Tulane University

Estimating the "Tax Gap" at the State Level:
The Case of Georgia's Personal Income Tax

James Alm   Kyle Borders
Department of Economics
Tulane University
jalm@tulane.edu

Abstract

Most studies of the so-called "tax gap" (or the amount of taxes that should be collected but are not) focus on national taxes. This study provides several estimates of the "tax gap" for the State of Georgias personal income tax. The methods use different estimation strategies for each of the three main components (underreporting of income, underpayment of tax liability, and nonfiling of a tax return), and then sum these separate estimates of the tax gap components to yield a range of estimates of the total tax gap in Georgia. The estimated range of the personal income tax gap is \$1.4 billion to \$2.9 billion, for a voluntary compliance rate that ranges from 89.8 percent to 80.8 percent. This study also provides some rough but suggestive estimates of the distributional effects of noncompliance, which indicate that noncompliance as a proportion of income may well be higher in lower income classes.

ESTIMATING THE "TAX GAP" AT THE STATE LEVEL: THE CASE OF GEORGIA'S PERSONAL INCOME TAX

James Alm and Kyle Borders*

Most studies of the so-called "tax gap" (or the amount of taxes that should be collected but are not) focus on national taxes. This study provides several estimates of the "tax gap" for the State of Georgia's personal income tax. The methods use different estimation strategies for each of the three main components (underreporting of income, underpayment of tax liability, and nonfiling of a tax return), and then sum these separate estimates of the tax gap components to yield a range of estimates of the total tax gap in Georgia. The estimated range of the personal income tax gap is $1.4 billion to $2.9 billion, for a voluntary compliance rate that ranges from 89.8 percent to 80.8 percent. This study also provides some rough but suggestive estimates of the distributional effects of noncompliance, which indicate that noncompliance as a proportion of income may well be higher in lower income classes.

INTRODUCTION

The "tax gap" is defined as the difference between tax revenues actually collected and the amount that should be collected if taxpayers fully complied with the tax laws. Studies to estimate the tax gap have grown in number within the past decade. However, these studies typically focus on estimates of <u>national</u> tax gaps, such as the federal government individual and corporate income tax, and are not often conducted for specific taxes at the <u>state</u> level. In particular, there are no prior estimates of the tax gap for the personal income tax (PIT) in the State of Georgia, so that there are no estimates of the uncollected tax revenues of the Georgia PIT. In this paper, we report on the results from several methods to estimate the total personal income tax gap in Georgia. Our tax gap estimates are in the range of $1.4 billion to $2.9 billion, which imply an estimated "voluntary compliance rate" (or the amount of taxes paid voluntarily as a percentage of legally due taxes) between 89.8 percent and 80.8 percent. We are also able to provide some rough but suggestive estimates of the distribution of the tax gap across different income levels, which indicate that noncompliance as a proportion of income may well be higher in lower income classes.

It should be recognized at the outset that estimating the tax gap faces a fundamental difficulty: measuring the tax gap requires measuring tax evasion, and there is no reliable information on the extent of tax evasion. After all, tax evasion is illegal, and individuals have strong incentives to conceal their cheating, given financial and other penalties that are imposed on individuals who are found cheating on their taxes. There have been many approaches to the measurement of evasion, but, as discussed in detail by Alm (2012), all of them are subject to various and serious criticisms. Any resulting estimates of the tax gap from any of these approaches – including our own estimates here – are necessarily subject to much imprecision.

2

Even so, researchers have been increasingly creative in their approaches to measurement of evasion, and the methods that we use here are reflective of these modern approaches.

The next section describes the tax gap and its main components, and also presents a discussion of previous tax gap studies of note. The following section gives a brief overview of the Georgia personal income tax. The datasets and the methods used in estimation of various components of the tax gap are then discussed, followed by the PIT tax gap estimates. We conclude in the final section.

DEFINING AND MEASURING THE "TAX GAP"

Defining the "Tax Gap"

The "tax gap" is typically defined as the difference between tax revenues actually collected in any given year and the amount that should be collected if taxpayers fully complied with the tax laws (Brown and Mazur, 2003; Mazur and Plumley, 2007; Toder, 2007). Researchers and policymakers often focus on two different measurements of the tax gap. The first measurement is the "gross tax gap", or the difference between tax liability paid and the true tax liability. The "net tax gap" is the gross tax gap less payments of the year's tax liability that are collected either via voluntary late payments or via a tax agency's enforcement efforts. Policymakers are often more concerned with the net tax gap than with the gross tax cap because the net tax gap is considered a better indication of the effectiveness of tax compliance efforts. There are also several separate components of the tax gap. These components include the "nonfiling gap" (or taxes not paid by individuals who do not file a return at all or who file after the due date), the "underreporting gap" (or taxes not paid by individuals who file a return but who misreport their true tax liability), and the "underpayment gap" (or taxes reported on filed tax

3

returns that are not paid by individuals in a timely way). The largest of these three components is typically the underreporting gap. These components can also be broken down by type of tax (e.g., individual income tax, corporation income tax, employment tax, estate tax, excise taxes), of which the individual income tax often accounts for the largest amount of the tax gap, at least at the federal level.

Again, it should be emphasized that any tax gap estimates are subject to much uncertainty, given the inherent difficulties in measuring the various tax gap components (Brown and Mazur, 2003; McManus and Warren; 2006; Toder, 2007; Mazur and Plumley, 2007). Even so, there is little question that tax gap estimates can in principle be quite useful to a tax agency in identifying ways in which its limited resources can be more efficiently targeted in its enforcement efforts.

Measuring the Tax Gap: Previous Tax Gap Studies

The existing literature for the tax gap estimates of underreporting consists of two main approaches, whose difference lies in the type and availability of the data. The more frequently used procedure relies upon <u>audit-based methods</u> by employing thorough line-by-line audits of a sample of individual tax returns to determine the underreporting of specific types of income from the selected sample; these results are then extrapolated to the entire taxable population in computing a measure of the total underreporting tax gap. The second method also uses a subset of tax returns but is not based upon tax audits. Instead, it estimates true income through more <u>indirect methods</u> that are by their nature somewhat less precise than audit-based measures. There is also a smaller literature that attempts to estimate the extent of nonfiling and the amount of underpayment. The following review details specific papers, studies, and government publications for these methodologies.

Underreporting: Audit-based Methods. The audit-based approaches have been popular with the IRS and with the few states that have attempted to estimate a state tax gap. The most prominent and widely cited tax gap estimate has been completed by the IRS originally through (in part) its Taxpayer Compliance Measurement Program (TCMP) and more recently through its National Research Program (NRP).

This methodology consists of a detailed line-by-line audit of a (stratified random) sample of individual tax returns. These audits yield IRS estimates of "true" reported items, which when compared to "actual" individual reported items allows the IRS to generate estimates of underreported income and/or underreported taxes.[1] There has also been an ongoing IRS compliance program that attempts to identify and to assess nonfilers (although its use in measuring the size of the nonfiling gap has been somewhat sporadic).

The benefits of this type of study are the precision and the flexibility to make observations beyond just the extent of the tax gap using the characteristics of individuals likely to be underreporters (McManus and Warren, 2006). The costs of this approach clearly relate to the extensive efforts that are required in actually conducting the audits.

The most recent IRS estimates were completed as part of the NRP program. For tax year 2006, the IRS (2012) estimated the federal gross tax gap to be $450 billion, for a voluntary compliance rate of 83.1 percent of the total true tax liability; the net tax gap was estimated to equal $385 billion; and the tax gap in 2006 associated solely with the individual income tax was $296 billion. In comparison, previous IRS (2006) estimates for tax year 2001 indicated a gross tax gap of $345 billion, a voluntary compliance rate of 83.7 percent, and a tax gap from the individual income tax gap of $245 billion. Within the 2006 gross tax gap estimate of $450

---

[1] Note that there are no longer line-by-line audits in the NRP.

billion, the IRS attributed $28 billion, $376 billion, and $46 billion to the nonfiling, underreporting, and underpayment gaps, respectively; for the individual income tax along, these respective tax gaps in 2006 were $25 billion, $235 billion, and $36 billion.

In addition to estimating the broad components of the tax gap, the IRS (2006, 2012) has also estimated the underreporting of income by different income sources. Through the audits, the IRS has established the "Net Misreporting Percentage" (NMP) for different sources of income, which measure the unreported income as a fraction of the estimated "true" income. As indicated in Table 1, the IRS estimates that the largest portion of the underreporting portion of the tax gap can be attributed to non-farm business income, partnerships and trusts, and other unearned income. Updated IRS (2012) estimates for the 2006 tax gap show a largely similar pattern, although the updated estimates report the NMP for broader categories.[2]

Underreporting: Indirect Methods. Within the past decade, several state revenue departments have attempted to quantify the tax gap for their states (McManus and Warren, 2006). In total, we have identified six states with published tax gap estimates: California, Idaho, Minnesota, Montana, New York, and Oregon. These studies have used very different methodologies, so their results are not strictly comparable, even when they rely largely upon imputations from federal estimates of compliance rates.

---

[2] The IRS (2012) reports the NMP only for broader income categories. These NMP estimates are:

Type of Income	NMP (%)	Percentage of Tax Gap (%)
Subject to <u>substantial</u> information reporting <u>and</u> withholding (wages and salaries)	1	5.3
Subject to <u>substantial</u> information reporting (pensions and annuities, unemployment compensation, dividends, interest, Social Security benefits)	8	5.8
Subject to <u>some</u> information reporting (deductions, exemptions, partnerships and S corporation income, capital gains, alimony income)	11	30.9
Subject to <u>little or no</u> information reporting (non-farm proprietor income, other income, rents and royalties, farm income, Form 4797 income, adjustments)	56	58.0

For example, New York (2005) and Minnesota (2004) employed a Census-based aggregate approach in which the state department of revenue estimated the income that should be reported according to Census data, and compared this estimate with the income that was actually reported on tax returns. The states then disaggregated the total tax gap into underreporting and nonfiling gaps by using the previously mentioned IRS (2006) tax gap component estimates. Other approaches include the less complex estimation strategies employed by California, Idaho, Montana, and Oregon, which simply apply the IRS-estimated misreporting percentages to calculate the state tax gap, given the distribution of income across different income sources in a particular state; for example, see California (2005). State estimated tax gaps (in millions of dollars) are listed in Table 2, where the estimated voluntary compliance rate again measures actual tax collections as a percentage of legally due taxes. Again, given state different methodologies, these estimates are not truly comparable.

While federal and state agencies have typically used auditing results and Census data to generate tax gap estimates, individual researchers have used more indirect approaches to estimate the tax gap. The advantage of these estimations is their ability to quantify the tax gap using less detailed data. However, these indirect approaches are also generally less reliable estimates for all components of the tax gap. Especially novel illustrations of these indirect approaches are by Pissarides and Weber (1989) and Feldman and Slemrod (2007), who use consumption-based or tax deduction-based measures as indirect indicators of tax evasion, on the assumption that a taxpayer's choices of consumption or of deductions are based on their "true" amount of income, not on their reported amounts of income.

For example, Feldman and Slemrod (2007) assume that an individual's reported wages approximate "true" wages and also that the ratio of one's charitable contributions to "true"

income does not vary by income source (e.g., income as reported Schedules C, D, E, and F).[3] If these assumptions hold, then the ratio of reported charitable contributions to "true" income from the different income sources should equal the ratio of charitable contributions to wages.[4] They use these assumptions to impute "true" income from these Schedules and then compare it to reported income on the Schedules. Their estimation results (Table 3) imply that the compliance rate for Schedule C income is 65.0 percent, which implies that reported Schedule C income must be scaled up by a factor of 1.54 (or 1/0.650) to estimate true Schedule C income. We apply here a variant of the Feldman and Slemrod (2007) method, as discussed later.[5]

Nonfiling: Audit-based Methods. Researchers have also focused on measuring the number of nonfilers. Erard and Ho (2001) use a special sample of filers and nonfilers whom the IRS was able to identify (or "locate"), from the 1988 TCMP, in order to estimate the characteristics of nonfilers versus filers in the federal individual income tax. Using the mean

---

[3] Note that Schedule C is for Non-farm Sole Proprietor Income/Loss, Schedule D is for Capital Gains Income/Loss, Schedule E is for Farm Rent, Rent, Royalties, and Estate Incomes/Loss, and Schedule F is for Farm Income/Loss.

[4] More precisely, Feldman and Slemrod (2007) use tax return data from the 1999 IRS Income Tax Microfile Data, and from a smaller IRS-University of Michigan panel data set for the years 1979 to 1985, to estimate charitable contributions as a function of reported income, the amount of income from different sources, the marginal price of charitable contributions, demographic variables, and binary independent variables equal to 1 if the taxpayer reports income from a source other than wage or salaries. The underlying premise behind their specifications is that, assuming no correlation between source of income and charitable contribution, any estimated difference in the relationship between charitable contributions and income earned from different sources can be attributed to underreporting of income.

[5] There is also a substantial literature that uses indirect methods to estimate various measures of compliance with other taxes, notably the value-added tax (VAT). A commonly used measure of VAT compliance compares actual VAT receipts with an estimate of hypothetical VAT revenues that could be generated with full VAT compliance, the latter estimate generally derived from household expenditure surveys. There are also other indirect measures such as "VAT collection efficiency" (defined as the ratio of VAT revenue to aggregate consumption, divided by the standard VAT tax rate) and "VAT productivity" (defined as the ratio of VAT revenue to the product of the standard VAT tax rate and GDP); both of these measures are straightforward to calculate, but both also combine revenue aspects with compliance aspects. For examples of these approaches, see Her Majesty's Customs and Excise (HMCE, 2002) for the United Kingdom, Gebauer and Parsche (2003) for estimates for several EU countries, and Reckon LLP (2009) also for estimates for EU countries, as well as many IMF studies. See Keen and Smith (2006) for a critical survey of these many indirect methods. Attempts to measure compliance with other indirect taxes (e.g., U.S. retail sales tax) are limited. Fox and Murray (2004) survey many of these studies. For a specific example, see Alm and Melnik (2010) for estimates of seller compliance with state sales taxes.

values of the characteristics of the subsample of "located" nonfilers, they compute the nonfiling

tax gap to be approximately $11billion and the number of nonfilers to be around 7.9 million.

These estimates suggest that, although nonfiling is not an especially widespread problem in the

federal income tax (only 7 percent of tax noncompliers are nonfilers), those who do not file

account for about 13 percent of the total federal income tax gap; that is, while underreporting of

income is the more widespread problem, nonfiling is a sizable portion of the tax gap.

Note that prior to 2006, the IRS based its own estimates of the nonfiling gap on the

"Census Exact Match Study", which essentially compared individuals in Census data with IRS

filing records, with those individuals in Census data who could not be "matched" with IRS

records classified as nonfilers (Childers and Hogan, 1984). The IRS now relies on IRS

administrative data that uses third-party information returns to identify income unreported on the

late returns as well as income that should have been reported on unfiled (but required) returns. A

tax calculator is then used to determine the tax gap associated with those two groups of people.

The late filers and the no-return people ended up contributing about equally to the estimated total

nonfiling gap in 2006, or $25 billion of the total tax gap from the individual income tax of $296

billion.

Underpayment: Audit-based Methods. The audit-based approaches have also been used

by the IRS to estimate the underpayment tax gap. Again, these estimates have been largely based

upon audits conducted by the IRS through the TCMP and the NRP.

THE GEORGIA PERSONAL INCOME TAX

The Georgia Personal Income Tax (PIT) is an individual income tax that uses federal

Adjusted Gross Income (AGI) as its starting tax base. Then, through a series of scheduled

adjustments, itemized or standard deductions, and personal exemptions, Georgia Taxable Income

(GTI) is derived.[6] A graduated tax rate schedule is applied to GTI, with the minimum tax rate

being 1 percent and the maximum tax rate being 6 percent. Given the relatively narrow income

ranges, the tax is effectively a 6 percent flat rate tax, and nearly two-thirds of the taxable

population reports income over the taxable income threshold at which income is taxed at the 6

percent tax rate.[7] The State of Georgia is heavily reliant on the personal income tax, and the PIT

has now surpassed sales taxes in terms of its importance in total tax revenues.

DATA AND METHODS

Data

Our tax gap estimates for the Georgia personal income tax use two major datasets: the

2006 Georgia personal income tax returns dataset from the Fiscal Research Center at the Andrew

Young School of Policy Studies at Georgia State University, and the IRS 2006 Income Tax

Master File (ITMF) tax return information. These data sets (and their respective years) were

---

[6] There are also several tax credits, such as a low income credit, a credit against income taxes paid by a Georgia resident to another state, a credit for the purchase of a new home with handicapped accessibility features (or for retrofitting expenses), and a credit for the purchase of a low emission motor vehicle in some designated high emission geographic areas.

[7] The tax rate schedules in 2006 were:

	GTI Income Range by Filing Status ($)		
Rate (%)	**Singles**	**Married Filing Separate**	**Married Filing Joint and Head of Household**
1	<750	<500	<1000
2	750-2250	500-1500	1000-3000
3	2250-3750	1500-2500	3000-5000
4	3750-5250	2500-3500	5000-7000
5	5250-7000	3500-5000	7000-10,000
6	>7000	>5000	>10,000

Personal exemptions in 2006 were $5400 for married couples filing a joint return and $2700 for other filers, and the standard deduction in 2006 was $2300 for single filers and heads of households, $1500 for married taxpayers filing separate returns, and $3000 for married taxpayers filing joint returns. The dependent exemption in 2006 was $2700. The rate schedules remain unchanged in 2014, but there have been some small increases in personal exemptions, the dependent exemption, and standard deductions since 2006.

chosen for comparability. The "Georgia PIT dataset" has 3,817,254 tax returns, which include a unique taxpayer identifier, total reported income, total deductions (both standard and itemized), and total credits. The "ITMF dataset" is a sample of 145,858 federal tax returns that have been weighted to be representative of the entire population of individuals who file tax returns. Of some importance, the ITMF dataset has a state identifier that we use to obtain tax returns for the State of Georgia. This smaller Georgia subset of the ITMF dataset contains 1,956 observations for Georgia, including 891 that use itemized rather than standard deductions. We use 2006 returns for Georgia (both from the Georgia PIT dataset and from the ITMF dataset) in order to apply IRS 2006 tax gap estimates that are based on 2006 information. Our procedures can in principle be applied to other years.

Two issues emerge with these datasets. First, regarding the Georgia PIT dataset, reported income is not broken down by different income sources. To deal with this first issue, we use the subset of Georgia tax returns from the ITMF dataset to create proportions of total income reported from each income source and type, and we then apply these percentages to the Georgia PIT dataset.

More precisely, we follow a two-stage process to determine the proportion of total income attributed to each source of income for which the IRS estimated misreporting percentages. Using the ITMF 2006 Georgia returns and a high income random sample (see below), we calculate for each taxpayer the following categories of (taxable) income: reported wage, interest and dividends, pensions and IRA accounts, unemployment income, S-corporation and partnership income, capital gains income, alimony income, business income, farm income, other gains (from the sale of business property), rents and royalties, other income, and social security income. We then combine these separate incomes to get aggregate income amounts for

11

the sample. Each of these aggregate incomes is divided by the sum of all of the income variables plus scheduled adjustments to yield the percentage of total income attributed to each income source. Table 4 includes the amounts and the estimated percentages of total income for each source of income.

A second data issue arises because the ITMF dataset is top-coded, meaning that the state identifier is removed from individuals who report total income in excess of $200,000. To get a more representative sample of high-income Georgia taxpayers, we use the Georgia PIT dataset to calculate the percentage of Georgia taxpayers who filed a tax return reporting income in excess of $200,000 as a percentage of total taxpayers and as a percentage of the total taxpayers who filed returns using itemized rather than standard deduction. The computed percentages were 2.76 percent of total taxpayers with income in excess of $200,000 and 5.68 percent of itemizing taxpayers filed income in excess of $200,000. Using these percentages and applying them to the total of 1956 Georgia tax returns in the ITMF dataset and the 891 itemized tax returns in the ITMF dataset yields simple random samples of 54 tax returns and 51 tax returns, respectively.

Tax Gap Estimation Methods

As mentioned earlier, the tax gap consists of three different components, the underreporting gap, the underpayment gap, and the nonfiling gap. Our basic approach is to estimate each of these gaps separately using various methods and then to add the different gap estimates together to obtain a range of possible values for the tax gap. The estimation strategies for each of the components of the tax gap are described below.[8]

---

[8] The State of Georgia, like the United States federal government, allows losses to be carried forward through years, thereby deducting the losses from income in future years. The over-reporting of losses diminishes the tax revenue in future years as taxable income is diminished by a yearly decreasing amount of past losses. We deal with this issue by assuming that all of the over-reported losses are (eventually) realized, and applying a 6 percent flat tax rate to all

It should be noted that we attempted two additional methods to estimate the Georgia PIT tax gap. One method attempted to estimate Georgia-specific compliance rates using the Feldman and Slemrod (2007) approach as applied to our Georgia data. Unfortunately, the resulting estimation results were not useful.[9] We also applied a method based on Census data to estimate the PIT tax gap, with somewhat more success.[10] Even though we have tried these other methods, we focus our discussion here on several different estimation strategies, as discussed in detail next.

---

over-reported losses in the same fashion as we do for all estimates of unreported income. This assumption that all losses are eventually realized is unlikely to be the case. For example, an anonymous referee indicated that roughly two-thirds of the corporate net operating losses generated by California corporations between 1985 and 1999 expired before they could be used to offset gains. Given the small relative magnitude of losses in Georgia data, this assumption does not have much impact on our estimates, and we use this assumption mainly for simplicity.

[9] We used the Georgia-specific returns plus a random sample of high income tax returns of the ITMF dataset to estimate a double-log regression equation of cash charitable contributions of itemized tax filers in the Georgia dataset as a function of the income attributed to different schedules, control variables, and demographic variables. The full equation was:

$$\ln(G+100) = \alpha_0 + \alpha_1 \ln(V + \sum_{ih} k_{ih}R_{ih} + \sum_j b_j S_j) + \alpha_2 \ln(Price) + \alpha_3 PEX + \alpha_4 MAR + u$$

where

G = Cash Charitable Contributions
V = Visible Income (e.g., income that cannot be misreported)
R = Reported income from each source of non-visible income
S = Dummy variable for each income source (Schedule D is the omitted category), equal to 1 if taxpayer files a tax return for that schedule and 0 otherwise.
Price = First-dollar marginal tax rate price of charitable contributions
PEX = Number of non-personal exemptions
MAR = Dummy variable equal to 1 if taxpayer is married and 0 otherwise
u = Error term.

and α_i, $k_{i,h}$, and b_j are coefficients. This equation was estimated using unweighted nonlinear least squares regressions. Unfortunately, given the small size, the estimated standard errors were quite large, so that the estimated slope coefficients on the schedule income variables were not always significant; even when significant, the estimated slope coefficients sometimes implied implausible results.

[10] The census-based approach assumes that all citizens report income correctly to the Census Bureau. Consequently, the difference between taxes calculated using income reported to the Census Bureau and taxes on income voluntarily reported income on state tax returns is the gross tax gap ((including the nonfiling, the underreporting, and the underpayment gaps). Since the Census Bureau top-codes all income above a given level for certain sources of income, this process must also be applied to the PIT sample. Also, types of income are defined differently in the two different samples, and a common definition must be applied in the PIT sample to correspond with the Census data sample. Finally, taxpayers who report income less than the standard deduction were removed from the dataset. Using these procedures, the estimated gross tax gap was $3.360 billion, with the tax gap associated with wages and salaries accounting for most all of the total tax gap.

Underreporting Tax Gap. To measure the tax revenues lost from the underreporting of personal income, we use two different strategies.

The first method uses the IRS-determined net misreporting percentages for different sources of income summarized in Table 1 to estimate the unreported income for each source, to which a flat 6 percent tax rate is then applied to estimate underreported taxes. This approach assumes that the Georgia subset of the ITMF dataset is representative of all Georgia taxpayers and that federal taxpayers and Georgia taxpayers misreport different sources of income equally. We call the resulting numbers "IRS-based Estimates". Note that this method applies federal compliance rates to Georgia, but also allows Georgia-specific factors (e.g., income sources) to affect the resulting dollar magnitudes of the underreporting gap.

The second approach applies the estimated compliance rates for Schedule C, D, E, and F sources of income from Feldman and Slemrod (2007) to the Georgia tax returns. The estimated compliance rates are listed in Table 3. This approach assumes: that the Georgia tax returns in the ITMF dataset used to determine the percentage of income filed under each source of income are representative of the actual Georgia taxpayer population; that federal taxpayers' noncompliance is equivalent to the Georgia taxpayers' noncompliance; and that underreporting of income only occurs on schedule C, D, E, and F incomes. We call these numbers "Deduction-based Estimates". As with the IRS-based Estimates, this method applies nationally estimated compliance rates from Feldman and Slemrod (2007) to Georgia, but also allows Georgia-specific factors (e.g., income sources) to affect the resulting dollar magnitudes of the underreporting gap.

Note that these estimates use individual-level data. As a result, they allow us to make rough if suggestive estimates of the distribution of underreporting. We discuss these distributional effects later.

Underpayment Tax Gap. We make two estimates of the underpayment tax gap, based on IRS (2012) estimates of the underpayment tax gap in the individual income tax for 2006. These estimates indicate that the underpayment tax gap from the individual income tax ($36 billion) is 15.3 percent of the underreporting tax gap ($235 billion). We use both the IRS-based Estimates and the Deduction-based Estimates of the underreporting tax gap, and then estimate the underpayment tax gap according to the 15.3 percent figure. These estimates assume that the underpayment tax gap in Georgia is the same share of the underreporting tax gap as it is for the U.S., but allow the total amount of underreporting in Georgia also to determine the Georgia underpayment tax gap. Note that this approach only allows us to estimate the aggregate amount of the underpayment tax gap, and not the amounts by income class.

Nonfiling Tax Gap. In a similar manner, we make two estimates of the nonfiling gap, based on IRS (2012) estimates of the 2006 nonfiling tax gap. These estimates indicate that the nonfiling gap from the individual income tax ($25 billion) is 10.6 percent of the underreporting tax gap ($235 billion). As with the underpayment tax gap, we estimate the aggregate amount of the Georgia nonfiling tax gap by applying the 8.47 percent figure to the IRS-based Estimates and the Deduction-based Estimates of the underreporting tax gap.

GEORGIA TAX GAP ESTIMATES

Underreporting Tax Gap Estimates

IRS-based Estimates. Applying the IRS misreporting percentages to Georgia income sources to estimate the underreporting in the state of Georgia yields an estimate of $1.381 billion dollars in lost tax revenue. These estimates are presented in Table 5. This method also allows for comparison of the tax gap liability across different income sources, as we discuss later. Non-farm

15

Business Income is the largest component of the underreporting tax gap, given the size of Non-farm Business Income together with its high rate of net misreporting.

Deduction-based Estimates. The second method for estimating the underreporting tax gap is based on the Feldman and Slemrod (2007) estimated compliance rates. Similar to the IRS-based Estimates, the ITMF Georgia returns are used to calculate the percentage of total income attributable to Schedule C, Schedule D, Schedule E, and Schedule F. However, unlike the IRS-based Estimates, positive and negative incomes are summed separately, and used to calculate tax gap estimates separately because Feldman and Slemrod (2007) are able to estimate different compliance rates for both positive income and negative income, rates that are often substantially different from one another. For this reason, we calculate underreporting tax gap estimates for these different forms of income.[11] Note that we do not calculate any underreported income for positive Schedule D income and for negative Schedule C income because Feldman and Slemrod (2007) computed that there is no noncompliance for these income sources. As with the IRS-based Estimates, a flat tax rate of 6 percent is applied to the reported income and the true income to obtain the reported tax liability and the true tax liability. The difference between the true tax liability and the reported tax liability for each source of income, summed over all income sources, yields the underreporting tax gap.

Table 6 summarizes these estimates. The underreporting gap from this approach is $2.907 billion. Note that there is far less distribution of the underreporting of income across different income sources, given the assumption that underreporting of income can only occur in Schedules

---

[11] For example, our estimate of the true tax liability for negative Schedule D income is calculated by multiplying the amount of negative Schedule D income reported on the ITMF Georgia returns (or -$2,434.6 billion) by the multiplier (-4.058) and then by the 6 percent individual income tax rate. This calculation gives $592.8 million as the negative Schedule D underreporting tax gap. The underreporting tax gaps for negative Schedule E and Schedule F incomes are calculated in a similar way.

C, D, E, and F income sources. Note also that much less of the tax gap is attributable to Schedule C income. Instead, Schedule E income is now a much higher portion of the underreporting gap. Also, a large proportion of the tax gap is attributable to over-reported tax losses.

While these Deduction-based Estimates are of interest, we believe that they are less reliable than the IRS-based Estimates, for several reasons. First, the IRS-based Estimates are derived from actual IRS audits. Second, the estimation results of Feldman and Slemrod (2007) indicate that there is full reporting of positive Schedule D incomes (or capital gains income) and full reporting also of (negative) Schedule C incomes (or Non-farm Business Income), both of which seem unlikely. Regardless, we report both estimates.

Underpayment Tax Gap Estimates

Recall that the estimates of the underpayment tax gap are dependent upon the estimated underreporting tax gaps, so that we scale the IRS-based and Deduction-based Estimates by the share of underpayment in the underreporting tax gap. The resulting estimates of the underpayment tax gaps are shown in Table 7, or $211.3 million and $444.8 million for the IRS-based and the Deduction-based Estimates, respectively.

Nonfiling Tax Gap Estimates

In a similar way, the estimates of the nonfiling tax gap are scaled versions of the estimated underreporting tax gaps. The resulting estimates of the nonfiling tax gaps are shown Table 8, or $146.4 million and $308.2 million for the IRS-based and the Deduction-based Estimates, respectively.

Summary

The above estimation strategies for each of the components of the overall tax gap generate a range of values of the personal income tax gap in the State of Georgia. In the

aggregate, the estimated range of the tax gap is $1.381 billion to 2.907 billion. Table 9 summarizes all of these calculations.

Compared with other state estimates, our estimates of the Georgia PIT tax gap appear to be representative of the results in other states. For example, we estimate that the voluntary compliance rate in Georgia (again defined as actual tax revenues divided by potential tax revenues) lies within a range of 80.8 percent to 89.8 percent. The six states discussed earlier had compliance rates that ranged from 78 percent to 90 percent, with a mean compliance rate of 84.8 percent. Of course, small differences in compliance rates have significant implications for revenues.

Distributional Effects

The distributional effects of taxation are also of interest, even though few state tax gap studies are able to consider such distributional effects. Because our underreporting tax gap estimates are based on individual-level data, we are able to assess the distribution of the underreporting part of tax gap across different income percentiles; we are not able to attribute the underpayment or nonfiling tax gaps to individuals.

It might be expected that the portion of total income attributable to each percentile would be the corresponding portion of the tax gap attributed to that percentile. However, this expectation is based on two assumptions that may not be accurate: that each income percentile contains the same proportional distribution of the different income sources, and that each income percentile complies with existing tax law at the same rate. Although we cannot test the latter assumption, we can estimate the distribution of the underreporting tax gap across income classes by using the Georgia ITMF dataset to measure the amount of each income source by income class and by then applying estimates of misreporting percentages to these incomes.

18

Specifically, we use the Georgia ITMF dataset to measure the amount of income or loss reported for different sources of income at each income percentile, and we then apply the IRS net misreporting percentages to estimate the tax gap at different income percentiles. The Georgia ITMF tax returns are used to determine the amount of income or loss reported for different sources of income at each income percentile. The predicted tax gap attributed to each income percentile is equal to the percentage of total income attributable to the relevant income percentile. We present our results separately for the bottom income percentiles of the taxpayers and for the higher income percentiles.

This distributional analysis yields some surprising results. It is commonly thought that those with the highest incomes are responsible for the most evasion, although evidence to support this claim is limited. In the most complete analysis of the distribution of income tax noncompliance in the U.S., Johns and Slemrod (2010) use the NRP of the IRS to assess these distributional effects for tax year 2001. They find that, when taxpayers are ranked by their "true" income, the ratio of aggregate misreported <u>income</u> to true <u>income</u> typically rises with income; however, the ratio of underreported <u>taxes</u> to true <u>taxes</u> is tends to fall with income, and in fact this ratio is highest for lower income taxpayers.

While we cannot control for a different propensity to evade at different income levels, we can demonstrate the distributional impact on the (restrictive) assumption of similar evading decisions across income deciles. Doing so yields a tax gap that is skewed to the lower income levels. Figures 1 and 2 illustrate this result. Figure 1 is for the lowest percentile of the tax payers; Figure 2 is for the highest percentile of taxpayers. In both figures, the left bar represents the estimated amount of the tax gap attributed to each income percentile using the total income percentages, while the right bar indicates the actual estimate of the tax gap. At all income

19

distributions for the bottom half of taxpayers, the actual tax gap estimate exceeds the predicted tax gap estimate, while the predicted tax gap exceeds the actual tax gap estimate for higher income percentiles.

However, while suggestive, these results should be taken with some caution. The main factor that drives these results is that certain forms of income (especially Non-farm Business Income, Rents and Royalties, and Farm Income) with higher noncompliance rates are concentrated disproportionately in lower income classes; other forms of income (e.g., Capital Gains) with estimated compliance rates that are higher tend to be received more heavily by higher income classes. The resulting patterns necessarily reflect these features of the data and the assumptions about misreporting.

CONCLUSIONS

Our methods for measuring the tax gap used in this report estimate the Georgia personal income tax to be within a range of $1.4 billion to $2.9 billion, corresponding to a compliance rate that ranges between 89.8 percent to 80.8 percent, a result that is comparable to the compliance rates of states that have completed personal income tax gap studies. We also find suggestive if not definitive evidence that the tax gap is distributed more towards lower income percentiles; that is, lower income households tend to be less compliant than higher income households.

It must be emphasized once again that all of these estimates should be used with much caution, given the many assumptions needed to generate them. Indeed, there are several potential issues with the estimates above, including the lack of Georgia-specific compliance rates, the representativenss of the ITMF Georgia returns and the high income sample, and the over-reporting of losses. For example, the attempt at estimating Georgia specific compliance

rates for different sources of income was largely unsuccessful because of potential data and weighting issues. The inability to generate Georgia-specific compliance rates forces us to assume that the compliance rates of national taxpayers and of Georgia taxpayers are the same (even though any resulting estimates of the tax gap still reflect in part Georgia patterns of income). The issue of whether the ITMF Georgia returns and the high income sample is representative of the Georgia populace is another potential concern. While the IRS claims that its dataset is representative of the state population, the IRS does not make clear whether aggregating the data and assuming the distribution of total income across different income sources is accurate. Also, it is likely that the assumption about the over-reporting of losses leads to an overestimate of the lost tax revenues due to over-reporting of tax losses. Finally, these results may reflect the idiosyncracies of the specific year (2006) that we examined. If the proportions of income from the various sources changes considerably over time, then the application of net misreporting percentages to these income sources will necessarily generate different estimates of the tax gap.

On balance, it is not clear how these potential biases may influence the overall tax gap estimate, and whether they tend to over- or underestimate the Georgia tax gap. Regardless, it seems clear that any estimate of the state tax gap may provide some useful guidance to help direct enforcement efforts, but any estimate must also be viewed with some skepticism.[12]

REFERENCES

Alm, James. 2012. "Measuring, Explaining, and Controlling Tax Evasion: Lessons from Theory, Experiments, and Field Studies." *International Tax and Public Finance* 19 (1): 54-77.
Alm, James, Jeremy Clark, and Kara Leibel. 2013. "Socio-economic Diversity, Social Capital,

---

[12] A natural question is what can be done to reduce the tax gap. Some Georgia-specific strategies are discussed in detail by Alm, Sjoquist, and Wallace (2006), and are similar to federal strategies outlined by the U.S. Department of the Treasury (2006).

and Tax Filing Compliance in the United States." New Orleans, LA: Tulane University Working Paper.

Alm, James and Mikhail I. Melnik. 2010. "Do eBay Sellers Comply with State Sales Taxes?" *National Tax Journal* 63 (2): 215-236.

Alm, James, David L. Sjoquist, and Sally Wallace. 2006. "Tax Collectability and Tax Compliance in Georgia," FRC Report No. 133. Atlanta, GA: Fiscal Research Center, Andrew Young School of Policy Studies, Georgia State University.

Brown, Robert E. and Mark J. Mazur. 2003. "IRS's Comprehensive Approach to Compliance Measurement." *National Tax Journal* 56 (3): 689-700.

Childers, Danny R. and Howard R. Hogan. 1984. "The IRS/Census Direct Match Study – Final Report." Washington, D.C.: Bureau of the Census, Statistical Research Division Research Report Series, CENSUS/SRD/RR-84/11.

California, Legislative Analyst's Office. 2005. "California's Tax Gap." Sacramento, CA.

Erard, Brian and Chih-Chin Ho. 2001. "Searching for 'Ghosts': Who Are the Non-Filers and How Much Tax do They Owe?" *Journal of Public Economics* 81 (1): 25-50.

Feldman, Naomi and Joel Slemrod. 2007. "Estimating Tax Compliance with Evidence from Unaudited Tax Returns." *Economic Journal* 117 (518): 327-352.

Fox, William F. and Matthew N. Murray. 2005. "Sales Taxation in a Global Economy." In Taxing the Hard-to-tax: Lessons from Theory and Practice," edited by James Alm, Jorge Martinez-Vazquez, and Sally Wallace, 221-244. Amsterdam, The Netherlands: Elsevier B. V. – North Holland Publishers.

Gebauer, Andrea and Rudiger Parsche. 2003. "Evasion of Value-added Taxes in Europe: IFO Approach to Estimating the Evasion of Value-added Taxes on the Basis of National Accounts Data (NAD)." *CESifo DICE Report* 2: 40-44.

Her Majesty's Customs and Excise (HMCE), UK Government. 2002. "Measuring Indirect Tax Losses." London, UK.

Internal Revenue Service. 2006. *Tax Year 2001 Federal Tax Gap (Extended Version).* Washington, D.C.: IRS Office of Research, Analysis, and Statistics.

Internal Revenue Service. 2010. *Final Report for Planning Theoretical Research*. Washington, D.C.: Department of the Treasury, Office of Research, Analysis, and Statistics.

Internal Revenue Service. 2012. *Federal Tax Compliance Research: Tax Year 2006 Tax Gap Estimation.* Washington, D.C.: IRS Office of Research, Analysis, and Statistics.

Johns, Andrew and Joel Slemrod. 2010. "The Distribution of Income Tax Noncompliance." *National Tax Journal* 63 (3): 397-418.

Keen, Michael and Stephen Smith. 2006. "VAT Fraud and Evasion: What Do We Know and What Can Be Done?" *National Tax Journal* 59 (4): 861-887.

Mazur, Mark J. and Alan H. Plumley. 2007. "Understanding the Tax Gap." *National Tax Journal* 60 (3): 569-576.

McManus, Jacqui and Neil Warren. 2006. "The Case for Measuring the Tax Gap." *eJournal of Tax Research* 4 (1): 61-79.

Minnesota Department of Revenue. 2004. "Individual Income Tax Gap: Tax Year 1999." Minneapolis, MN.

New York Department of Taxation and Finance, Office of Tax Policy Analysis. 2005. "New York State Personal Income Tax Compliance Baseline Study." Albany, NY.

Oregon Department of Revenue. 2009. "Personal Income Tax Compliance Report 2009." Salem,

OR.

Pissarides, Christoper A. and Guglielmo Weber. 1989. "An Expenditure-based Estimate of Britain's Black Economy." *Journal of Public Economics* 39 (1): 17–32.

Reckon LLP. 2009. *Study to Quantify and Analyse the VAT Gap in the EU-25 Member States.* Report for the Director General Taxation and Customs Union, European Commission.

Toder, Eric. 2007. "What is the Tax Gap?" *Tax Notes* 117 (4) (October 22, 2007): 367-378.

U.S. Department of the Treasury. 2006. "A Comprehensive Strategy for Reducing the Tax Gap." Washington, D.C.: Office of Tax Policy.

Wallace, Sally. 2009. "The Buoyancy of Georgia's Personal Income Tax." FRC Report No. 190. Atlanta, GA: Andrew Young School of Policy Studies, Georgia State University.

TABLE 1
IRS Estimates of Net Misreporting Percentages

Source of Income	Net Misreporting Percentage (%)
Wages and Salaries	1.2
Interest and Dividends	3.7
Pensions and IRA Income	4.1
Unemployment Income	11.1
S Corps, Partnerships, and Trusts	17.8
Capital Gains	11.8
Alimony Income	7.2
Non-farm Business Income	57.1
Farm Income	72.0
Other Gains	64.4
Rent and Royalties	51.3
Other Income	63.5

Source: IRS (2006).

TABLE 2
State Estimates of Tax Gaps and Voluntary Compliance Rates

State	Year	Tax Gap ($ millions)	Voluntary Compliance Rate (%)
California	2004	6,500	85.0
Idaho	2005	NA	82.9
Minnesota	1999	604	89.5
Montana	2006	NA	78.0-82.0
New York	2005	2,838	86.1
Oregon	2006	1,247	81.5-88.9

Source: Oregon Department of Revenue (2009). Note that Idaho and Montana do not generate estimates of the tax gap, but only of the voluntary compliance rate, and "NA" (for "Not Available") refers to these states.

TABLE 3
Feldman and Slemrod (2007) Estimates of Compliance Rates by Income Type

Source of Income	Parameter Estimate (Standard Error)	Implied Compliance Rate (%)
Schedule C_{POS}	1.539 (0.074)*	65.0
Schedule D_{POS}	0.895 (0.010)*	100
Schedule E_{POS}	4.544 (0.154)*	22.0
Schedule F_{POS}	3.868 (0.691)*	25.9
Schedule C_{NEG}	-0.279 (0.381)	NA
Schedule D_{NEG}	-4.058 (0.714)*	NA
Schedule E_{NEG}	-3.354 (0.310)*	NA
Schedule F_{NEG}	-3.036 (0.918)*	NA

Source: Feldman and Slemrod (2007). Note that their method does not generate plausible estimates for all categories, and "NA" (for "Not Available") refers to these categories. Note also that "POS" refers to returns with positive amounts of reported income, by type of schedule; "NEG" refers to returns with negative amounts of reported income, by type of schedule. * denotes a significant difference from one at the 5 percent level.

TABLE 4
Amounts and Percentages of Georgia Total Income from Different Sources as Reported on Georgia Tax Returns

Source of Income	Estimated Reported Income ($ thousands)	Percentage of Total Income (%)
Wages and Salaries	173,070,278.5	67.43
Interest and Dividends	11,575,105.0	4.51
Pensions and IRA Income	17,518,780.8	6.83
Unemployment Income	210,853.0	0.08
S Corps, Partnerships, and	9,441,246.1	3.68
Capital Gains	21,230,655.4	8.27
Alimony Income	176,127.2	0.07
Non-farm Business Income	11,154,116.0	4.35
Farm Income	-1,272,780.4	-0.50
Other Gains	2,535,776.5	0.99
Rent and Royalties	6,277,153.7	2.45
Other Income	1,017,891.5	0.40
Social Security Income	3,746,547.1	1.46

Source: Calculations by authors.

TABLE 5
IRS-based Estimate of Underreporting Tax Gap

Source of Income	Reported Tax Liability ($ thousands)	Estimated True Tax Liability ($ thousands)	Underreporting Tax Gap ($ thousands)
Wages and Salaries	10,384,216.7	10,508,827.3	124,610.6
Interest and Dividends	694,506.3	720,203.0	25,696.7
Pensions and IRA Income	1,051,126.8	1,094,223.0	43,096.2
Unemployment Income	12,651.2	14,055.5	1,404.3
S Corps, Partnerships	566,474.8	667,307.3	100,832.5
Capital Gains	1,273,839.3	1,424,152.4	150,313.0
Alimony Income	10,567.6	11,328.5	760.9
Non-farm Business Income	669,247.0	1,051,387.0	382,140.0
Farm Income	-76,366.8	131,350.9	207,717.8
Other Gains	152,146.6	250,129.0	97,982.4
Rents and Royalties	376,629.2	569,840.0	193,210.8
Other Income	61,073.5	99,855.2	38,781.7
Social Security Income	224,792.8	239,134.6	14,341.8
Total	15,400,905.0	16,781,793.7	1,380,888.6

Source: Calculations by authors.

TABLE 6
Deduction-based Estimate of Underreporting Tax Gap

Source of Income	Reported Tax Liability ($ thousands)	Estimated True Tax Liability ($ thousands)	Underreporting Tax Gap ($ thousands)
Schedule C (POS)	698,146.4	1,074,447.3	376,300.9
Schedule D (POS)	1,305,270.7	1,305,270.7	0
Schedule E (POS)	406,637.1	1,847,759.0	1,441.121.9
Schedule F (POS)	19,761.9	76,439.1	56,677.2
Schedule C (NEG)	NA	NA	NA
Schedule D (NEG)	0	592,778.2	592,778.2
Schedule E (NEG)	0	357,199.3	357,199.3
Schedule F (NEG)	0	83,116.9	83,116.9
Total	2,429,816.1	5,337,010.4	2,907,194.3

Source: Calculations by authors. Note that the deduction-based approach does not generate estimates for all categories. "NA" (for "Not Available") refers to these categories. Note also that "POS" refers to returns with positive amounts of reported income, by type of schedule, and "NEG" refers to returns with negative amounts of reported income, by type of schedule.

TABLE 7
Underpayment Tax Gap

IRS-based Estimate	
Underreporting Tax Gap ($ thousands)	1,380,888.6
Underpayment Tax Gap as Percentage of Underreporting Tax Gap (%)	15.3
Underpayment Tax Gap ($ thousands)	211,276.0
Deduction-based Estimate	
Underreporting Tax Gap ($ thousands)	2,907,194.3
Underpayment Tax Gap as Percentage of Underreporting Tax Gap (%)	15.3
Underpayment Tax Gap ($ thousands)	444,800.7

Source: Calculations by authors.

TABLE 8
Nonfiling Tax Gap

| IRS-based Estimate	
Underreporting Tax Gap ($ thousands)	1,380,888.6
Nonfiling Tax Gap as Percentage of Underreporting Tax Gap (%)	10.6
Nonfiling Tax Gap ($ thousands)	146,374.2
Deduction-based Estimate	
Underreporting Tax Gap ($ thousands)	2,907,194.3
Nonfiling Tax Gap as Percentage of Underreporting Tax Gap (%)	10.6
Nonfiling Tax Gap ($ thousands)	308,162.6

Source: Calculations by authors.

TABLE 9
Summary of Tax Gap Estimates

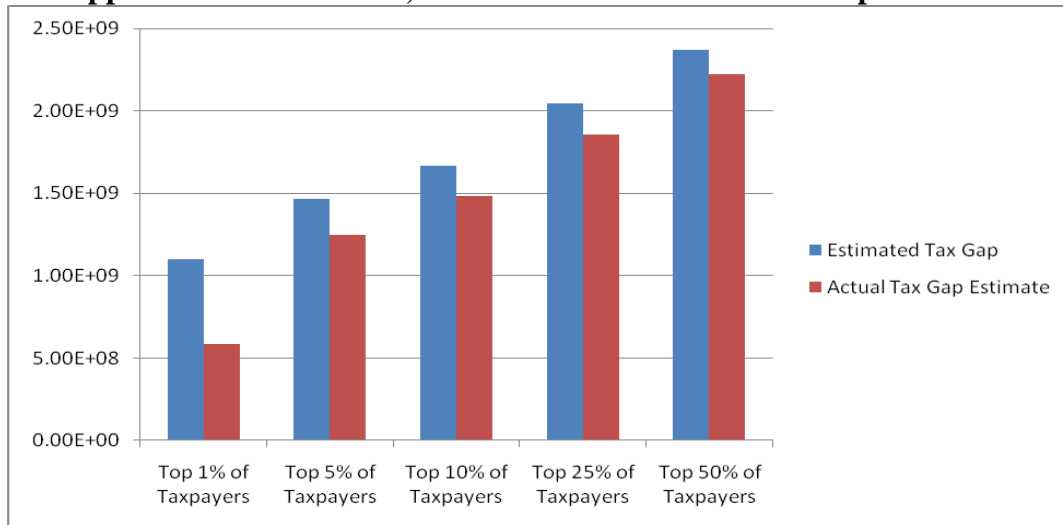| Type of Tax Gap	IRS-based Estimates	Deduction-based Estimates
Underreporting Tax Gap ($ thousands)	1,380,888.6	2,907,194.3
Underpayment Tax Gap ($ thousands)	211,276.0	444,800.7
Nonfiling Tax Gap ($ thousands)	146,374.2	308,162.6
Total: Tax Gap Estimates ($ thousands)	1,738,538.8	3,660,157.6
Voluntary Compliance Rate (%)	89.8	80.8

Source: Calculations by authors.

FIGURE 1
Bottom Income Percentile, Predicted versus Actual Tax Gap Estimates
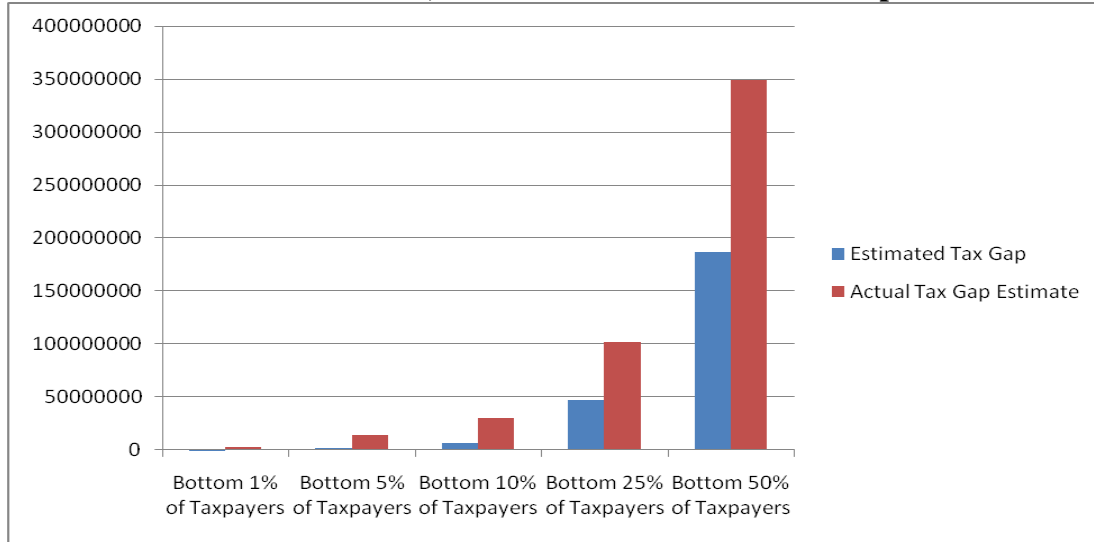
Source: Calculations by authors.

FIGURE 2
Upper Income Percentile, Predicted versus Actual Tax Gap Estimates

Source: Calculations by authors.