Abstract - The Suits Index is often used in tax policy analysis to measure the degree of progressivity of a tax, or to analyze changes in progressivity under alternative tax regimes. As a point estimator, however, the Index provides researchers with no assistance in assessing whether changes are in fact statistically significant. We present a bootstrap methodology by which researchers can estimate confidence intervals for differences in Suits Indices. We also illustrate the use of that methodology with an application for the U.S. income tax, simulating the effects of removing housing deductions.

INTRODUCTION

Tax policy researchers often employ the Suits Index in order to measure the degree of progressivity of a tax system. They also compute Suits Indices to analyze changes in the progressivity of a tax over time, or under alternative tax regimes. While this index is widely used and very helpful for analysis of tax regime changes, it has one fundamental limitation. The index is a point estimator and provides no confidence interval estimates for use in assessing whether changes in the index are in fact statistically significant. Given the popularity of the Suits Index and its pervasive use in policy analysis, it is important to move beyond simple use of the index as a point estimator and begin to compute confidence intervals for the index.

The Suits Index is a summary measure of the distribution of a tax. Like any average, the index theoretically has a sampling distribution. If the exact finite sampling properties of the Index were known, including its standard deviation, we could construct confidence intervals in the usual manner. In reality, however, we have no information about the distributional properties of the Suits Index (just as we have no knowledge of the distributional properties of the Gini Coefficient). Furthermore, we typically do not have sufficient data to compute a large number of Suits Indices for a particular problem and appeal to the Central Limit Theorem for a convenient application of the normal distribution. Hence, we cannot rely on the usual methods of constructing confidence intervals. The purpose of this paper is to present a methodology by which tax policy researchers can estimate confidence intervals for changes in the Suits Index. We present a step-by-step method by which confidence intervals and standard errors for differences in the index can be estimated using a bootstrap methodology.
In order to illustrate the use of that methodology, we present a policy application from the U.S. income tax. We simulate the effects of removing two housing deductions that have been the subject of recent policy debate—deductions for mortgage interest and local property taxes. Tax reform discussion since the early 1980s has questioned whether these housing deductions are justified and policy debate has included discussion of the distributional impact of removing these deductions. Several reform plans have suggested their elimination. The Armey–Shelby flat tax proposal, for example, would eliminate both deductions, while the USA Tax and the Gephardt plan would eliminate local property tax deductions (Aaron and Gale, 1996). Estimating confidence intervals for the change in the Suits Index with and without these deductions, we illustrate that their removal would significantly affect the progressivity of the income tax.

The second section of the paper presents an overview of the Suits Index and its use in tax policy analysis. The third section contains our step-by-step procedure for bootstrapping confidence intervals for the index. Our income tax policy application using this method is presented in the fourth section. The final section contains a summary and conclusions. Readers interested in the details of the bootstrap procedure will find a step-by-step description in the Appendix.

THE SUITS INDEX AND ITS USE IN TAX POLICY ANALYSIS

The Suits Index as a Measure of Tax Progressivity

The Suits Index has become one of the most widely used measures of tax progressivity since its development by D. B. Suits (1977). He devised the index by adapting the method of estimating a Gini coefficient based on the popular Lorenz Curve in the income distribution literature. For the Lorenz Curve, the cumulative percent of total family income is calculated and plotted on the vertical axis against the cumulative percent of families on the horizontal axis. The 45 degree line from bottom left corner of the diagram to top right corner represents the line of perfect equality. Along this line, the income distribution is perfectly proportionate to the population. The first quintile of the population earns exactly 20 percent of the income, the first two quintiles earn 40 percent, and so on. Of course, income distributions are typically not proportional. The first quintile of the population earns less than 20 percent of the income while the top quintile earns much more than 20 percent. Consequently, the Lorenz Curve sags below the line of perfect equality. The farther below the line of perfect equality the Lorenz Curve sags, the less equitable is the income distribution.

The Gini coefficient is computed as the ratio of the area from the Lorenz Curve sag to the 45 degree line to the total area under the 45 degree line. If there is no income inequality, the Lorenz Curve is simply the line of perfect equality, the 45 degree line, and thus the area between the Lorenz Curve and the line of perfect equality is zero making the Gini coefficient zero. At the other extreme, if there is extreme inequality with one household holding all the income, the Lorenz Curve sags all the way down to the axes and the area between the line of perfect equality and the Lorenz Curve approaches the full area under the line of perfect equality making the Gini coefficient one. Thus, the Gini coefficient ranges from zero for no income inequality to one for the most extreme inequality in which all income is concentrated in a single family.

In order to adapt this measure of income inequality to the application of estimating the progressivity of a tax system, Suits envisioned a figure similar to the Lorenz curve, but one in which the cumulative percent of tax burden is plotted on
Confidence Intervals for the Suits Index

the vertical axis against the cumulative percent of income on the horizontal axis. Figure 1 illustrates the Lorenz Curve for an income tax.

If the income tax were strictly proportional to income, then the Lorenz curve, $OCB$, would coincide with the 45 degree diagonal line $OB$—the line of perfect equality. In this tax policy application we will call this the line of proportionality. Suppose the income distribution is divided into quintiles. For a progressive tax, the percent of tax burden borne by the first quintile is less than 20 percent, and is increasing for subsequent quintiles, with the top quintile accounting for more than 20 percent of the tax. In this case, the Lorenz Curve sags below the line of proportionality. Of course, if the tax is regressive, the first quintile of the income distribution pays more than 20 percent of the tax while the top quintile pays less than 20 percent. In that case the Lorenz Curve would bow above the line of proportionality.

To construct the Suits Index, we can let $K$ denote the area below the line of proportionality, and let $L$ denote the area below the Lorenz curve. The Suits Index is defined as: $S = 1 - L/K$. For a proportional tax, $L$ approaches $K$, so the Suits Index $S$ approaches zero. Since the Lorenz Curve corresponding to a progressive tax sags below the line of proportionality, the area $L$ is smaller than $K$. As a result, the index $S$ is positive for a progressive tax. In the limiting case where the highest income household bears the entire tax burden, the Lorenz curve lies along sides $OA$ and $AB$, so $L$ equals zero and hence $S = 1$. With a regressive tax, the Lorenz curve arches above the line of proportionality making the area $L$ larger than $K$, so $S$ is negative. An index of minus one indicates that a tax system is completely regressive with the lowest income household paying all of the tax. An index value of zero identifies a tax system as proportional.

It is important to notice that the Suits Index is a measure of the average progressivity of a tax system over the full income range. Some tax systems may be progressive over one range of income and regressive over another range. This is a familiar problem that arises with any av-

---

**Figure 1.** Lorenz Curve

Cumulative proportion of total tax (%)

Cumulative proportion of total income (%)

83
verage measurement. As a summary measure of the progressivity of a tax system, the Suits Index is incapable of capturing subtleties that require information about higher moments of the tax distribution.

Mathematically, we can write the Suits Index as:

\[ S = 1 - \frac{1}{5000} \int_0^{100} T(Y) dy, \]

where \( Y \) and \( T(Y) \) are the cumulative percentage of total income and the corresponding cumulative percentage of tax burden, respectively. The \((1/5000)\) term in this expression comes from the area of the triangle below the line of proportionality whose legs are both 100 percent. In reality, however, the cumulative distribution \( T(Y) \) is often available for only a few discrete values of \( Y \).

Researchers typically compute the Suits Index by dividing taxpayers into quartiles, quintiles, or deciles. Hence they work with four, five, or ten observations. In our application, however, taxpayers are not grouped into such quantiles. Rather, we consider the cumulative share of total income received by all taxpayers in a large sample. In the income tax application we present later in this paper, for example, we explore data from an IRS sample of taxpayers, giving us \( n = 22,683 \) values of \( Y \). A discrete approximation of the Suits Index is given by the expression:

\[ S = 1 - \frac{1}{5000} \int_0^{100} T(Y) dy \approx 1 - \frac{1}{5000} \sum_{i=0}^{n-1} [T(Y_i) + T(Y_{i+1})][Y_{i+1} - Y_i], \]

where \( n \) is the total number of taxpayers in the sample.

**Applications of the Suits Index**

Numerous methods to measure tax progressivity have been developed over the years. Kiefer (1984) identifies many of these and their nuances and inconsistencies. Likewise, Greene and Balkan (1987) compare and contrast a collection of progressivity measures used in academic studies, and Seetharaman and Iyer (1995) critique the seven most widely used indices of tax progressivity, including the Suits Index. None of the progressivity measures have been proven to be superior to the others; however, the Suits Index may be the most widely used measure (Congressional Budget Office, 1988, p. 41).

Progressivity measures are useful for comparing changes in vertical equity over time or between different tax regimes. For example, Pierce (1989) measured the changes in vertical equity resulting from alternative tax preferences made available to taxpayers. She used one tax year and recalculated several alternative tax liabilities under differing assumptions and observed the changes in vertical equity. Ricketts (1990) observed changes in vertical equity over time, the 1980s, resulting from the counterbalancing effects of increased social security taxation and lower income tax rates. These types of analyses are common in the accounting and economics literature (e.g., Formby and Sykes, 1984; Dunbar and Nordhaser, 1991; Wallace et al., 1991; Stevens, 1992; Scott and Triest, 1993; Seetharaman, 1994; Seetharaman and Iyer, 1995; Shoemaker, 1995; and Ryan, 1997). The aforementioned studies used the Suits Index for all or a portion of their analyses.

While changes in the Suits Index over time or between tax regimes may indicate corresponding changes in progressivity to tax legislation, small changes may lead to false conclusions, particularly where small sample sizes are involved. Large samples are less sensitive to broad fluctuations of one, or several, data point observations than are small samples. The sensitivity of a small sample to even one outlier observation may cause a researcher to conclude that differences in comparative samples exist when they in fact do not. Such would be the case on certain classes of taxpayers where large sample sizes are not possible.
to obtain. For example, Dunbar and Nordhauser (1991) examined the regressivity of the childcare credit. Their sample was significantly reduced because only 4 to 8 percent of their sample taxpayers claimed the childcare credit. Stevens (1992) narrowed his sample to include only Nebraska taxpayers to measure the progressivity of several tax expenditures.

Furthermore, researchers are inclined to report changes in vertical equity even if minute, in the absence of a test for significance. However, small changes in progressivity, as measured over two time periods or between tax regimes, may be caused by unobserved confounding variables. Shoemaker (1995) computed a change in the Suits Index from 0.208 to 0.232 from one year to the next and concluded that taxation of the elderly becomes more progressive after turning age 65. Dunbar and Nordhauser (1991) use a difference in Suits Indices of 0.1391 and 0.1377 to justify the progressivity of the childcare credit. Ryan (1997) draws a similar conclusion for a change from –0.35 to –0.37 over a two–year period, and Rick- ets (1990) does similarly for a change from 0.213 to 0.268 over a four–year period. Perhaps these changes are significant; perhaps they are not. Unless a test of significance for the Suits Index is used, speculation on significance may lead to wrong conclusions.

A METHOD FOR ESTIMATING SUITS INDEX CONFIDENCE INTERVALS

Bootstrap Methodology

In this section we describe the methodology of computing bootstrap–t–intervals in order to test the significance of a change in the degree of progressivity between two tax system regimes. A detailed step–by–

step description of a bootstrapping procedure for constructing confidence intervals is provided in the Appendix. Bootstrapping is a computer–intensive resampling method that can provide standard error estimates and confidence intervals for parameters of interest without a parametric specification. The bootstrap method depends on the notion of bootstrap samples and uses all possible sample information for the estimation of standard errors and confidence intervals. There is no standard statistical package for the estimation of bootstrap–t–intervals, hence one must program this task and we did so using the FORTRAN language.1 Bootstrap confidence intervals for the Suits Index can also be programmed using GAUSS or other common mathematical programs. In fact, any program that is written in FORTRAN, C, C++, Java, or VB can also be written in GAUSS. In addition, Gauss has a feature called the “Foreign Language Interface” (FLI) that allows users to create functions written in C, FORTRAN, or other languages, and call them from a GAUSS program. Therefore one could use our FORTRAN program, but run it using GAUSS. In the case of a small sample there are other options such as using SAS.

Using the typical econometric approach we work within the world of classical statistics where experiments can be repeated indefinitely. Using that framework, we attempt to discover the underlying data generating process, representing that process parametrically. When we do not know the finite sample properties of the statistic we are attempting to estimate we rely upon the implications of the Central Limit Theorem and use the normal distribution. In this context, inference from the sample data is viewed in terms of discovering the probability distribution of the

---

1 Researchers interested in obtaining a copy of our computer program used in bootstrap estimation of Suits Index confidence intervals may write to request a copy. The program will be made available at no cost with the understanding that researchers using the program will appropriately cite the source.
parameter estimates. In reality, however, we are often prevented from doing this analytically and we resort to asymptotic distribution theory. Bootstrap methodology offers an alternative, based on resampling. The idea is to use the sample data at hand to empirically determine the distribution. The key concept is that resampling, with replacement, provides information on the stochastic nature of the data generating process. Using this method we obtain an empirical distribution function for the Suits index and use that distribution to test hypotheses.

In order to conduct a significance test, the following null and alternative hypotheses are specified. The null hypothesis is specified as where the difference in Suits Indices for two regimes is computed: \( H_0: S_d \leq 0 \), where \( S_1 \) and \( S_2 \) are population values of Suits Index computed for tax regimes one and two, respectively, and \( S_d = S_2 - S_1 \). This null hypothesis states that the difference between the two indices is non-positive; implying that tax regime two is no more progressive than tax regime one. The alternative hypothesis is specified as \( H_a: S_d > 0 \), stating that tax regime two is more progressive than tax regime one. We estimate the bootstrap confidence interval of the difference between the two Suits Indices: \( S_d \). If the confidence interval for the difference contains only positive values we can reject the null hypothesis and conclude that tax regime two is more progressive than tax regime one.

**APPLICATION: ASSESSING THE IMPACT OF ELIMINATING HOUSING DEDUCTIONS**

This section illustrates an application of computing the confidence interval for the change in Suits Index under two tax regimes. We examine the vertical equity effect of removing housing tax deductions. In particular, we simulate removal of both mortgage interest and property tax deductions on the U.S. personal income tax. Using a cross-section data set of individual income tax returns from the IRS for 1990 we assess both the revenue-neutral and revenue non-neutral vertical equity effects of eliminating these housing tax deductions. The change in Suits Index between the existing federal income tax that allows full nominal mortgage interest and property tax deductions, and an alternative income tax regime that disallows deductions for nominal mortgage interest and property tax payments is used to gain insight on the vertical equity impact of removing these deductions.

The estimated Suits Index for the current federal income tax, \( \hat{S}_1 \), and the estimated Suits Index for a tax regime that excludes the tax deductibility of mortgage interest and property tax payments, \( \hat{S}_2 \), are computed. If \( \hat{S}_2 \) is strictly greater than \( \hat{S}_1 \) the removal of housing tax deductions increases the progressivity of the federal income tax system.

Our data is taken from the Internal Revenue Service (IRS) Statistics of Income panel of tax returns. The Statistics of Income (SOI) division of the IRS draws a random sample of tax returns each year. These samples of approximately 100,000 returns are then compiled by the Office of Tax Policy Research (OTPR) at the University of Michigan into Individual Tax Files. These files are then used by the OTPR to construct a panel data set, tracking taxpayers over time. All taxpayers whose last four Social Security Number digits matched randomly selected numbers were included in the 1990 Panel File produced by the OTPR. We used the OTPR 1990 Panel File that contained returns for \( n = 22,683 \) taxpayers. Based on the sample taxpayer data, the estimated Suits Index increases from \( \hat{S}_1 = 0.1577 \) with full deductibility of mortgage interest and property taxes to \( \hat{S}_2 = 0.2337 \) with no deductibility. This yields an estimated change in vertical equity of \( \hat{S}_d = \hat{S}_2 - \hat{S}_1 = 0.076 \). The revenue-neutral Suits Index increases from with full deductibility to \( \hat{S}_2 \).
Confidence Intervals for the Suits Index

87

= 0.2379 with zero deductibility, yielding an estimated change in vertical equity of $S_d = S_2 - S_1 = 0.08$. In this case, we scaled existing marginal tax rates back proportionately in order to maintain the same revenue yield on the income tax. Results for both revenue-neutral and revenue non-neutral cases suggest that eliminating housing tax deductions will increase the progressivity of the income tax.

We compute bootstrap–t confidence intervals following the procedure outlined above for the population value of the difference between Suits Indices, $S_d$, for a hypothetical tax regime that excludes mortgage interest and property tax deductions and the existing federal tax regime including both housing tax deductions. Table 1 reports the bootstrap–t values. Using those $t$ values, the $100(1 - 2\alpha)$ percent bootstrap–t confidence intervals are formed. Table 2 presents both 90 percent and 95 percent bootstrap–t confidence intervals for $S_d$. Those bootstrap–t confidence intervals show that $S_d$ is positive at both levels of significance. Hence, we can reject the null hypothesis of a non–positive difference in Suits Indices. We are left with the alternative that the difference is strictly positive. Thus, we find a statistically discernible increase in progressivity from eliminating mortgage interest and property tax deductions in the U.S. income tax. This example illustrates the advantage of having an interval estimate upon which statements can be made with a specified level of confidence.

### TABLE 1
BOOTSTRAP t TABLE

<table>
<thead>
<tr>
<th>2.5%</th>
<th>5%</th>
<th>10%</th>
<th>16%</th>
<th>50%</th>
<th>84%</th>
<th>90%</th>
<th>95%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.41</td>
<td>-2.81</td>
<td>-2.24</td>
<td>-2.25</td>
<td>-0.73</td>
<td>0.26</td>
<td>0.54</td>
<td>0.82</td>
<td>1.07</td>
</tr>
</tbody>
</table>

### TABLE 2
BOOTSTRAP t CONFIDENCE INTERVALS FOR $S_d$

<table>
<thead>
<tr>
<th>90% Lower</th>
<th>90% Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.071</td>
<td>0.093</td>
</tr>
<tr>
<td>95% Lower</td>
<td>95% Upper</td>
</tr>
<tr>
<td>0.069</td>
<td>0.097</td>
</tr>
</tbody>
</table>

### SUMMARY AND CONCLUSIONS

Researchers have commonly used the Suits Index to provide insight on the degree of progressivity of a tax system. The index has also been used, however, to measure changes in the degree of progressivity of a tax system, despite the fact that there has been no means of determining whether the observed change in the index is discernible in a statistical sense. This paper has remedied that situation, providing a means by which confidence intervals can be constructed for the Suits Index using the bootstrap method. This methodology for computing confidence intervals, which we detail in the Appendix, allows researchers to state with greater certainty whether proposed tax policy changes will have a significant impact on progressivity.

### REFERENCES


Kiefer, Donald.  

Pierce, Bethane J.  

Rao, Jon N. K., and Chien–Fu J. Wu.  

Ricketts, Robert C.  

Ryan, David.  


Seetharaman, Ananth.  

Seetharaman, Ananth, and Govind S. Iyer.  

Shoemaker, Paul.  

Stevens, Kevin T.  

Suits, Daniel B.  

U.S. Congress, Congressional Budget Office.  

Wallace, Sally, Michael Wasylenko, and David Weiner.  

**APPENDIX: COMPUTING THE BOOTSTRAP–t INTERVAL**

The bootstrap–t interval procedure estimates the distribution of the bootstrap–t statistic directly from the data. In other words, this method computes the bootstrap–t statistic in order to build a student–t table that can be used to construct a confidence interval around the point estimate provided by the Suits Index. The bootstrap–t table constructed from the data using this procedure is only appropriate for the data at hand, however. The algorithm for computing the bootstrap–t confidence interval for $\tilde{S}_j$ follows Efron and Tibshirani (1993) and is summarized in the following steps:

1. $\tilde{S}_1$ and $\hat{S}_j$ are estimates of Suits Indices $S_1$ and $S_j$ computed from samples $X_1$ and $X_j$, respectively. The samples $X_1$ and $X_j$ consist of $n$ cross–section observations on income and tax liability drawn from the population of tax returns filed under tax regimes one and two, respectively. Note that each sample contains observations on two variables: income and tax liability. Hence, $X_1$ and $X_j$ are both matrices of dimension $nx2$.

2. $N$ independent bootstrap random samples are generated with replacement from the sample $X_1$ and are denoted by the $nx2N$ matrix, $x_1^* = [x_{11}^*, x_{12}^*, x_{13}^*, \ldots, x_{1n}^*]_{n\times2N}$, where
Confidence Intervals for the Suits Index

3. Each random sample in both matrices, \( x^*_i \), represents the \( i \)-th bootstrap random sample drawn with replacement from \( X_i \) and is denoted by the \( nx2 \) matrix \( x^*_{ij} = [x_{1,ij}, x_{2,ij}] \) for all \( i = 1, \ldots, N \). In this matrix, \( x_{1,ij} \) is the first column vector and \( x_{2,ij} \) is the second column vector, each containing \( n \) observations on income and tax liability. Similarly, \( N \) independent bootstrap random samples are drawn with replacement from the sample \( X_s \) and are denoted as the \( nx2N \) matrix, \( x^*_s = [x^*_{11}, x^*_{12}, x^*_{13}, \ldots, x^*_{1N}] \), where \( x^*_s \) represents the \( i \)-th bootstrap random sample drawn with replacement from sample \( X^*_s \). We can denote the resulting \( nx2 \) matrix as \( x^*_{2j} = [x^*_{2j,1}, x^*_{2j,2}] \) for all \( i = 1, \ldots, N \), where \( x^*_{2j,1} \) is the first column vector and \( x^*_{2j,2} \) is the second column vector containing \( n \) observations on income and tax liability, respectively. Since each random sample in \( x^*_1 \) and \( x^*_2 \) is selected with replacement, some might appear twice, and some might appear one time, and some might appear twice in each of bootstrap samples.

Each random sample in both matrices, \( x^*_i \) and \( x^*_2 \) is sorted by income. Sorted bootstrap samples are then used for the estimation of bootstrap replications of Suits Indices. Corresponding to each bootstrap sample in matrix, \( x^*_s \), a bootstrap replication of the Suits Index is calculated. \( N \) bootstrap replications of Suits Indices are represented by an \( Nx1 \) vector, \( s^*_i = [s^*_{1i}, s^*_{22}, s^*_{21}, \ldots, s^*_{2N}] \), where \( s^*_i = s^*_i(x^*_i) \) for all \( i = 1, \ldots, N \). Similarly, corresponding to each bootstrap sample in vector \( x^*_s \), a bootstrap replication of the Suits Index is computed. \( N \) bootstrap replications of Suits Indices are defined by the \( Nx1 \) vector, \( s^*_s = [s^*_{1s1}, s^*_{2s2}, s^*_{2s1}, \ldots, s^*_{2sN}] \), where \( s^*_s = s^*_s(x^*_s) \) for all \( i = 1, \ldots, N \).

4. The difference between the two vectors, \( s^*_i \) and \( s^*_2 \), is computed and is denoted by the vector, \( s^*_{ij} = [s^*_{1ij} - s^*_{2ij}, (s^*_{12} - s^*_{22}), (s^*_{13} - s^*_{23}), \ldots, (s^*_{1N} - s^*_{2N})] \), where \( s^*_{ij} = s^*_{ij}(x^*_{ij}) \) for all \( i = 1, \ldots, N \). The computation of the bootstrap-\( t \) statistic requires the standard error of each bootstrap replication of Suits Indices. Since the distribution of the Suits Index is unknown, there is no statistical formula for estimating standard errors of the differences between the Suits Indices in vector \( s^* \). Thus, the next step in the process is to calculate the bootstrap estimate of standard error for each element in vector \( s^* \). The bootstrap algorithm for estimating standard errors is as follows.

4a. \( m \) bootstrap random samples (with \( m < N \)) are generated with replacement, corresponding to each random sample in vector \( x^*_s \) and is denoted by the \( nx2mn \) matrix:

\[ Y^*_s = [y^*_{s1}, y^*_{s2}, y^*_{s3}, \ldots, y^*_{sN}] \]

where each \( y^*_{sj} \) is a matrix of dimension \( nx2n \) and contains \( m \) random samples of size \( n \) and can be written as \( y^*_{sj} = [y^*_{s1j}, y^*_{s2j}, y^*_{s3j}, \ldots, y^*_{sNj}] \) for all \( i = 1, \ldots, N \). Each \( y^*_{sj} \) is the \( j \)-th random sample drawn with replacement from the \( j \)-th sample in matrix, \( x^*_s \), and is denoted by the \( nx2 \) matrix:

\[ y^*_{sj} = [y^*_{sj1}, y^*_{sj2}, y^*_{sj3}, \ldots, y^*_{sjN}] \]

for all \( j = 1, \ldots, m \), where \( y_{sj1} \) is the first column vector and \( y_{sj2} \) is the second column vector containing \( n \) observations on income and tax liability, respectively. Similarly, \( m \) (where \( m < N \)) bootstrap random samples are also generated with replacement corresponding to each random sample in vector \( x^*_s \) and is also denoted by the \( nx2mn \) vector:

\[ Y^*_s = [y^*_{s1}, y^*_{s2}, y^*_{s3}, \ldots, y^*_{sN}] \]

where each \( y^*_{sj} \) is a matrix of bootstrap random samples \( nx2m \) drawn with replacement from sample and can be denoted as for all \( i = 1, \ldots, N \), where \( y^*_{sj} \) is the \( j \)-th random sample of size \( n \) drawn from the \( j \)-th sample in matrix \( x^*_s \) and is denoted by the \( nx2 \) matrix:

\[ y^*_{sj} = [y_{sj1}, y_{sj2}] \]

for all \( j = 1, \ldots, m \) and \( i = 1, \ldots, N \), where \( y_{sj1} \) is the first column vector and \( y_{sj2} \) is the second column vector containing \( n \) observations on income and tax liability, respectively.

4b. For the purpose of computing bootstrap replications of Suits Indices from \( mN \) random samples contained in matrices, \( Y^*_s \) and \( Y^*_s \), the samples are sorted by income. Then a bootstrap replication of the Suits Index is calculated corresponding to each sample in
the $i$th matrix, $y_{i1}^{*}$, and denoted by the $mx1$ vector, $s_{i}^{*} = \{s_{2i1}^{*}, s_{2i2}^{*}, s_{2i3}^{*}, \ldots, s_{2im}^{*}\}'$, for all $i = 1, \ldots, N$. Similarly, a bootstrap replication of the Suits Index is calculated corresponding to each sample in the $i$th matrix, $y_{i1}^{*}$, and denoted by the $mx1$ vector, $s_{i}^{*} = \{s_{2i1}^{*}, s_{2i2}^{*}, s_{2i3}^{*}, \ldots, s_{2im}^{*}\}'$, for all $i = 1, \ldots, N$.

4c. The difference between the vectors $s_{ii}^{*}$ and $s_{i}^{*}$ is estimated for all $i = 1, \ldots, N$ and denoted by the vector, $s_{ii}^{*} = \{s_{2i1}^{*} - s_{2i1}^{*}, \ldots, (s_{2im}^{*} - s_{2im}^{*})\}'$, where, $s_{2im}^{*} = s_{2im}^{*}(y_{i1,i}^{*} - y_{2i}^{*})$ for all $i = 1, \ldots, N$ and $j = 1, \ldots, m$. There are $N$ $s_{ii}^{*}$ vectors.

4d. The bootstrap standard errors of each $s_{ii}(i = 1, \ldots, N)$ are calculated by using the formula $\hat{\sigma}_{s_{i}} = \sqrt{\sum_{i=1}^{N}(s_{ii}^{*} - s_{ii}^{*})^{2}/(m-1)}$ for all $i = 1, \ldots, N$, where $s_{ii}^{*} = \sum_{i=1}^{N} s_{ii}^{*}/m$.

5. The bootstrap–t statistic is then computed for each $s_{ii}^{*}$ and is defined by $t_{i}^{*} = s_{ii}^{*} - \hat{\bar{S}}_{j}/\hat{\sigma}_{j}$, where $\bar{S}_{j}$ is a function of $X_{i}$ and $X_{j}$, and is the estimate of the population value $S_{j}$, and $s_{ii}^{*}$ is the $i$th bootstrap replication computed from the bootstrap samples $X_{i}^{*}$ and $X_{j}^{*}$. Thus, $N$ bootstrap–t statistics are computed using the above formula. The bootstrap–t table is then constructed from the ordered values of $t_{i}^{*}$. For example, for $N = 1,000$ the estimates of the 5 percent and 95 percent points are the 50th and 950th largest values of $t_{i}^{*}$, respectively.

6. The estimate of the confidence interval for $\bar{S}_{j}$ requires the standard error of its estimate, $\hat{\sigma}_{j}$. The bootstrap standard error, $\hat{\sigma}_{j}$, of $\bar{S}_{j}$ is estimated utilizing the bootstrap method for calculating standard error.

7. Finally, the 100$(1 - 2\alpha)$ percent confidence interval for $\bar{S}_{j}$ is given by the expression $[\bar{S}_{j} - t_{1-\alpha}^{*}\hat{\sigma}_{j}, \bar{S}_{j} + t_{1-\alpha}^{*}\hat{\sigma}_{j}]$. If $N\alpha$ is not an integer, the following procedure can be used. Assuming $\alpha \leq 0.5$, let $k = \lceil (N + 1)\alpha \rceil$, the largest integer $\leq (N + 1)\alpha$. Then the $\alpha$ and $(1 - \alpha)$ quantiles are defined by the $k$th largest and $(N + 1 - k)$th largest values of $(t^{*})$. The number of bootstrap samples, $N$, used in our income tax application is $N = 1,000$. Generally, the estimation of the bootstrap–t interval requires $N$ to be at least 1,000.

Standard error estimation requires the number of bootstrap samples to be in the range of 25 to 100. Here, we use 50 bootstrap samples for estimating the bootstrap standard errors. The student–t statistic is called an approximate pivot. Approximate pivot means that the distribution of the student–t statistic is approximately the same for each population value of the parameter of interest. In fact, this property allows the construction of bootstrap–t intervals from the distribution of $(t^{*})$.

This method of computing bootstrap–t intervals is applicable when you have an independent and identically distributed (iid) sample of fixed size $n$ from an unknown distribution. Our application, however, uses an Internal Revenue Service (IRS) tax data file that is a stratified random sample. Consequently, we use the Rao and Wu (1988) method applicable for stratified random samples. This method requires resampling from each stratum and then scaling each observation in each stratum prior to the calculation of bootstrap intervals. For example, assume that the IRS tax data file contains $n$ strata and let $X_{1h}$ and $X_{2h}$ denote the $h$th stratum in samples $X_{1}$ and $X_{2}$, respectively. $X_{1h}$ and $X_{2h}$ are random samples drawn from $X_{1}$ and $X_{2}$, respectively. The Rao and Wu method is described in the following three steps.

1. For all $h$ strata in sample $X_{1}$, a simple random sample $X_{1h}^{*}$ of size $m_{h}(\geq 1)$ is drawn with replacement from $X_{1h}$ of size $n_{h}$. Similarly, for all $h$ strata in sample $X_{2}$, a simple random sample $X_{2h}^{*}$ of size $m_{h}(\geq 1)$ is drawn with replacement from $X_{2h}$ of size $n_{h} \cdot m_{h}$ is required to be less than $n_{h}$.

2. Each observation $X_{1h}^{*}$ and $X_{2h}^{*}$ are scaled to obtain $\hat{X}_{1h}^{*}$ and $\hat{X}_{2h}^{*}$, respectively, using the following formulae: $\hat{X}_{1h}^{*} = \frac{X_{1h}^{*} - \bar{X}_{1h}}{s_{X_{1h}}} + m_{h}^{1/2}(n_{h} - 1)^{-1/2}$ and $\hat{X}_{2h}^{*} = \frac{X_{2h}^{*} - \bar{X}_{2h}}{s_{X_{2h}}} + m_{h}^{1/2}(n_{h} - 1)^{-1/2}(X_{1h}^{*} - \bar{X}_{1h})$.

3. Each observation, $X_{1h}^{*}$ in the $h$th stratum, $X_{1h}^{*}$ is replaced by $\hat{X}_{1h}^{*}$. Similarly, each observation, $X_{2h}^{*}$ in the $h$th stratum, $X_{2h}^{*}$ is replaced by $\hat{X}_{2h}^{*}$.