of deductions by S corporations, we need to know information about the shareholders, including their individual income tax returns. The effective marginal tax rate of the underreported S corporation income will depend not only on the S corporation but ultimately on the various Form 1040 return

characteristics of the shareholders. Although aggregate macromeasures of noncompliance may be possible by averaging the results of the 2 tax years, a reliable estimate of the tax impact requires pooling the 2 tax years at the microlevel. Pooling the data requires choosing a target tax year, with Tax Year 2004 the logical choice because it is the most recent tax year and the year from which most cases were sampled. The weights could be adjusted following the procedure outlined previously in the section discussing the individual income tax multiyear study to ensure that the sum of the weights from the pooled sample matches the 1120-S population for Tax Year 2004. Dollar amounts could also be adjusted for inflation by using the CPI index or adjusted so that the weighted average income for each item from Tax Year 2003 matches the weighted average income for each item for Tax Year 2004. This method could be applied at the stratum level as described in the individual section or at the aggregate level.

Underreporting of income on Form 1120-S and underreporting of tax on Form 1040 must be linked. NRP is tracking the audits of shareholders which would supply information on the flowthrough adjustments, including the resulting change in individual income tax. However, if auditors expand the audits of the shareholders beyond the flowthrough income, then the flowthrough adjustments may need to be backed out from the remaining adjustments. We may then need to use a tax calculator calibrated to Tax Year 2004 to estimate the change in tax due to the backed-out flow through adjustments. In fact, the pooling of Tax Year 2003 returns with Tax Year 2004 returns likely necessitates the use of a tax calculator because of the changing individual income tax brackets, standard deductions, exemption amounts, and deduction thresholds, etc. between Tax Years 2003 and 2004.

One of the objectives of the S corporation study is to incorporate the results with the overall estimate of the gross tax gap. The previous discussion only pertains to the underreported income that examiners are actually able to detect, whereas the tax gap includes undetected income. The individual underreporting portion of the latest estimate of the tax gap is for Tax Year 2001 and is partially based on random audits of Tax Year 2001 individual income tax returns. The Tax Year 2001 returns included individuals who were shareholders of S corporations, and normal audit procedures were followed when determining whether to audit the actual S corporation. Therefore, in some cases, the entity corporation was audited along with the individual shareholders, and any adjustments would be reflected in the NRP raw results. Adding the results of the S corporation study directly to the Tax Year 2001 individual results would not be appropriate because of this overlap, even if the Tax Year 2001 results were projected forward to 2004.

The Tax Year 2001 tax gap estimate also includes an adjustment to Schedule E income, where S corporation ordinary, royalty, and rental incomes are reported, for income not detected during the audits. The adjustment for undetected income results in additional overlap between the Tax Year 2001 Schedule E tax gap estimate and the results of the S corporation study. The tax gap estimate for Schedule E income consists of two separate estimates: 1) rental and royalty income reported on page 1 of Schedule E and 2) the aggregate of partnership, S corporation, real estate mortgage investment conduit (REMIC), and farm rental incomes reported on page 2. Therefore, the adjustment for undetected Schedule E income also includes income other than S corporation income. Theoretically, the adjustment for Schedule E undetected income can be apportioned in the following manner: 1) S corporation income that was not detected among S corporations that were audited, 2) S corporation income that was not detected because the S corporation was not audited, 3) S corporation income that was not detected and would not have been detected had the S corporation been audited, and 4) income that was not detected that is unrelated to S corporation income.

S corporation income also flows through to the interest, dividend, and capital gain income lines on Form 1040. Some S corporation deductions like charitable contributions also flow through to Schedule A of the shareholders' individual income tax returns. As with Schedule E, identifying the different portions of overlap is one of the challenges that lie ahead with incorporating the results of the S corporation study with the overall estimate of the individual underreporting tax gap.

Summary

The National Research Program represents the commitment the IRS has made to improve the efficiency and fairness of the tax administration process. With measures of strategic compliance, the IRS will be able to make more informed strategic decisions about workload allocation, resource planning, and taxpayer communication and support. These measures will also provide a benchmark against which the IRS can measure the effectiveness of programs to promote voluntary compliance with the Tax Code.

The NRP has made significant progress in the development and delivery of its strategic measures. The IRS now has Filing Rate measures for Tax Years 1992 through 2001, and the Service is working on developing new filing compliance measures using new data sources for more recent years. Working with the Census Bureau, the IRS has developed Nonfiling

tax gap estimates for Tax Years 2001 and 2003. The IRS has also developed Voluntary and Cumulative Payment Compliance Rates for Tax Years 1999 through 2005. Mechanisms are in place to deliver these measures on an ongoing basis.

NRP successfully completed the first significant research effort to obtain estimates of reporting compliance from individual income tax returns in nearly 20 years and is close to completing the first reporting compliance study of Subchapter S corporations in nearly 25 years. Building on a new sample design methodology, NRP is embarking on an ongoing annual study of individual income taxpayers that will allow for more frequent updates to workload identification formulas and compliance estimates. The new methodology will also lead to greater opportunities for researching individual issues and line items, but it will require care in the merging of data from multiple tax years. Future reporting compliance studies will surely follow, focusing on other types of business returns, which have not been systematically studied in this way since the mid-1980s.