

Understanding the Tax Gap¹

***Abstract** - The Tax Gap is defined as the difference between the amount of tax imposed by the Tax Code and the amount that is reported and paid with timely filed returns. For the federal government, the gross tax gap is estimated at \$345 billion for Tax Year 2001 (after the collection of late and enforced payments, the net tax gap is estimated at \$290 billion for Tax Year 2001). This paper explains the concept of the tax gap, discusses how it is estimated, and points out some limitations with the estimates.*

INTRODUCTION

The term “tax gap” has been used repeatedly by pundits and policy makers. These individuals exhibit a wide range of understanding of the concepts underpinning the tax gap estimates. Our aim with this article is to provide public finance professionals with a basic understanding of the various aspects of the tax gap so these professionals can help shape and guide future debate on this subject.²

DEFINING THE TAX GAP

The federal “tax gap” is defined for a given tax year as the difference between the amount of tax owed by taxpayers under the Tax Code and the amount that is actually paid to the federal Treasury on time. This definition seems straightforward, but in many cases there are some difficulties applying it in practice. For instance, consider the case of a complex multinational enterprise, with far-flung operations and complicated financial undertakings. The amount of tax imposed by the Tax Code for this entity may not be viewed as certain by the taxpayer and its advisers and investors. Often, many interpretations of tax law are required to match the entity’s fact patterns, and these interpretations may differ between the taxpayer and the tax administrator. If there is some uncertainty about the true tax liability imposed by the Tax Code, then this uncertainty will flow right through to estimates of the tax gap.

In principle, the tax gap covers all taxpayers and all taxes (individual and corporate income tax, payroll taxes,

**Mark J. Mazur &
Alan H. Plumley**
*Office of Research,
Analysis, and
Statistics, Internal
Revenue Service,
Washington, D.C.
20024*

National Tax Journal
Vol. LX, No. 3
September 2007

¹ This paper is an elaboration of a short piece contained in the February 2007 NTA Network newsletter. The views expressed in this article are those of the authors and should not be construed as the official position of the Internal Revenue Service or the U. S. Department of the Treasury.

² See also U.S. Department of the Treasury (2006).

estate and gift taxes, and excise taxes). In practice, however, the Internal Revenue Service (IRS) does not yet have estimates of all components of the tax gap for a few types of taxes.

The Internal Revenue Service limits its estimates of the tax gap to tax due on legal-sector activity only. That is, taxes that may be due on income from activities such as illegal gambling, prostitution, and drug dealing are not included in the estimates of the tax gap. The primary rationale for this exclusion is somewhat philosophical, in that the overriding government interest in these areas is not in eliminating the tax gap that might be associated with illegal activities, but rather to eliminate the activity itself. A secondary (though more practical) rationale is that it would be extraordinarily difficult to estimate the amount of taxable income from illegal activities, net of deductible expenses associated with earning that income.

The tax gap is not synonymous with the so-called "underground economy," though there is some overlap. The underground economy is made up of activities that are not very visible to tax and other government authorities. Some of these are legal-sector activities and some, illegal-sector. As noted above, illegal-sector income is not included in estimates of the tax gap. While some of the tax gap arises from legal-sector income generated by underground economy participants, some of it arises from noncompliance that is completely unrelated to the underground economy—such as claiming the wrong filing status or overstating exemptions or tax credits. So, while there is substantial overlap between the tax gap and the underground economy, it is best to maintain the distinction between these two concepts.

COMPONENTS OF THE TAX GAP

The Internal Revenue Service classifies every dollar of the tax gap into one of

the following three mutually exclusive categories.

Nonfiling gap—The tax not paid on time associated with returns that are filed after the due date (with extension, if any) or not filed at all. Any withholding (or other timely payments) paid on behalf of the taxpayer is netted out in computing estimates of the nonfiling gap.

Underreporting gap—The additional tax due on timely filed returns arising from the misreporting of tax liability on those returns (compared to the true liability owed under the Tax Code).

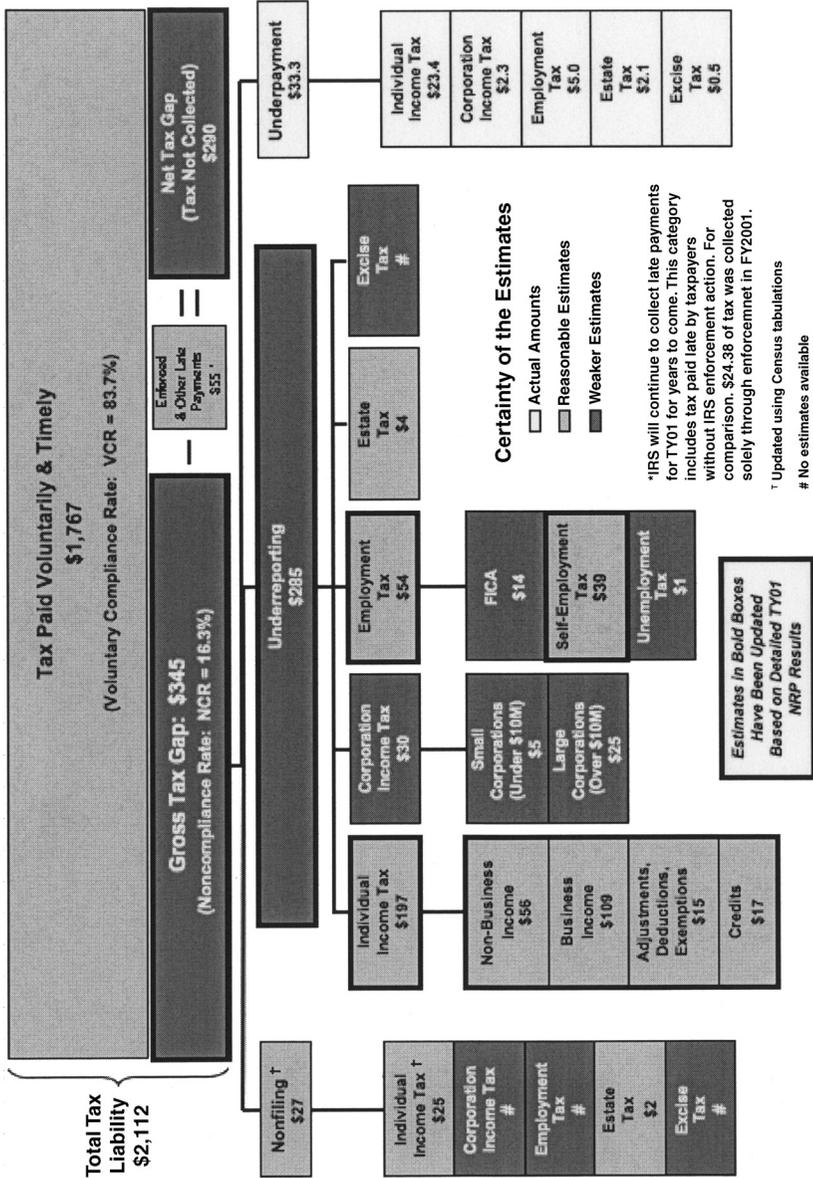
Underpayment gap—The tax that is reported on timely filed returns, but that is not paid on time.

For each of these components, the gross tax gap is defined as underpayments net of overpayments, so taxpayer errors in favor of the federal government are appropriately accounted for in these estimates.

These three components are depicted in the chart displayed below, called the Tax Gap Map. From the Map it is apparent that for Tax Year 2001 (the most recent year for which estimates are available) underreporting accounts for over 80 percent of the \$345 billion gross tax gap, while nonfiling and underpayment each account for less than ten percent of the total gross tax gap.

The Map also indicates that the largest portion of the tax gap is attributable to the individual income tax (which is primarily because it is the federal government's largest source of revenue). The biggest contributor to the individual income tax gap is underreported business income (from sole proprietors or business owners, farmers filing Schedule F, or recipients of partnership, S-Corporation, rental, or royalty income). This is due largely to the

Figure 1. Tax Gap Map for Tax Year 2001 (in \$ Billions)



fact that much of this income is much less “visible” to the IRS through third-party information reporting or withholding than are other types of income.

The Tax Gap Map organizes estimates of the various components of the tax gap (where these estimates exist). It also shows that an estimated \$55 billion of Tax Year 2001 liabilities will eventually be collected through enforcement actions taken by the Internal Revenue Service, or via other late payments, leaving a net tax gap of \$290 billion for Tax Year 2001. Finally, the Tax Gap Map provides an indication of the strength of the estimates. White boxes (no shading) indicate that the figures are generated from internal accounting data and can be considered very solid. This characterization applies only to the underpayment data, which are derived from Internal Revenue Service accounting data. Light shading indicates that the estimates were generated from relatively recent studies and, thus, can be considered quite reasonable. This covers the individual income tax and self-employment tax estimates, which were generated as a result of a Tax Year 2001 study of the reporting compliance of individual taxpayers. The darkest shading indicates that the estimates are based on older data, with results projected to Tax Year 2001 assuming constant compliance behavior. These estimates are generally extrapolated from studies of taxpayer compliance covering periods prior to 1988 and are clearly weaker than the others.

ESTIMATING THE TAX GAP

The elements of the Tax Gap are estimated separately, with the various components combined on the Tax Gap Map. As noted above, the underpayment gap is computed directly from IRS accounting data. IRS knows both how much tax is reported on timely filed returns and

how much of that is paid on time. The underpayment gap is merely the difference between the two. Since these figures are derived from internal accounting data, these calculations can, in principle, be done on a fairly timely basis.

Nonfiling gap estimates have been derived for the individual income tax and the estate tax. In both cases, the estimates result from comparing estimates of overall tax liability among the population required to file a return developed from a non-tax data set (such as the Current Population Survey) with corresponding figures derived from tax data pertaining to timely filers. The estimate of overall tax liability associated with late filers and nonfilers is then reduced by the amount of tax withheld or otherwise paid on time by nonfilers and late filers to get an estimate of the nonfiling gap. Since these estimates require comparisons of multiple data sets, generally there is a noticeable time lag between when returns are (or should be) filed and when these estimates are generated.

Underreporting gap estimates typically are derived either from thorough analyses (including audits – sometimes called “research audits”) of randomly selected returns, or from the regular operational audits that the IRS conducts. For Tax Year 2001, the IRS undertook a reporting compliance study of 46,000 randomly selected individual income tax returns under the new National Research Program (NRP).³ Examinations started in October 2002 and continued through September 2004. Comprehensive tax gap estimates were produced from these data in February 2006, illustrating the substantial time required to develop these sorts of estimates.

One of the reasons for the lag in analysis is that examinations of complex returns take a significant amount of time to bring to completion. Statistically valid estimates for the population require that high qual-

³ See Brown and Mazur (2003).

ity data be collected for all observations in the sample, and there is a necessary time delay in this process. A second reason for the lag is the time needed to review and clean the data and address any anomalies that may exist in the data.⁴ A third reason for the lag is that comprehensive tax gap estimates require an adjustment to take into account the extent to which auditors do not detect all underreported income. In prior efforts to estimate the individual income tax gap, the Internal Revenue Service used results from a study of 1976 data that compared income reported on third-party information documents with the unreported income detected by examiners who did not make use of those documents to estimate how much unreported income could not be detected by unaided examiners. These “multipliers” were used to adjust the results from the examinations of randomly selected returns to account for undetected underreported income.⁵ This approach suffered from age and the belief that taxpayer behavior may have changed since the advent of widespread third-party income reporting.

In response, the Internal Revenue Service applied a new statistical technique for the Tax Year 2001 individual income tax underreporting gap estimates for estimating the amount of income that auditors did not detect. Called “Detection-Controlled Estimation,” the technique controls for which returns were audited by the same agent.⁶ This technique makes use of the fact that individual auditors have strengths and weaknesses in their detection capabilities, which are reflected in somewhat consistent results for a given auditor, but in significant differences across auditors. The estimating procedure essentially constructs a “composite best examiner” and implicitly applies

this income detection ability to all of the face-to-face audits in the sample. Due to data limitations, this adjustment is applied to four discrete categories: high- and low-income taxpayers and high- and low-visibility income (where “visibility” refers to the degree of third-party information reporting). For each category, a multiplier was constructed and applied to the detected but unreported income to generate an estimate of the total amount that should have been reported—covering both the portion detected and the portion undetected by the examiner. The IRS also supplemented the data from the analyses of randomly selected returns with an independent estimate of unreported tip income.

The Self-Employment Tax underreporting gap estimate is also derived from the NRP study of reporting compliance for individual taxpayers. Most of this gap arises because individuals who misreport self-employment income for income tax purposes also tend to misreport that income for purposes of computing their self-employment tax liability.

The corporation income tax underreporting gap estimates for Tax Year 2001 were based on estimates published for tax years in the 1980s. Those earlier estimates were projected to 2001 based on the assumption that reporting compliance rates were constant for each major component during the intervening years. That is, the estimate grew in proportion to the growth in actual tax receipts. The estimate for corporations under \$10 million in assets was based on representative audits under the Taxpayer Compliance Measurement Program (TCMP). The estimates for mid-size corporations (those with assets between \$10 and \$100 million) were based on yield curves derived from regular

⁴ For example, if a taxpayer in the sample filed a return as single and should have filed a joint return with a spouse, the data need to be adjusted to include the information from the other spouse’s return and the weight placed on this observation generally will need adjustment.

⁵ See U.S. Department of the Treasury (1996).

⁶ See Feinstein (1990) and Feinstein (1991).

operational audits that expressed average audit yield as a function of audit coverage rate; these curves were projected to 100 percent coverage to estimate the total gap. The estimates for the largest corporations (over \$100 million in assets) were based on actual operational audit results with no projection to the rest of the population; this was because back in the 1980s, the coverage rate among these large corporations approached 100 percent, meaning no projection was necessary. The estimates for the small corporations accounted for what was not detected by the auditors following a methodology similar to that used for individual income tax (the "multiplier" methodology described above). The estimates for mid-size and large corporations, however, made no such adjustment. For these taxpayers, the estimates were based on the amount of additional tax recommended by the operational audits, rather than on the amount that was ultimately assessed (or somewhere in between). The underreporting gap (as explained earlier) is based on a comparison of the liability reported to the true tax liability, and neither the amount recommended in operational audits nor the amount ultimately assessed perfectly represents true tax liability—particularly since it is likely that there is much non-compliance that is not detected at all by the auditors. On balance, the IRS believed that the undetected noncompliance missing from the recommended amounts was at least as large as the extent to which the recommended amounts overstate true liability on the issues that were detected, so using the recommended amounts would be somewhat conservative. This estimate, therefore, represented an imperfect approximation of the reporting tax gap for these taxpayers.

Perhaps the most heroic assumption in this area, though, is that reporting compliance behavior has remained roughly constant for two decades. With substantial changes in financial planning, the charac-

terization of corporate tax departments as "profit centers" instead of "compliance centers," and the perceived increased use of aggressive (and sometimes abusive) tax planning strategies, it seems very unlikely that reporting compliance behavior has not deteriorated over at least part of that time period. Given all these concerns, the corporate income tax gap estimates should be considered among the weakest in the entire Tax Gap Map.

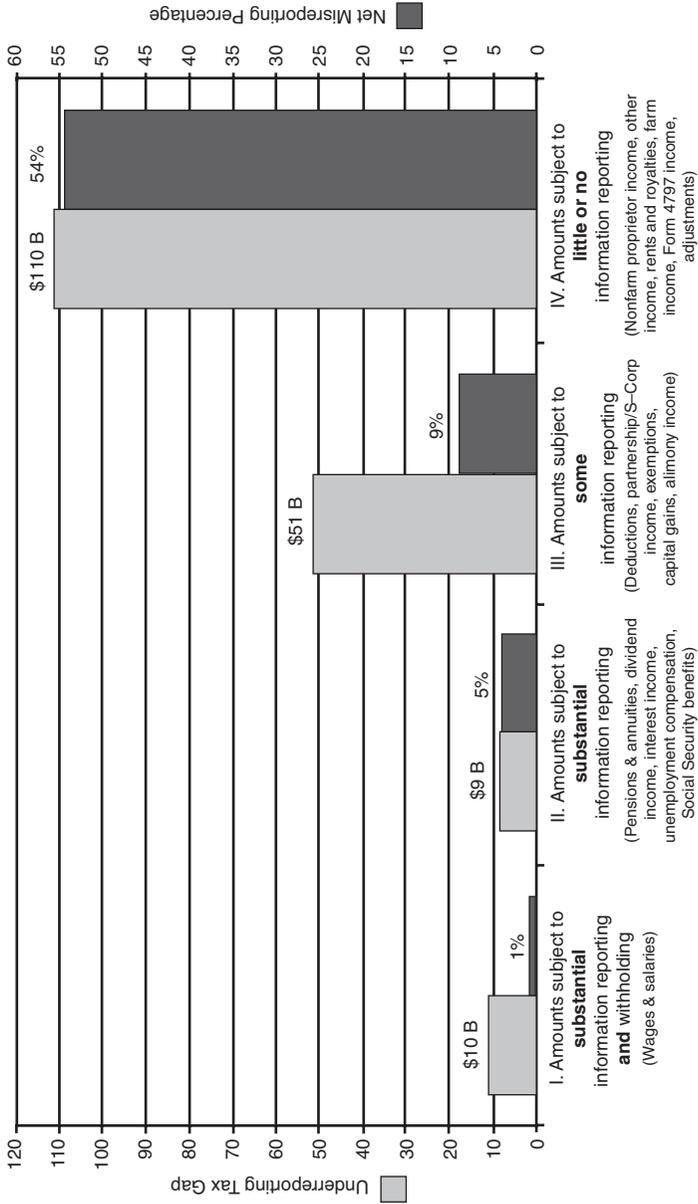
The Internal Revenue Service is currently undertaking a reporting compliance study of S-Corporations (the first study of these entities in almost 20 years). This study encompasses about 5,000 returns filed in Tax Years 2003 and 2004. Since the income and expense items for S-Corporations flow through to the owners, this study will help refine the underreporting tax gap estimates for the individual income tax. Results should be available sometime in 2008.

The payroll tax gap estimates are derived from studies undertaken in the 1980s, with the results projected forward assuming constant reporting compliance behavior. These estimates have many of the same concerns described for the corporate income tax, though there is somewhat less scope for aggressive tax avoidance (or evasion) in the payroll tax area. So while the estimates are derived from dated data, they likely are somewhat more reliable than those for the corporate income tax.

THE ROLE OF INCOME VISIBILITY

The consistent finding of studies conducted by the Internal Revenue Service is that income subject to third-party reporting and withholding (such as wages and salary payments) is reported with about 99 percent compliance. Income with substantial third-party reporting but no withholding (such as interest and dividend payments) is reported with about 95 percent compliance. Income with some, but not complete, third-party reporting

Figure 2. Relationship between Income Visibility and Reporting Compliance Level



Based on updated estimates derived from the TY01 National Research Program study of individual income tax reporting compliance.

(such as capital gains and alimony) is reported with about 90 percent compliance. And income with little or no information reporting (such as sole proprietor income) is reported at about a 50 percent compliance rate. (See the accompanying "visibility chart.")

FUTURE PLANS

Estimates of the tax gap will always lag estimates of tax receipts and other important items. In large part, this is because noncompliance can be assessed only some time after the tax period has ended. However, a robust research agenda for the Internal Revenue Service can help provide improved estimates and also lead to more regular updates of the tax gap. The Internal Revenue Service recently announced a new set of individual income tax reporting compliance studies, commencing with Tax Year 2006 returns. These studies will occur annually and will involve relatively small samples (for example, approximately 13,000 Tax Year 2006 returns will be included). Estimates of the amount of individual income tax underreporting (and updated estimates of the tax gap) will require knitting together the results for several years of observations.

The President's FY 2008 Budget Request includes a significant amount of support for reporting compliance studies, in part to help improve the reliability of the resulting tax gap estimates. Obtaining better data can lead to improved decision-making within the Internal Revenue

Service and by oversight bodies. Since the IRS Oversight Board has established a voluntary compliance goal of 86 percent for 2009, there will be ongoing demand for these estimates. This increased attention will also provide a spur for developing improved estimation techniques. This, in turn, makes us cautiously optimistic that the quality of tax gap estimates will increase over the coming years.

REFERENCES

- Brown, Robert E., and Mark J. Mazur.
 "The National Research Program: Measuring Taxpayer Compliance Comprehensively." *University of Kansas Law Review* 51 No. 5 (December, 2003): 1255-71.
- Feinstein, Jonathan S.
 "Detection Controlled Estimation." *Journal of Law and Economics* 33 No. 1 (April, 1990): 233-76.
- Feinstein, Jonathan S.
 "An Econometric Analysis of Income Tax Evasion and Its Detection." *Rand Journal of Economics* 22 No. 1 (Spring, 1991): 14-35.
- U.S. Department of the Treasury. Internal Revenue Service.
 "Federal Tax Compliance Research: Individual Income Tax Gap Estimates for 1985, 1988, and 1992." Publication 1415 (rev. 4-96). Washington, D.C., 1996.
- U. S. Department of the Treasury. Internal Revenue Service.
 "Tax Gap Figures." Press release. Washington, D.C., February 14, 2006.