Estimating Confidence Intervals in a Tax Microsimulation Model

Robert McClelland[1]*, Surachai Khitatrakun[1], Chenxi Lu[1]

[1]The Tax Policy Center, Washington, DC, USA

Abstract Since the creation of microsimulation models, the need to measure uncertainty has been recognized, but little research has been conducted in spite of the widespread use of these models. In this article, we calculate confidence intervals for a large tax microsimulation model, comparing a normal approximation to a bootstrap estimator. We estimate confidence intervals for five proposed changes to tax law. We explore the relationship between the size of proposals' point-estimated impacts and their confidence intervals by considering high, central, and low scenarios for three proposals. To explore uncertainty around small-magnitude estimates, we consider a proposal designed to keep tax revenue unchanged. Finally, we consider a proposal that affects a small number of taxpayers but has a large heterogeneous effect. We find that, overall, confidence intervals in the model fit tightly around the point estimates, but there are exceptions. We also find that in many cases the normal approximation is close to the bootstrap estimator but may differ for policy changes that affect a small number of taxpayers.

JEL classification: H2, C5

DOI: https://doi.org/10.34196/ijm.00216

1. Introduction

In 1957, Guy Orcutt introduced microsimulation modeling by making a radical proposition: rather than using aggregate data to predict the behavior of populations, researchers could model the decisions made by households and aggregate the results.[1] Data on households would be drawn from surveys to create a sample that would represent the population; the decisions of each household would be separately modeled and the aggregate decisions would predict the behavior of the population. Further, researchers could learn how various assumptions and parameters affect the model's outcomes by recalculating the results under different assumptions and parameters. The benefit of this approach to study government policies was clear: by altering model parameters, the difference in outcomes attributable to different policy choices could be measured. Importantly, uncertainty could be estimated by repeatedly running the model on random subsamples and calculating standard errors and confidence intervals.

Scholars quickly embraced this new approach and by the early 1960s, Joseph Pechman and colleagues at the Brookings Institution had developed a microsimulation model of individual taxation that calculated both total federal revenue and the distribution of taxes across income groups. Since then the microsimulation approach has become the standard method for evaluating the United States individual income tax system, and it is used by organizations such as the Congressional Budget Office (CBO), the Treasury Department's Office of Tax Analysis, the Joint Committee on Taxation (JCT) and the Urban-Brookings Tax Policy Center (TPC).

In microsimulation analyses, measures of uncertainty, such as standard errors or confidence intervals, can play at least two important roles. First, they can help us understand whether economically meaningful differences in estimates by different organizations reflect different modeling choices or

*For correspondence:
RMcClelland@urban.org

---

1. See *Orcutt (1957)*. The concept applies to studies of not only households but also individuals or firms.

mere sampling variation. Second, they can also help us infer whether apparent differences in outcomes among tax units reflect differences in the population or just sampling variation.[2] For example, a proposed change in tax law may appear progressive when the average tax cut for a lower income group is larger than the tax cut for those in a higher income group in a sample. Yet if the estimates are not significantly different in a statistical sense, the policy change may not be progressive if applied to the population. Unfortunately, neither standard errors nor confidence intervals are routinely calculated or published in the U.S. although some recent research using EUROMOD, the tax and benefit micro-simulation model for the European Union, has included asymptotic standard errors.[3]

For many years, measures of uncertainty were not routinely estimated because they required faster computers than were available. For example, *Citro et al. (1991)* surveyed the state of microsimulation modeling and recommended the use of "new computer technologies to provide enhanced capabilities, such as the ability for a wider group of analysts to apply the models; conduct timely and cost-effective validation studies, including variance estimation and sensitivity analyses." Computational power has since improved substantially yet the few attempts at measuring uncertainty either use simplifying distributional assumptions (*Pudney and Sutherland, 1994*), or simple policy changes that allow for the use of normal approximations (*Goedemé et al., 2013*).

In this paper, we demonstrate the use of bootstrapping to create standard errors and confidence intervals for estimates from TPC's individual income tax model. We also rely on a fundamental insight that may have escaped prior researchers: resampling does not require recalculating taxes for each draw. Instead, in each resample the weights are adjusted and applied to the original sample of tax units. These resamples' weights only need to be produced once and can be used to compute confidence intervals for point estimates of any policy proposal. We also calculate confidence intervals using a normal approximation and compare the results to our bootstrapped intervals. To this end, we estimate five potential changes in tax policy. For (i) a partial restoration of the personal exemption, (ii) an increase in the standard deduction, and (iii) an increase in the top marginal tax rate, we estimate the effects of high, central, and low scenarios. In addition, we consider a revenue-neutral proposal to replace the current-law farm loss deduction with a refundable tax credit based on the total farm loss that should create confidence intervals around point estimates of zero dollar for the average change in tax burden. Finally, we examine a rescission of the tax-free treatment of interest accruing from certain government bonds that should only affect a small group of heterogeneous tax units.

Our results are as follows. First, in most cases the standard errors of the estimated change in taxes for most income groups are small and the 95 percent confidence intervals fit tightly around the point estimates. This means that differences across income groups are statistically significant and, if the same holds true for other tax models, economically meaningful differences among models are also statistically significant. Further, the normal approximation are quite close to the bootstrapped estimates. Second, for the policy alternatives that have high, central and low scenarios, the scenarios leading to larger effects also lead to wider confidence intervals. Finally, when reform only affects a small number of tax units, the average change in taxes may be estimated imprecisely. In these cases, it may be difficult to determine if a proposed policy change is progressive or regressive. In addition, the normal approximation to a confidence interval may differ from the bootstrapped version.

2. Literature review

Orcutt (1957) is the foundational paper on microsimulation. He hypothesized that simulations could be calculated on "a large electronic machine, such as the IBM 704 or the UNIVAC II, or some improved successor to these powerful giants." Both the IBM and the UNIVAC II used vacuum tubes and the IBM could calculate up to 12,000 floating point additions per second.[4] He also noted that for such models "All predictions could be obtained in the form of expected values plus some measure of uncertainty. Or, if desired, they could be in the form of confidence interval estimates."

Pudney and Sutherland (1994) test the finite sample properties of the Central Limit Theorem by estimating the sampling variation of microsimulation models under the assumption of normality. They conclude that "the baseline simulations are reasonably accurate, but that some widely-used

---

2. A tax unit is an individual, or a married couple, that files a tax return or would file a tax return if their income were high enough, along with all dependents of that individual or married couple.
3. See, for example, *Paulus et al. (2019)*.
4. According to Wikipedia, accessed June 19, 2019.

measure of the effects of policy changes may be very imprecise estimates of population effects." For example, when estimating the number of winners and losers from a revenue-neutral reform of family benefits, the confidence interval for the population as a whole is plus or minus four percent, but the confidence interval for single parent families is plus or minus 29 percent. We follow Pudney and Sutherland (1994) and calculate confidence intervals under the normality assumption.

However, our research is closest to *Fiorio (2003)*, who runs a tax microsimulation model "backwards" on a small dataset. The author starts with after-tax income and estimates pre-tax income by adding calculated taxes. The author then uses a bootstrap to estimate the 90 percent confidence intervals due to sampling variation and concludes that, while the confidence interval for the entire sample is small, the distribution is very asymmetric (and therefore badly approximated by the normal distribution) and the interval can be wide in subpopulations.

Creedy et al. (2007) use a parametric bootstrap to sample from the estimated distribution of labor supply elasticities. Their algorithm runs extremely slowly, however, and even with a sample of only 7,000, calculations take weeks to complete. They address the problem by using a small sample to estimate the mean and standard deviation and then create confidence intervals by assuming the parameters are normally distributed.

Goedemé et al. (2013) suggest that if there are no behavioral parameters, or if they are ignored, many changes in the output of tax microsimulations are nearly linear functions of income. In this case, it can be straightforward to calculate the variance of a difference between a baseline distribution of taxes and an alternative scenario using a Taylor first-order linearization of a variance estimate.[5] However, even absent behavioral parameters such as labor supply elasticities, decisions involving whether or not to use various credits or deductions can create difficult-to-map nonlinear function of income.

3. Data

3.1. Public use file

We use TPC's microsimulation individual income tax model in this research. The model is based on a base-year data set composed of a sample of tax-filing 'tax units', which were constructed to be nationally and state representative of the tax filing population for the 2011 tax year, and a sample of non-filers in 2011. The tax-filing sample was a projection of the 2006 public-use file (PUF) produced by the Statistics of Income (SOI) Division of the Internal Revenue Service (IRS). The 2006 PUF contains 145,858 tax unit records with detailed information from federal individual income tax returns filed in calendar year 2006. Non-filers were derived by matching the projected 2011 tax-filing tax units with the March 2012 Current Population Survey (CPS) of the U.S. Census Bureau.[6] Additional variables were imputed for analyzing a wide range of tax policy proposals such as education, health, retirement, estate, and consumption taxes and tax benefits.

To perform revenue and distribution analyses for future years, the 2011 data are further extrapolated for years beyond 2011 using a two-step process based on forecasts and projections from CBO, JCT, the IRS, and U.S. Census Bureau. First, the dollar amounts of income, adjustments, deductions, and credits on each record are inflated by their appropriate forecasted per capita growth rates. Second, TPC uses a linear programming algorithm to adjust the weights on each record so that the major income items, adjustments, and deductions match aggregate targets. TPC also adjusts the overall distribution of adjusted gross income (AGI) to match published information from SOI for available years and projections from CBO for further years.

---

5. This is also known as the delta method.

6. TPC uses published tax data to calculate per-return average growth rates for income, deduction, and other items between 2006 and 2011 by adjusted gross income (AGI) class. These growth rates are used to adjust the dollar amounts on each PUF record. Then, it uses a constrained optimization algorithm to reweight the records to match an extensive set of about 100 national targets and 39 to 51 state targets, depending on AGI classes, for both return counts and dollar amounts. The resulting file is referred to as the 2011 "Look Alike Public Use File" (LAPUF). Afterward, TPC adds information on other demographic characteristics and unreported sources of income by matching the LAPUF with data from the March 2012 Current Population Survey (CPS) of the U.S. Census Bureau. That match also generates a sample of individuals who do not file individual income tax returns ("non-filers"). Finally, the tax model database contains imputations for wealth, education, consumption, health, and retirement-related variables. The full tax model database is a representative national sample of the US population for calendar year 2011.

Table 1. Population and Sample Size of TPC Microsimulation Model, 2019

Expanded Cash Income Level (thousands of 2019 dollars)	Population (Tax Unit)		Sample Size	
	Number (thousands)	Percent of Total	Number (thousands)	Percent of Total
Less than 10	12,323	7.1	8	3.3
10–20	21,651	12.4	16	6.3
20–30	19,523	11.2	16	6.3
30–40	15,976	9.2	14	5.5
40–50	13,399	7.7	12	4.8
50–75	25,288	14.5	27	10.4
75–100	16,983	9.7	20	7.8
100–200	31,667	18.1	43	16.8
200–500	14,026	8.0	35	13.8
500–1,000	1,826	1.1	16	6.3
More than 1,000	837	0.5	40	15.7
All	174,690	100.0	256	100.0

Source: Urban-Brookings Tax Policy Center Microsimulation Model (version 0319-1).
Note: The sample exclude tax units with negative AGI or ECI.

3.2. TPC tax model

Using the extrapolated 2011 data, TPC simulates policy options using a detailed tax calculator that captures most features of the federal individual income tax system. The model reads the data for each tax unit, calculates its tax liability, and outputs the same data with the liability appended. Importantly, this includes the weight for that tax unit. The model's current law baseline reflects major income tax legislation enacted through early 2019, including the Tax Cuts and Jobs Act of 2017 (TCJA).

In its analysis of the distributional effects of a simulated policy change, TPC includes the following federal taxes in its calculation of effective tax rates: individual and corporate income taxes; payroll taxes for Social Security and Medicare; excise taxes; and the estate tax. TPC calculates effective tax rates using a broad measure of income called Expanded Cash Income (ECI), which is defined as adjusted gross income (AGI) plus: above-the-line adjustments (e.g., IRA deduction, student loan interest deduction, self-employed health insurance deduction, etc.), employer paid health insurance and other nontaxable fringe benefits, employee and employer contributions to tax-deferred retirement savings plans, tax-exempt interest, nontaxable Social Security benefits, nontaxable pension and retirement income, accruals within defined benefit pension plans, inside buildup within defined contribution retirement accounts, cash and cash-like (e.g., SNAP) transfer income, employer's share of payroll taxes, and imputed corporate income tax liability.

The effects of tax policy are reported by ECI group. We focus here on groups based on dollar values of ECI. (TPC also reports groups based on percentiles.) ECI is a broad measure of pre-tax income which serves as a proxy for tax units' economic well-being and their ability to pay taxes. In this exercise, we first separate out tax units with either negative ECI or AGI. Then, we categorize the remaining tax units into 11 groups by ECI level. *Table 1* shows these ECI groups and their sample sizes.

4. Method

The bootstrap is a common method for estimating the sampling variation and confidence interval of nonlinear functions applied to a set of observations. It is particularly useful when the distribution of these observations, such as the change in tax liability from a tax policy proposal, is asymmetric. We start by following a description of the bootstrap given in *Kolenikov (2010)*. We then describe the bootstrapping method used in this paper.

We start with a population distribution function $F(x)$ and a sample of n independently and identically distributed observations $(x_1,\ldots,x_n)$. We wish to estimate a statistic θ as a function of the distribution $\theta = T(F)$. For example, θ could be the average change in after-tax income for tax units in a given range of income, and $T(F)$ could be a function of the tax code for those tax units. We estimate this statistic using our sample, $\hat{\theta}_n = T(F_n)$, where F_n is the empirical distribution. In a stratified data set such as the one used in this paper, $\hat{\theta}_n$ is created using a weighted sample. Nevertheless, we do not need to assume that the point estimate $\hat{\theta}_n$ is an unbiased estimate of θ in finite samples.

Of course, the exact estimate depends on our sample. If other samples had been used, the estimate would be different. In general, the estimate will vary across those samples so that we cannot be certain about the distance between the estimate and the statistic in the population. If $T(F)$ is a linear function, we can calculate the variance of the estimate using $V(\hat{\theta}) = E(\hat{\theta} - E\hat{\theta})^2$, where E is the expected value operator.

The Central Limit Theorem implies that if a large number of independent tax units are affected, the distribution of an average tax change can be approximated using the normal distribution. We calculate confidence intervals under this assumption, as $\hat{\theta} \pm C \times \sqrt{V(\hat{\theta})}$, where C is the appropriate two-sided critical value for the normal distribution.

But under the U.S. individual income tax system, taxes owed are a highly nonlinear function of income. In addition, if the sample size is small, confidence intervals may be poorly estimated by methods that assume asymptotic normality of the statistics.

As an alternative, the bootstrapping method estimates various aspects of the distribution of $\hat{\theta}_n$ by repeatedly drawing subsamples. The steps are: (i) draw with replacement R samples of size m $(x_1^*,\ldots x_m^*)$ from the original observations $(x_1,\ldots,x_n)$, (ii) calculate the statistic $\hat{\theta}_m^* = T(F_m^*)$ on each bootstrap sample, where F_m is the empirical distribution function of the bootstrapped sample, and (iii) calculate the appropriate moments F_m^* from the R statistics $\hat{\theta}_m^*$. In the first step the size of the bootstrap sample m is usually set to n, the size of our sample. The number of samples R can be set to any number, although *Kolenikov (2010)* indicates that R is usually set to be between 100 and 1,000. In the second step the statistic is calculated all R times.

The third step relies on the insight that simulated sampling via bootstraps mimics the original sampling process, so that $\hat{\theta}_m$ relates to $\hat{\theta}_n$ as $\hat{\theta}_n$ relates to θ. For example, the potential bias of $\hat{\theta}_n$, defined as $E(\hat{\theta}_n - \theta)$, can be estimated as $\hat{b}_n = E^*(\hat{\theta}_m^* - \hat{\theta}_n)$, where E is an unweighted average across the R samples. Similarly, the variance of θ_n is estimated as $E^*(\hat{\theta}_m - E^*\hat{\theta}_m)^2$, and the calculation of the standard error of $\hat{\theta}_m$ follows naturally.

As with the bias and variance, we can estimate the cumulative distribution function of the estimated parameter $Prob(\hat{\theta}_n - \theta < t)$ with its analogue based on the bootstrapped sample $Prob(\hat{\theta}_m^* - \hat{\theta}_n < t)$. A naïve estimate of a confidence interval would be based on $E^*I(\hat{\theta}_m^* - \hat{\theta}_n \leq t)$, where I is an indicator function equal to 1 if the condition holds and 0 otherwise. It has long been recognized that the bootstrap can also be used to estimate a bias-corrected confidence interval as $EI(2\hat{\theta}_n - \hat{\theta}_m \leq t)$.[7]

Our goal is to estimate confidence intervals for the point estimates $\hat{\theta}_n$, such as average tax changes and percent change in after-tax income, of the proposed changes in the individual income tax from the TPC's microsimulation model. To obtain confidence intervals of these point estimates, we bootstrap Tax Model runs in two steps: (i) producing a series of alternative analytical weights of tax units in the Tax Model and (ii) calculating relevant bootstrapped statistics using these alternative weights.

4.1. Producing a series of alternative analytical weights

Similar to a standard bootstrap exercise, we want to draw observations with replacement from the original data set after tax liabilities have been calculated. The idea is that this set of alternative analytical weights can be used to infer the population characteristics of the original Tax Model sampled tax

---

7. This correction is made under the assumption that the bias is a constant. For biases that are a function of θ, a "bias-corrected and accelerated" confidence interval may be used.

units. However, because the Tax Model observations carry different weights (and these observations represent one hundred eighty-four million US tax units in 2019), the process is not straightforward. Here we rely on the Stata bsweight command formulated in *Kolenikov (2010)* to produce a series of replicated observations' weights. In particular, each set of alternative analytical weights is constructed such that observations are sampled with replacement, for each ECI group h, set $m_h=n_h-1$, and the observations weights w_i within each ECI group are scaled such that totals by ECI group match the totals derived using the original weights.[8]

In the current exercise, we produce 201 sets of alternative analytical weights.[9] Producing a large set of alternative analytical weights is time consuming although they only need to be produced once and can be used to compute confidence intervals for point estimates of any policy proposal. In addition, we can use a smaller number of these alternative analytical weights to speed up the calculations of bootstrapped statistics (more on this below).

4.2. Calculating relevant statistics

In this exercise, we calculate confidence intervals using both the normal approximation and the boot-strapping approach. A complication occurs with the normal approximation because there are several formulas for calculating standard errors of weighted averages. We address this by following *Gatz and Smith (1995)* and use the standard errors calculated from our bootstrap procedure. The normal approximation of the 95 percent confidence interval is thus calculated as the average change in taxes, plus or minus 1.96 times the bootstrapped standard error (BSE). Gatz and Smith also describe several formulas for approximating normal standard errors of weighted averages and for comparison we use one of their formulas GSSE = $\sqrt{\frac{1}{n}\left(\frac{1}{w_i}\right)\sum w_i\left(x_i-\bar{x}\right)^2}$. We then calculate confidence intervals as the average change in taxes plus or minus 1.96 times the GSSE.

To calculate the bootstrap standard errors and confidence intervals we rely on built-in bias-corrections (BC) available in the Stata bootstrap command bs4rw.[10] In particular, we use bs4rw with the summarize command to obtain bootstrapped statistics for totals and averages and with the ratio command to obtain bootstrapped statistics for percentages.

In practice, the program runs progressively more slowly as the set of alternative analytical weights become larger. In this exercise, we use a set of 201 alternative weights, which allow us to theoretically rank the bootstrapped estimates with an interval of 0.5% from 0th to 100th percentiles.[11]

As a caution, it is a good practice to inspect the derived bootstrapped statistics in detail. For example, when a policy proposal only affects a few tax units in an income group, many of these 201 simulations may not pick up any tax units affected by the policy and as a result the bootstrapped statistics will be calculated based on only the subset of simulations where affected tax units were drawn. In such case, it may be useful to focus as well on bootstrapped statistics of the fraction of tax units affected by the proposal and the average tax changes per tax units affected.

5. Results

We apply the bootstrapping methodology described in the previous section to examine TPC's micro-simulation distributional estimates of five alternative tax policy proposals if implemented in 2019. These proposals are (i) a partial restoration of the personal exemption eliminated by TCJA, (ii) an

---

8. If other values of m_h are used, the weights must be scaled to correct for a bias. This is of greater concern for very small datasets. See *Rao and Wu (1988)* for more information.

9. It is 201 because it allows us to calculate percentiles precisely. For example, if we want to calculate percentiles (from 0 to 100) precisely, we need 101 points. Likewise, if we want to calculate 1/2 of percentiles (0 to 200) precisely, we need 201 points.

10. As described in *Kolenikov (2010)*, bs4rw "is an analogue of the official bootstrap command that uses the replicate weights instead of actually resampling the data in Stata memory". The command produces two types of biased corrections for confidence intervals, bias-corrected (BC) and bias-corrected and accelerated (BCA). We opt for the BC option and hence implicitly assume that bias is not a function of the parameter of interest. It also helps that we use 201 alternative weights, a relatively large number of replicates which help ensure approximate normality.

11. It took about 80 minutes to obtain bootstrapped statistics for all point estimates in the Tax Model's standard summary distribution table using 201 alternative weights. In contrast, it took slightly more than ten hours (approximately 7.5 times longer) when using 1,001 alternative weights.

increase in the standard deduction, (iii) an increase in the top statutory income tax rate, (iv) replacing the current-law farm loss deduction with a revenue-neutral refundable tax credit based on the total farm loss, and (v) a rescission of the tax-free treatment of interest accruing from certain government bonds. We chose these policy proposals because point estimates should be relatively more uncertain when proposals affect fewer, more diverse tax units. In addition, to allow for confidence intervals to vary with the point estimates, we also varied the relevant parameters of the first three proposals to three levels: high, central and low scenarios. In our discussion, we only present a subset of the estimates for tractability. The complete estimates are available in an online appendix.[12]

5.1. Partially restore the personal exemption

Before 2018, tax units could claim a personal exemption for every adult and qualified dependent. In 2017 the exemption amount was $4,050. For many low-income households, this exemption eliminated their tax liability.[13] The exemption phased out at high levels of adjusted gross income (AGI), with the 2017 phase-out beginning at an AGI of $261,500 for singles, $287,650 for heads of household, $313,800 for married couples filing a joint return and $156,900 for married couples filing separately. Personal exemptions were completely phased out when AGI reached $384,000 for singles, $410,150 for heads of households, $436,300 for married couples and $218,150 for married couples filing separately. TCJA eliminated the personal exemption and increased the standard deduction and the child tax credit between 2018 and 2025.[14]

We consider a scenario that would restore the personal exemption (and phase-out) partially in 2019 to $4,000, and estimate how taxes would change. We then repeat this exercise with scenarios that would restore the personal exemption to $2,000 or $1,000. In each scenario, each tax unit's tax change under the proposal is calculated as the difference between the tax unit's tax liability with the exemption restored and its tax liability under current law. We then apply the bootstrapping procedure and calculate estimates, such as the average change in income for tax units in a given range of income, their standard errors, and confidence intervals.

Recall that we examine uncertainties around a point estimate due to sampling variation. For this proposal, sampling variation come from three sources: (i) variation in the number of exemptions across tax units, (ii) variation in tax units' tax rates and (iii) variation in the amount of the exemption each tax unit can claim. For (ii), a tax reduction for a full $4,000 per person exemption reduces tax by only $480 for a single tax unit with no dependents in the 12 percent tax bracket but $1,280 for a similar tax unit in the 32 percent bracket. For (iii), low income tax units whose taxable income is less than the available exemptions under the proposal ($4,000 multiplied by the number of personal exemptions) will only be able to claim an exemption amount up to their taxable income while high income tax units may face a phase-out and can only use part or none of the available exemptions.

To conserve space, we present our results for a small number of ECI groups. Here we use three: a low-income group of $75,000 to $100,000; a middle-income group of $200,000 to $500,000; and a high-income group of more than $1 million. The results for all income groups are available in the online appendix. Because ECI accounts for components of income that are not part of AGI and tax proposals generally depend on AGI, a proposal may affect tax units in an ECI group that at first glance seems counterintuitive. For example, a small number of tax units with ECIs in excess of $1 million have AGI below the phase-out limits of the personal exemption and, as a result, may receive a tax cut under this proposal.

The point estimates in the last column of *Table 2* indicate that restoring the personal exemption at $4,000 would reduce the average tax burden of tax units in the low-income group by $1,023, in the middle-income group by $2,459, and in the high-income group by $8.[15] The middle-income group get

---

12. The complete online appendix for this paper, including tables and links to Stata code, is available for downloading at the Urban Institute's Data Catalog: https://datacatalog.urban.org/dataset/estimating-confidence-intervals-tax-microsimulation-model

13. TPC Briefing Book https://www.taxpolicycenter.org/briefing-book/what-are-personal-exemptions

14. If the personal exemption had been entirely restored in 2019 it would be $4,200 using the chained CPI to index the exemption and $4,250 using the CPI-U to index the exemption.

15. In all the tables we present, tax cuts are shown as negative tax changes, and tax increases are shown as positive tax changes.

Table 2. Partially Restore the Personal Exemption to $4,000, Selected Income Groups, 2019

Expanded Cash Income Level (thousands of 2019 dollars)	Tax Units with Tax Cut	Average Tax Change ($)	Percent Change in After-Tax Income	Average Federal Tax Change ($)
	Percent of Tax Units			
75–100	94.239 (0.314)	−1,085.1 (4.5)	1.370 (0.008)	−1,022.6 (5.7)
	[93.599, 94.843]	[−1,094.0, −1,076.8]	[1.351, 1.383]	[−1,032.4, −1,009.9]
200–500	98.470 (0.080)	−2,496.9 (10.7)	1.077 (0.006)	−2,458.6 (10.9)
	[98.296, 98.615]	[−2,515.1, −2,475.3]	[1.065, 1.088]	[−2,479.3, −2,435.9]
More than 1,000	0.620 (0.176)	−1,214.3 (242.8)	0.000 (0.000)	−7.5 (2.6)
	[0.334, 1.054]	[−1,680.9, −789.1]	[0.000, 0.001]	[−13.2, −2.9]
All	62.332 (0.131)	−1,216.4 (2.8)	0.933 (0.004)	−758.2 (1.7)
	[62.033, 62.542]	[−1,220.9, −1,210.7]	[0.927, 0.943]	[−761.6, −754.6]
Addendum: Confidence Intervals Calculated Using Bootstrap and Using Normal Approximation of BSE				
75–100 Bootstrap	[93.599, 94.843]	[−1,094.0, −1,076.8]	[1.351, 1.383]	[−1,032.4, −1,009.9]
Normal with BSE	[93.624, 94.854]	[−1,093.9, −1,076.2]	[1.354, 1.386]	[−1,033.7, −1,011.4]
200–500 Bootstrap	[98.296, 98.615]	[−2,515.1, −2,475.3]	[1.065, 1.088]	[−2,479.3, −2,435.9]
Normal with BSE	[98.313, 98.628]	[−2,517.9, −2,475.9]	[1.065, 1.089]	[−2,479.9, −2,437.4]
More than 1,000 Bootstrap	[0.334, 1.054]	[−1,680.9, −789.1]	[0.000, 0.001]	[−13.2, −2.9]
Normal with BSE	[0.275, 0.966]	[−1,690.2, −738.4]	[0.000, 0.001]	[−12.6, −2.4]
All Bootstrap	[62.033, 62.542]	[−1,220.9, −1,210.7]	[0.927, 0.943]	[−761.6, −754.6]
Normal with BSE	[62.075, 62.589]	[−1,221.8, −1,210.9]	[0.926, 0.941]	[−761.6, −754.8]

Source: Urban–Brookings Tax Policy Center Microsimulation Model (version 0319-1).

Notes: Proposal would partially restore the personal exemption of $4,000 in 2019.

Numbers in parentheses are bootstrapped standard errors (BSE). Numbers in brackets are confidence intervals.

the largest average tax cut. The corresponding width of each point estimate's confidence interval is $23, $43, and $10, respectively.

But the first column of *Table 2* shows that only a very small number of tax units in the high-income group (0.6 percent) receives a tax cut. High-income tax units whose AGI are larger than their phase-out limits do not benefit from the proposal because their available personal exemptions are completely phased out.

As the second column of *Table 2* shows, the average tax cut among tax units receiving a tax cut under the proposal are $1,085, $2,497 and $1,214 for the low, middle, and high-income group, respectively. The average tax cut for the middle-income group is more than twice the size of the cut for the low-income group mainly because the middle-income group faces a higher income tax rate and a greater share of tax units in the middle-income group file as married-filing-jointly, allowing them one more exemption than those filing as single or as head of household. In contrast, the smaller benefit for high-income tax units -- even with the large average tax rate -- occurs because some of them are in the phase-out range. As a result, their tax cut is only half of the middle-income group's average based on the point estimates.

Below the point estimates and standard errors in *Table 2* are the confidence intervals around these point estimates of average tax changes among tax units receiving a tax cut. The confidence interval around the low-income group's tax cut of $1,085 is [$1,077, $1,094], the confidence interval around the middle-income group's tax cut of $2,497 is [$2,475, $2,515], while the confidence interval around the high-income group's tax cut of $1,214 is [$789, $1,681]. Note that the point estimate for the low-income group is within the confidence interval of the high-income group. That is, the estimated effects of the high and low groups are not statistically different. In the addendum, we compare confidence intervals using a normal approximation with BSE to the confidence intervals created with the bootstrap. Overall, the normal approximation works well, although its confidence interval of the high-income group is not as wide as the bootstrap interval.

The high-income tax units' point estimate is much less precise because the proposal only affects a very small number of sample observations in this income group and the change among the few observations is large. *Table 3* shows that 40,100 tax unit observations represent 836,500 tax units in the top-income group. However, only 71 observations represent 5,190 tax units that receive a tax cut under the $4,000 personal exemption proposal. In other

Table 3. Comparison of Population and Sample Size Affected by Partially Restoring the Personal Exemption to $4,000, Selected Income Groups, 2019

Expanded Cash Income Level (thousands of 2019 dollars)	All Tax Units		Tax Units with Tax Change					
			Population		Sample		Tax Change ($)	
	Population Number	Sample Number	Number	Share within Income Group	Number	Share within Income Group	Minimum	Maximum
200–500	14,026,200	35,400	13,815,800	98.50%	32,220	91.00%	14	11,000
More than 1,000	836,500	40,100	5,190	0.62%	71	0.18%	15	3,840

Source: Urban–Brookings Tax Policy Center Microsimulation Model (version 0319-1).

words, the proposal affects 0.62% of tax units in the high-income population but only 0.18% of those in the sample. Tax units' tax cuts vary from $15 to $3,840 so the weighted average tax cuts from different mixes of observations can be quite different. In contrast, *Table 3* shows that out of 35,400 tax unit observations representing 14.0 million tax units in the middle-income group, 32,220 observations representing 13.8 million tax units would receive a tax cut under the $4,000 personal exemption proposal. Their tax cuts vary from $14 to $11,000 but because most of the sampled tax units in the middle-income group receive a tax cut, the weighted average tax cuts from different mixes of observations should be quite similar.

5.2. Increase the standard deduction

We next consider a more complicated change in tax liability. When filing tax returns, tax units can reduce their taxable income using either itemized deductions or the standard deduction. Because tax units with the same taxable income and filing status may have different amounts of expenses eligible for itemized deductions, some will choose to itemize and others will not.

In 2019, the current law standard deduction amounts are $12,200 for single filers and married couples filing separately, $18,350 for heads of household and $24,400 for married couples filing jointly. We consider scenarios that would increase the standard deduction amounts by $4,000, $2,000 and $1,000.[16] In each scenario we compare the itemized deductions to the standard deduction for each tax unit and select the deduction that results in the lowest tax liability. We describe the effects on tax units with incomes between $50,000 and $75,000 (low-income group), between $200,000 and $500,000 (middle-income group), and greater than $1 million (high-income group).

Sampling variation in estimating the effects of this change comes from three sources: (i) variation in the amount of the increased standard deduction that can be used (ii) variation in the amount of itemized deductions and, (iii) variation in the tax rates. The first occurs because married couples filing jointly receive twice the amount of single filers but they may not be able to deduct the full additional amount available if the increase makes the total deduction greater than their taxable income. The second comes from variation in the amount of expenses that can be deducted. Tax units who were originally claiming the standard deduction before the increase will continue to do so. Tax units with itemized deductions greater than the increased standard deduction will still itemize. Those with itemized deductions greater than the old level of standard deductions but less than the new level will stop itemizing.

Because there are more potential sources of variation from raising the standard deduction than from the partial restoration of the personal exemption, this proposal may generate more variation in average benefits for each income group, and possibly wider confidence intervals. In particular, there may be more variation among tax units with similar incomes.

The variation across income groups can be seen in *Table 4*. Under the high scenario in which the standard deduction is increased by $4,000, tax units in the low-income group see an average tax cut of $501, those in the middle-income group see a tax cut of a $1,286, and those high-income group see an average tax cut of $491.

This pattern reflects two offsetting factors. On the one hand, high-income tax units are more likely to have a sufficient amount of deductions to itemize even when the standard deduction is increased. As shown in the first column in *Table 4*, nearly 80 percent of tax units in low and middle-income groups benefit from a higher standard deduction, but only 23 percent for those tax units with incomes greater than $1 million benefit.[17] On the other hand, among those units that take the standard deduction, higher-income tax units face higher marginal rates, so an increased deduction would reduce tax liabilities more for higher-income tax units than for lower-income tax units. This is shown in the second column in *Table 4*, in which those in the low-income group that use the standard deduction see an average tax reduction of $626 while those in the middle-income group see an average tax reduction of $1,622. Those in the high-income group see an average reduction of $2,131.

---

16. These increases are for single filers and married couples filing separately. Heads of household and married couples receive 1.5 times and double of the singles' increases, respectively.

17. The TPC table TM18-0001 shows that under current law, the share of itemizers in 2018 is only 7 percent among those with income between $50,000 and $75,000, 47 percent among those with income between $200,000 and $500,000, and 82 percent for those who earn more than $1 million. https://www.taxpolicycenter.org/model-estimates/impact-itemized-deductions-tax-cuts-and-jobs-act-jan-2018/t18-0001-impact-number

Table 4. Increase the Standard Deduction to $4,000, Selected Income Groups, 2019

Expanded Cash Income Level (thousands of 2019 dollars)	Tax Units with Tax Cut	Average Tax Change ($)	Percent Change in After-Tax Income	Average Federal Tax Change ($)
	Percent of Tax Units			
50–75	79.993 (0.399)	−625.8 (2.4)	0.921 (0.006)	−500.6 (2.9)
	[79.140, 80.676]	[−629.7, −620.9]	[0.909, 0.931]	[−505.9, −494.4]
200–500	79.289 (0.329)	−1,621.6 (5.3)	0.563 (0.003)	−1,285.7 (7.0)
	[78.622, 79.919]	[−1,631.7, −1,612.3]	[0.556, 0.569]	[−1,296.7, −1,268.5]
More than 1,000	23.048 (0.548)	−2,130.7 (26.4)	0.023 (0.001)	−491.1 (13.4)
	[22.019, 24.250]	[−2,180.3, −2,077.4]	[0.021, 0.025]	[−521.0, −466.6]
All	58.636 (0.141)	−889.3 (1.8)	0.642 (0.003)	−521.4 (1.2)
	[58.391, 58.910]	[−893.3, −886.0]	[0.638, 0.648]	[−523.8, −519.0]
Addendum: Confidence Intervals Calculated Using Bootstrap and Using Normal Approximation of BSE				
50–75	Bootstrap	[79.140, 80.676]	[−629.7, −620.9]	[0.909, 0.931]	[−505.9, −494.4]
	Normal with BSE	[79.211, 80.775]	[−630.4, −621.2]	[0.910, 0.932]	[−506.3, −494.9]
200–500	Bootstrap	[78.622, 79.919]	[−1,631.7, −1,612.3]	[0.556, 0.569]	[−1,296.7, −1,268.5]
	Normal with BSE	[78.645, 79.934]	[−1,632.0, −1,611.1]	[0.556, 0.570]	[−1,299.5, −1,271.9]
More than 1,000	Bootstrap	[22.019, 24.250]	[−2,180.3, −2,077.4]	[0.021, 0.025]	[−521.0, −466.6]
	Normal with BSE	[21.974, 24.122]	[−2,182.5, −2,079.0]	[0.021, 0.024]	[−517.3, −464.9]
All	Bootstrap	[58.391, 58.910]	[−893.3, −886.0]	[0.638, 0.648]	[−523.8, −519.0]
	Normal with BSE	[58.361, 58.912]	[−892.7, −885.8]	[0.637, 0.647]	[−523.8, −519.1]

Source: Urban-Brookings Tax Policy Center Microsimulation Model (version 0319-1).

Notes: Proposal would increase the standard deduction by $4,000 for single filers and married couples filing separately, $8,000 for married couples filing jointly, and $6,000 for heads of household in 2019.

Numbers in parentheses are bootstrapped standard errors (BSE). Numbers in brackets are confidence intervals.

Unlike the personal exemption, a sizable share of each income group receives a benefit and the confidence intervals are small for all groups, although the confidence interval is larger in the high-income group, similar to the personal exemption proposal. The confidence interval for the average tax cut ranges from $494 to $506 for the low-income group, from $1,269 to $1,297 for the middle-income group and from $467 to $521 for the high-income group. Overall, the confidence intervals using normal approximation with BSE are similar to the bootstrapped confidence intervals.

5.3. Increase in the top tax rate

TCJA lowered the top rate from 39.6 to 37 percent starting in 2018. The top tax rate applies to taxable income over $500,000 for single filers and over $600,000 for married couples filing jointly. Tax brackets are adjusted annually for inflation. We consider an increase in the top rate to 41, 39 or 38 percent. Increasing the top rate to 41 percent would only affect about 0.6 percent of all tax units, almost exclusively those with expanded cash income of over $500,000. We therefore describe the effects on tax units with incomes between $200,000 and $500,000 (low-income group), between $500,000 and $1 million (middle-income group), and with incomes greater than $1 million (high-income group).

Sampling variation comes from four sources: (i) from the mix of single tax filers and married couples filing jointly in an income group, (ii) from the share of tax units in each income group to which the top rate applies, (iii) from the amount of income above the threshold, which determines the amount of the tax increase, and (iv) from the small number of tax units that will be affected by the complex interaction of the regular income tax system and the Individual Alternative Minimum Tax (AMT).[18]

Successively higher income groups have greater shares of affected tax units and higher average changes among those affected, as shown in *Table 5*. Only 0.06 percent of tax units in the low-income group would see a tax increase while about 18 percent of the tax units in the middle-income group and more than three quarters of those in the high-income group would pay more taxes. The average increase among units with a tax change in the low-income group is just over $2,000. This increases to over $4,000 for the middle-income group and to more than $52,000 in the high-income group.

The small number of affected records in the low-income group results in relatively large confidence intervals. The interval around the share of affected tax units ranges from 0.038 to 0.091, while the interval around the average tax cut among those affected ranges from $1,349 to $2,956. The confidence interval calculated from the normal approximation with BSE, ranging from $1,173 to $2,916, is slightly wider than the bootstrapped interval. The other income groups have much smaller confidence intervals and the normal approximation is much closer to the bootstrapped confidence intervals. Overall, the confidence intervals using normal approximation with BSE are similar to the bootstrapped confidence intervals.

5.4. Revenue neutral replacement of a deduction for farm losses with a refundable tax credit

Having found generally small confidence intervals in the above three cases, we examine two additional cases in which the confidence intervals might be more sensitive to policy changes. In this section, we estimate a revenue-neutral change of replacing the current-law farm loss deduction with a refundable tax credit based on the total farm losses in 2019. Under the current law, tax units with a farm loss can deduct the loss up to a limit.[19] The proposal would repeal this farm loss deduction but allow tax units to claim a refundable tax credit equal to 82.5% of the product of tax units' statutory income tax rates and their total, i.e. unlimited, farm losses (the 82.5% factor brings about the proposal's revenue neutrality.) Because the proposal is revenue-neutral, it would generate both winners and

---

18. Some tax units are required to calculate their liability under the rules for the regular income tax and under the AMT rules and then pay the higher amount. TCJA reduced the number of affected tax units to only 200,000 filers in 2018. For more on the Alternative Minimum Tax, see http://www.taxpolicycenter.org/briefing-book/what-amt
19. There are three relevant limits: (a) at-risk limits limiting deductions for losses from most business or income-producing activities including farming, (b) passive activity limits generally limiting deductions for losses from passive activities to not be greater than income from passive activities, (c) excess business loss limitation limiting deductions for losses of businesses including farming (losses above this limit can be carried over to the next tax year). See https://taxmap.irs.gov/taxmap/pubs/p225-017.htm for more detail.

Table 5. Increase the Top Income Tax Rate to 41 Percent, Selected Income Groups, 2019

Expanded Cash Income Level (thousands of 2019 dollars)	Tax Units with Tax Increase		Percent Change in After-Tax Income	Average Federal Tax Change ($)
	Percent of Tax Units	Average Tax Change ($)		
200–500	0.062 (0.014)	2,044.6 (444.8)	−0.001 (0.000)	1.3 (0.3)
	[0.038, 0.091]	[1,349.4, 2,956.2]	[−0.001, 0.000]	[0.8, 1.9]
500–1,000	18.225 (0.530)	4,127.8 (111.2)	−0.147 (0.006)	752.3 (29.8)
	[17.236, 19.150]	[3,909.4, 4,363.9]	[−0.156, −0.135]	[691.6, 800.8]
More than 1,000	77.965 (0.571)	52,878.8 (1,006.8)	−1.917 (0.049)	41,226.8 (808.0)
	[76.552, 78.925]	[51,126.9, 54,963.1]	[−1.996, −1.827]	[39,944.2, 43,177.1]
All	0.569 (0.006)	36,127.7 (759.5)	−0.253 (0.005)	205.7 (3.9)
	[0.555, 0.581]	[34,865.4, 37,852.8]	[−0.264, −0.246]	[199.5, 214.6]
Addendum: Confidence Intervals Calculated Using Bootstrap and Using Normal Approximation of BSE				
200–500 Bootstrap	[0.038, 0.091]	[1,349.4, 2,956.2]	[−0.001, 0.000]	[0.8, 1.9]
Normal with BSE	[0.034, 0.089]	[1,172.8, 2,916.3]	[−0.001, 0.000]	[0.7, 1.9]
500–1,000 Bootstrap	[17.236, 19.150]	[3,909.4, 4,363.9]	[−0.156, −0.135]	[691.6, 800.8]
Normal with BSE	[17.186, 19.264]	[3,909.9, 4,345.6]	[−0.158, −0.135]	[693.9, 810.6]
More than 1,000 Bootstrap	[76.552, 78.925]	[51,126.9, 54,963.1]	[−1.996, −1.827]	[39,944.2, 43,177.1]
Normal with BSE	[76.846, 79.084]	[50,905.5, 54,852.0]	[−2.013, −1.822]	[39,643.2, 42,810.5]
All Bootstrap	[0.555, 0.581]	[34,865.4, 37,852.8]	[−0.264, −0.246]	[199.5, 214.6]
Normal with BSE	[0.557, 0.582]	[34,639.1, 37,616.4]	[−0.262, −0.244]	[198.1, 213.2]

Source: Urban-Brookings Tax Policy Center Microsimulation Model (version 0319-1).
Notes: Proposal would increase the top income tax rate to 41 percent in 2019.
Numbers in parentheses are bootstrapped standard errors (BSE). Numbers in brackets are confidence intervals.

losers. Specifically, this proposal would help tax units with large farm losses at the expense of those with small losses. However, it is not clear how it would affect tax units in different income groups.

We examine a proposal with farm losses because of the large sampling variation due to: (i) few tax units report a farm loss; (ii) there is wide variation in farm losses across observations; and (iii) the tax rate will have a tangible impact on those with a farm loss. [20]

As shown in *Table 6*, this proposal would on average minimally reduce tax burden of tax units with income between $40,000 and $50,000, slightly increase tax burden of tax units with income between $500,000 and $1 million, and slightly reduce tax burden of tax units with income more than $1 million. By construction, the proposal would not change tax burden of tax units overall. Although the bootstrapped confidence intervals and the interval based on normal approximation are quite close in most cases, there are several exceptions.

Based on the bootstrapped confidence intervals, it is inconclusive whether the average change in the tax burden was positive or negative for tax units with incomes between $40,000 and $50,000 or between $500,000 and $1 million. In each of these income groups, some tax units would see an increase in their tax burden and others would see a decrease, leading to confidence intervals for the average tax change that include zero. On the other hand, the confidence interval for the average change in the tax burden for tax units with incomes of $1 million or more is strictly less than zero.

In contrast, the confidence interval assuming a normal distribution with BSE is strictly greater than zero while the bootstrapped confidence interval for tax units with incomes between $500,000 and $1 million straddles zero. Although the bootstrapped and normal confidence intervals are similar for average tax increases, they diverge for tax units facing a tax cut. For example, the bootstrapped confidence interval for tax units in the $500,000 to $1 million income group is [−8,022, −1,417] but [−6,541, −391] for the normal confidence interval. This leads to a bootstrapped confidence interval for the overall change of [−2.7, 32.6] which includes zero but a normal confidence interval of [1.0, 36.5] which omits zero.

5.5. Rescind the tax-free treatment of interest accruing from certain government bonds

Finally, we estimate the effect of rescinding the tax-exempt status of interest from government bonds. Sampling variation in this case comes from three sources: (i) few tax units have tax-exempt interest; (ii) there is wide variation in interest income across observations; and (iii) the tax rate will have a tangible impact on those with tax-exempt interest income.

Municipal (i.e. state and local) bond interest is exempted from federal income tax. In 2018, there were $3.8 trillion in municipal bonds outstanding.[21] The exemption primarily benefits higher-income individuals. In our dataset, 17% out of approximately 256,000 records receive tax exempt interest, representing 3% out of 175 million tax units in the population.[22] The unweighted distribution of tax-exempt interest income is extremely wide and asymmetric, with a median of $1,647 and a mean of $12,605.

As shown in *Table 7*, rescinding the tax-exempt status of interest from government bonds would increase tax liabilities of the top income group the most – an average of $13,545 for those with ECI more than $1 million. We find that the confidence intervals are tightly wrapped around the point estimates for most income groups and that the normal approximation with BSE is very similar to the bootstrapped confidence interval.

However, for the group with income between $40,000 and $50,000, the confidence interval is not as tight, and the results are asymmetric, with the average tax increase among tax units with a tax increase being $383, and a confidence interval ranging from $174 to $816. In the addendum, the confidence interval calculated with the normal approximation is noticeably different, ranging

---

20. In our dataset, 5,380 (2%) out of approximately 256,000 records reported a farm loss, representing 1.4 million (0.8%) out of 175 million tax units in the population. The losses range from $4 to $23 million.
21. SIFMA Research: https://www.sifma.org/resources/research/us-municipal-securities-holders/
22. Discussing tax incidence is beyond the scope of this paper, but *Galper et al. (2013)* show that the current distributional methodology does not capture the shift of benefits of tax-exempt interest. Since the tax exemption affects the relative prices of taxable and tax-exempt bonds, making yields on tax-exempt bonds fall and yields on taxable bonds rise, the benefit of the exemption for holders of tax-exempt bonds is overstated, while some of the benefit of the exemption accrues to holders of taxable bonds in the form of higher yields.

Table 6. Replace Farm Loss Deduction with a Revenue-Neutral Refundable Tax Credit, Selected Income Groups, 2019

Expanded Cash Income Level (thousands of 2019 dollars)	Tax Units with Tax Cut		Tax Units with Tax Increase		Average Federal Tax Change ($)
	Percent of Tax Units	Average Tax Change ($)	Percent of Tax Units	Average Tax Change ($)	
40–50	0.109 (0.029)	−1,368.9 (537.1)	0.267 (0.065)	404.1 (101.8)	−0.4 (0.7)
	[0.062, 0.180]	[−2,743.8, −685.0]	[0.166, 0.462]	[234.1, 614.2]	[−2.0, 1.0]
500–1,000	0.531 (0.109)	−3,465.7 (1,568.9)	1.338 (0.170)	2,779.1 (400.8)	18.8 (9.1)
	[0.386, 0.812]	[−8,022.0, −1,416.6]	[0.990, 1.652]	[2,082.2, 3,748.6]	[−2.7, 32.6]
More than 1,000	0.994 (0.077)	−15,499.8 (4,948.4)	0.922 (0.078)	7,473.2 (787.6)	−85.1 (50.9)
	[0.852, 1.171]	[−26,531.3, −7,571.9]	[0.781, 1.073]	[6,012.7, 9,182.7]	[−176.3, −3.0]
All	0.207 (0.013)	−1,909.9 (221.3)	0.485 (0.020)	770.2 (47.8)	−0.2 (0.5)
	[0.186, 0.230]	[−2,420.9, −1,566.6]	[0.450, 0.522]	[700.7, 884.2]	[−1.3, 0.8]

Addendum: Confidence Intervals Calculated Using Bootstrap and Using Normal Approximation of BSE

		Percent of Tax Units	Average Tax Change ($)	Percent of Tax Units	Average Tax Change ($)	Average Federal Tax Change ($)
40–50	Bootstrap	[0.062, 0.180]	[−2,743.8, −685.0]	[0.166, 0.462]	[234.1, 614.2]	[−2.0, 1.0]
	Normal with BSE	[0.052, 0.166]	[−2,421.6, −316.2]	[0.140, 0.394]	[204.6, 603.6]	[−1.8, 1.0]
500–1,000	Bootstrap	[0.386, 0.812]	[−8,022.0, −1,416.6]	[0.990, 1.652]	[2,082.2, 3,748.6]	[−2.7, 32.6]
	Normal with BSE	[0.318, 0.744]	[−6,540.8, −390.7]	[1.005, 1.670]	[1,993.5, 3,564.6]	[1.0, 36.5]
More than 1,000	Bootstrap	[0.852, 1.171]	[−26,531.3, −7,571.9]	[0.781, 1.073]	[6,012.7, 9,182.7]	[−176.3, −3.0]
	Normal with BSE	[0.843, 1.145]	[−25,198.5, −5,801.0]	[0.769, 1.076]	[5,929.6, 9,016.9]	[−184.8, 14.6]
All	Bootstrap	[0.186, 0.230]	[−2,420.9, −1,566.6]	[0.450, 0.522]	[700.7, 884.2]	[−1.3, 0.8]
	Normal with BSE	[0.182, 0.232]	[−2,343.6, −1,476.2]	[0.445, 0.524]	[676.5, 863.9]	[−1.3, 0.8]

Source: Urban-Brookings Tax Policy Center Microsimulation Model (version 0319-1).
Notes: Proposal would replace current-law farm loss deduction with a refundable tax credit equal to 82.5% of total farm loss times the statutory income tax rate in 2019. Numbers in parentheses are bootstrapped standard errors (BSE). Numbers in brackets are confidence intervals.

Table 7. Rescind the Tax-free Treatment of Interest Accruing from Certain Municipal Bonds, Selected Income Groups, 2019

Expanded Cash Income Level (thousands of 2019 dollars)	Tax Units with Tax Increase		Percent Change in After-Tax Income	Average Federal Tax Change ($)
	Percent of Tax Units	Average Tax Change ($)		
40–50	0.598 (0.105)	382.8 (142.4)	−0.006 (0.002)	2.3 (0.9)
	[0.412, 0.830]	[174.3, 815.8]	[−0.011, −0.002]	[1.0, 4.5]
200–500	10.871 (0.289)	2,693.2 (124.2)	−0.128 (0.006)	292.8 (14.4)
	[10.348, 11.518]	[2,442.2, 2,965.2]	[−0.141, −0.116]	[263.7, 323.0]
More than 1,000	50.162 (0.535)	27,001.5 (1,425.9)	−0.630 (0.033)	13,544.6 (729.8)
	[49.325, 51.238]	[24,610.9, 30,036.2]	[−0.697, −0.576]	[12,302.0, 15,143.7]
All	2.924 (0.050)	4,272.1 (140.4)	−0.154 (0.005)	124.9 (3.9)
	[2.826, 3.014]	[4,021.8, 4,613.3]	[−0.164, −0.145]	[118.0, 133.8]
Addendum: Confidence Intervals Calculated Using Bootstrap and Using Normal Approximation of BSE				
40–50 Bootstrap	[0.412, 0.830]	[174.3, 815.8]	[−0.011, −0.002]	[1.0, 4.5]
Normal with BSE	[0.393, 0.803]	[103.8, 661.9]	[−0.010, −0.001]	[0.6, 4.0]
200–500 Bootstrap	[10.348, 11.518]	[2,442.2, 2,965.2]	[−0.141, −0.116]	[263.7, 323.0]
Normal with BSE	[10.305, 11.437]	[2,449.8, 2,936.6]	[−0.141, −0.116]	[264.6, 321.1]
More than 1,000 Bootstrap	[49.325, 51.238]	[24,610.9, 30,036.2]	[−0.697, −0.576]	[12,302.0, 15,143.7]
Normal with BSE	[49.113, 51.212]	[24,206.9, 29,796.1]	[−0.695, −0.565]	[12,114.2, 14,975.1]
All Bootstrap	[2.826, 3.014]	[4,021.8, 4,613.3]	[−0.164, −0.145]	[118.0, 133.8]
Normal with BSE	[2.825, 3.022]	[3,996.9, 4,547.4]	[−0.163, −0.144]	[117.3, 132.5]

Source: Urban-Brookings Tax Policy Center Microsimulation Model (version 0319-1).

Notes: Proposal would rescind the tax-free treatment of interest accruing from certain municipal bonds in 2019.
Numbers in parentheses are bootstrapped standard errors (BSE). Numbers in brackets are confidence intervals.

International
Microsimulation
Association

Table 8. Comparison of Confidence Intervals with Standard Errors Calculated from Bootstrapping and the Gatz and Smith Formula

| Proposal | Average Tax Change for All Tax Units ($) | | | | |
| | Mean | Standard Error | | Confidence Interval | |
		BSE	GSSE	Bootstrap	Normal Approximation with GSSE
Restore Personal Exemption	−758.2	1.7	1.9	[−761.6, −754.6]	[−761.9, −754.4]
Increase Standard Deduction	−521.4	1.2	1.2	[−523.8, −519.0]	[−523.8, −519.1]
Increase Tax Rate	205.7	3.9	19.3	[199.5, 214.6]	[167.9, 243.4]
Rescind Tax-free Treatment of Certain Bonds	124.9	3.9	11.2	[118.0, 133.8]	[102.9, 146.9]
Farm Loss Deduction Replaced with Credit	−0.2	0.5	2.2	[−1.3, 0.8]	[−4.5, 4.1]

Source: Urban-Brookings Tax Policy Center Microsimulation Model (version 0319-1).

Notes: The BSE column shows bootstrapped standard errors. The GSSE columns shows the Gatz and Smith formula for calculating standard errors. The Bootstrap Confidence Interval column shows the bias-corrected bootstrap confidence intervals. Finally, the Normal Approximation with GSSE Confidence Interval column shows confidence intervals derived as the means plus or minus the GSSE multiplied by 1.96. Due to rounding, that interval differs slightly from an interval calculated using the values in the table.

from $104 to $662. The asymmetry of the bootstrapped confidence interval is caused by the small number of affected records and the skewness of the underlying distribution of interest income from tax-exempt bonds. While more than 95 percent of records in this group have no tax-exempt interest income, several records have substantial amounts of tax-exempt interest income. If these records are included in the bootstrapped samples, then the resulting tax changes of the policy are extremely large and the normal approximation may not be accurate.

5.6. Discussion

Although the bootstrapped confidence intervals show that most point estimates of our analyses of the distributional impact of policy alternatives are precisely estimated, it also indicates that the model more precisely estimates some policy analyses than the others. For example, for tax units with income between $200,000 and $500,000, two policy alternatives would have almost the same magnitude of impact: rescinding tax-exempt interest would increase average taxes by $2,693 among those with tax-exempt interest income, and restoring the personal exemption by $4,000 would decrease average taxes by $2,497. However, with bootstrapped standard errors of 124 and 11, respectively, the confidence interval of the former is much wider than the latter.

Differences in the number of tax units affected in each income group only explains some of the differences in confidence intervals among the proposals. Differences in the source of income facing a change in tax treatment also plays a role, as do differences in demographic characteristics, such as the number of dependents, and interactions with other parts of the tax system, such as the Alternative Minimum Tax.

In most cases, the confidence interval constructed using the normal approximation with BSE is close to the bootstrapped confidence interval. However, it is not clear whether this finding holds if we calculate the standard error using a formula rather than a bootstrap. To explore this issue, *Table 8* shows the bootstrapped confidence intervals and the confidence intervals calculated using GSSE for each proposal's average tax change of all tax units, as well as both the BSE and the GSSE. The GSSE is comparable to BSE for proposals that affect many tax units, such as restoring the personal exemption and increasing the standard deduction. The GSSE is noticeably larger than the BSE for proposals that affect few tax units, such as increasing the top tax rate, rescinding tax-free treatment of certain bonds and replacing the farm loss deduction with a credit. In a similar fashion, confidence intervals calculated from the GSSE are comparable to the bootstrapped confidence intervals for proposals that affect many tax units and noticeably wider for proposals that affect few tax units.

6. Conclusion

The need for measuring the uncertainty of estimates from microsimulation modeling has been under-stood since its inception, yet there exists little research on the topic. In this article, we use a bootstrap technique to estimate the sampling variation from TPC's microsimulation model of the U.S. individual income tax system. The bootstrapped exercises show that point estimates produced by the model are generally very precise. This supports the standard practice of not reporting associated standard errors and confidence intervals of point estimates. However, there are exceptions, and the methodology demonstrated in this paper can be used to address uncertainties of the point estimates, with the goal of providing adequate information for better policy formation.

Calculating confidence intervals using a normal approximation works best when many tax units are affected and the point estimates are precisely estimated. When calculating confidence intervals for income categories with a small number of impacted tax units, the normal approximation may differ from the bootstrapped confidence intervals.

Sampling variation is not the only source of uncertainty in microsimulation modeling. *Cohen (1991)* discusses other sources of uncertainty, such as the use of "various control totals and regression equa-tions, from the use of imputation and statistical matching, from the use of demographic and macro-economic projections, and from the use of aging modules." Estimating the uncertainty arising from some of these sources is substantially more difficult than from sampling variation, and is a subject for future research.

ORCID iD

Robert McClelland https://orcid.org/0000-0003-1064-7177

Conflict of Interest

No competing interests reported.

Data and Code Availability

The data for this paper are based on confidential data purchased from the U.S. Internal Revenue Service. Information about purchasing the data may be found at https://www.irs.gov/statistics/soi-tax-stats-individual-public-use-microdata-files

The Stata code used for this paper is open-source and it is available at https://github.com/UrbanInstitute/confidence-intervals-tax-microsimulation

Reference

Citro CF, Hanushek EA, Citro CF, Hanushek EA. 1991. Improving information for social policy decisions: the uses of microsimulation modeling. In:*Improving Information for Social Policy Decisions -- The Uses of Microsimulation Modeling: Volume I, Review and Recommendations.* Washington, DC: The National Academies Press. DOI: https://doi.org//10.17226/1835

Cohen M. 1991. Variance estimation of microsimulation models through sample reuse. *Improving information for social policy decisions: the uses of microsimulation modeling* **2**:237–254.

Creedy, J., Kalb, G., & Kew, H. (2007). Confidence intervals for policy reforms in behavioral tax microsimulation modelling. *Bulletin of Economic Research*, **59**(1), 37-65.

Fiorio C. V. 2003. Assessing the reliability of microsimulation models using the bootstrap: an analysis of the sampling error when population is not infinite. European Economic Association and Econometric Society European Meeting, London.

Galper H, Rosenberg J, Rueben K, Toder E. 2013. Who Benefits From Tax-Exempt Bonds?: An Application of the Theory of Tax Incidence. Working Paper. Washington, DC: Urban-Brookings Tax Policy Center.

Gatz DF, Smith L. 1995. The standard error of a weighted mean concentration—I. bootstrapping vs other methods. *Atmospheric Environment* **29**:1185–1193. DOI: https://doi.org/10.1016/1352-2310(94)00210-C

Goedemé T, Van den Bosch K, Salanauskaite L, Verbist G. 2013. Testing the statistical significance of microsimulation results: a plea. *International Journal of Microsimulation* **6**:50–77.

Kolenikov S. 2010. Resampling variance estimation for complex survey data. *The Stata Journal: Promoting communications on statistics and Stata* **10**:165–199. DOI: https://doi.org/10.1177/1536867X1001000201

Orcutt GH. 1957. A new type of socio-economic system. *The Review of Economics and Statistics* **39**:116–123. DOI: https://doi.org/10.2307/1928528

Paulus A, Sutherland H, Tasseva I. 2019 EUROMOD Working Paper Series EM 3/19. Indexing out of poverty? Fiscal drag and benefit erosion in cross-national perspective. https://www.euromod.ac.uk/sites/default/files/working-papers/em3-19.pdf.

Pudney S, Sutherland H. 1994. How reliable are microsimulation results? an analysis of the role of sampling error in a UK tax-benefit model. *Journal of public economics* **53**:327–365.

Rao JNK, Wu CFJ. 1988. Resampling inference with complex survey data. *Journal of the American Statistical Association* **83**:231–241. DOI: https://doi.org/10.1080/01621459.1988.10478591