
Jordi ARCARONS
Samuel CALONGE*

Department of Statistics, Econometrics and Spanish Economy
University of Barcelona

ABSTRACT

The Suits index is often used in tax policy analysis to measure progressivity and to detect changes in progressivity over time and between different fiscal policies. It is surprising however that little attention has been given to inference issues. In this paper, the limiting distribution of the Suits index estimator is deduced and easily computed plug-in formulae for the index estimator and its sampling variance are provided. By means of a simulation analysis, we prove that inference based on first-order asymptotics performs well for moderately large samples. Bootstrap-t -which uses the plug-in variance estimator- appears to be favourable. The accuracy of the proposed inference tests is also illustrated using real data from the Spanish Income Tax.

KEYWORDS: Suits index, Asymptotic inference, U-statistics, Bootstrap-t, Plug-in variance estimator.

Words: 7573.

*Corresponding author scalonge@ub.edu.
1. INTRODUCTION

A relevant aspect concerning fiscal policies is related to the effects of taxes and benefits on income distribution. From an empirical point of view, the computation of indices such as Kakwani, Suits and Reynolds-Smolensky gives rise to an aggregate picture on tax-progressivity and income-redistribution patterns which is of primary interest, not only for applied welfare analysis, but also for policy makers to acquire knowledge about the policies therein embedded. The use of microdata –usually household surveys or samples drawn from administrative data– is the source of information for estimating income and tax distributions and consequently, statistical inference has to be considered even in case of large samples being available (Maasoumi 2000, Bishop et al. 1998).

A number of authors have focused on deriving the statistical properties of economic indices, with the Gini index of income inequality as the most paradigmatic case (i.e., Glasser 1962, Gastwirth 1974, Cowell 1989). Other authors have emphasized the point of obtaining a simple and reliable variance formula for the Gini index estimator to undertake inference, see Glasser (1962), Dadvison (2009), or Schechtman (1999) and Karagiannis and Kovacevic (2000) who used Jackknife as a shortcut to obtain a variance formula without the need to resample.

In the case of tax-progressivity and income-redistribution indices, there have been few references so far. Specifically, Bishop et al. (1998) derived the asymptotic distribution of the Kakwani and Reynolds-Smolensky index estimators and proposed variance estimators which have a $U$-statistics structure. These estimators were applied to test changes in liability and residual progression (using Musgrave’s terminology) throughout the fiscal reforms that several Western countries underwent in the 1980 decade. However, these variance estimators have cumbersome analytical expressions (see formulae B.1 to B.8 in the quoted paper).
The Suits index has also been profusely applied to tax and public expenditure analysis, although its use goes beyond this literature. Generally speaking, it measures disproportionality between two economic variables by means of the *relative concentration curve* (RCC), which is essentially a *Lorenz*-type curve. In this paper we shall use the index to measure progressivity, that is, the degree of disproportionality between the pre-tax income and tax distributions. As recipients' post-tax income is used in its computation, instead of taxes, it is also interpreted as an income-redistribution index, Pfähler (1983).

There is no literature providing an explicit formula for its variance estimator which allows for setting confidence intervals and hypothesis tests in a tractable way. Actually, inference on the Suits index has been driven by using resampling methods. The most interesting reference is Anderson *et al.* (2003) who applied what they called a ‘step-by-step’ bootstrap- procedure to set confidence intervals. They used a very large sample of taxpayers drawn from the 1990 US Income-Tax Returns file for testing the effect of removing housing deductions on tax-progressivity. As these authors remarked, relying on bootstrapping was necessary as there was no information about the distributional properties of the index.

In this paper, the asymptotic distribution of the Suits index is deduced by applying non-degenerate *U*-statistic theory. We provide simple analytical formulae for the index estimator and its sampling variance. The expressions obtained are based on the *plug-in* estimator of the first component of Hoeffding’s Decomposition (HD), from which an approximate result for the asymptotic variance of *U*-statistics can be derived.

Our formulae make confidence intervals and hypothesis tests easy to set up. Furthermore, the variance formula can be applied to the bootstrap-*t* method, thus avoiding the double-nested resampling needed to estimate the standard error of the pivotal statistic.

---

1 Pfähler (1985) is a useful reference for the analysis of RCC. In this paper, the Suits index is dated back to Hainsworth (1964), who used the index to measure (im)balance of trade analysis across countries, and to productivity and factor properties analysis across countries.
The next step of the analysis consists in comparing the finite sampling behaviour of the asymptotic indices with bootstrap inference through a simulation analysis and in an empirical application. The simple asymptotic method performs reasonably well for samples which are moderately large, although the bootstrap-\(t\) method combined with the asymptotic variance estimator to form the \(t\)-statistic is found to have a superior performance.

Finally, our results are applied to evaluate changes in tax-progressivity and income-redistribution by using real data drawn from the Spanish Income Tax. First, the effect of inflation on tax-progressivity is analyzed (fiscal drag). Hypothesis tests provided strong evidence of a significant reduction on tax-progressivity, and this effect is probably due to the tax burden for low-income taxpayers increasing by more than the high-income taxpayers.

Second, a sensitivity analysis is performed to find out whether inference tests are able to detect slight differences on tax-progressivity among three different neutral-revenue tax cuts simulated from the Spanish Income Tax. As is common in this research area, the data used are weighted samples, where weights are the inverse of the inclusion probabilities. Our theoretical results assumed i.i.d. random samples for the sake of simplicity, however the assumption of i.n.i.d. distributed samples can be easily accommodated in the analysis by constructing weighted consistent estimators of the different components of the empirical estimators.

This paper is organized as follows: Section 2 offers a mathematical description of the Suits index and its properties and introduces the index as a function of regular functionals. Further, consistent plug-in estimates for the index and its sampling variance are presented together. The detailed process of deriving the asymptotic distribution of the Suits index estimator is explained with the demonstration included in the Appendix, in which the formula of the asymptotic variance is included. Section 3 illustrates the Monte Carlo exercise performed for the comparison of the asymptotic and bootstrap confidence intervals. In Section 4, inference
methods are used to investigate changes in tax-progressivity and redistribution by using the two empirical applications mentioned before. Section 5 provides brief concluding remarks.

2. STATISTICAL PROPERTIES OF THE SUITS INDEX.

2.1 Preliminaries.

Let \( X \) represent the pre-tax income variable and \( Y \) denote a general variable that in some cases will represent tax or benefit (\( T \)) and in others, the post-tax income, defined as \( Y = X - T \). We assume that \( X \) and \( Y \) are non-negative random variables.

The Suits index, which is interpreted as an average measure of tax-progressivity, is based on the RCC of the \( Y \)-variable, which plots its cumulative percentage against the cumulative percentage of pre-tax income \( X \) when both variables have been ordered in ascending order of \( X \). As an example, the curve OCD in Figure (1) represents the RCC of the 2006 Spanish Income Tax. We shall denote the RRC of \( Y \) as \( L_Y(q) \), where \( q \) is the value of the Lorenz curve associated to the population rank \( p \in (0,1) \).

Geometrically, the Suits index is calculated as \( S = 1 - 2A \) or, equivalently, \( S = 2\overline{A} - 1 \), where \( A \) is the area compressed between the RCC and the \( x \)-axis. For convenience, let us consider the latter expression. By integrating the RCC over the \( y \)-axis, we have:

\[
S = 2 \int_0^1 q dL_Y(q) - 1
\]

(1)

As the RCC evaluated at \( q \) is equivalent to the TCC at the point \( p, L_Y(p) \), the Suits index is rewritten as:

\[ 2 \]

Figure (1) also overlays the Lorenz curve (LC) and the tax-concentration curve (TCC), representing the cumulative percentages of pre-tax income and taxes in the \( y \)-axis against the cumulative percentage \( p \) of the population in the \( x \)-axis (curves OBD and OGD respectively).
\[ S = 2 \int_{0}^{1} q dL_Y(p) - 1 \]  

Considering that \( q = \frac{1}{\mu_x} \int_{0}^{p} x(t) dt, \ L_Y(p) = \frac{1}{\mu_y} \int_{0}^{p} y(t) dt, \) where \( x(p), \ y(p) \) are the \( p \)th quantiles of the pre-tax income and tax distributions respectively, and differentiating the TCC, \( dL_Y(p) = \frac{y}{\mu_y} dp, \) the following mathematical expression for the index is obtained:

\[ S = \frac{2}{\mu_x \mu_y} \int_{0}^{1} \left( \int_{0}^{p} x(t) dt \right) y dp - 1 \]  

or alternatively

\[ S = \frac{2}{\mu_x \mu_y} \int_{0}^{1} x(p) \int_{0}^{p} xdp \ y dp - 1 \]  

If the \( Y \) variable is proportionally distributed, then RCC coincides with the \( 45^0 \)-line, the area \( \bar{A} \) is equal to \( \frac{1}{2} \), and consequently the Suits index equals zero. The index ranges from -1 to 1, a nice property not shared by the Kakwani index, indicating respectively the cases of lowest tax-regressivity, in which the tax burden is totally supported by the poorest individual, and highest tax-progressivity, in which the tax burden is concentrated in the richest individual (a detailed discussion of the index and its properties can be found in Lambert 2001).

### 2.2 The Suits index as a function of regular functionals.

Let \((X_i, Y_i)\) be i.i.d. pairs of continuous bivariate random variables \((1 \leq i \leq n)\) with distribution function (d.f.) \( F(x, y) \) and continuous marginal d.f. \( F(x) \) and \( F(y), \) respectively. It is assumed that first and second moments exist.

Consider the following regular functional \( \theta(F) \) of the population d.f. \( F(x, y): \)

\[ \theta(F) = \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} y_1 x_2 I(x_2 < x_1) dF(x_1, y_1) dF(x_2) \]  

5
where \( h(x_1,y_1;x_2,y_2) = y_1x_2I(x_2 < x_1) \) is the kernel of the functional \( \theta(F) \) and \( I(.) \) is the usual indicator function.\(^3\) We shall define the population Suits index as

\[
S = \frac{2\theta}{\mu_x \mu_y} - 1
\]

with \( \mu_x > 0, \mu_y > 0 \). In the same way as the Gini and Kakwani indices, the Suits index also has a ‘covariance’ form which has been traditionally used for estimation purposes. However, formula (6) is easier to estimate once the plug-in estimator of \( \theta \) is defined (the covariance formula of the index is derived from equation (6) in appendix A.1).

### 2.3 The Suits index estimator.

Let \((X_1,Y_1), \ldots,(X_n,Y_n)\) be a i.i.d. random sample drawn from the d.f. \( F(x,y) \). An estimator of the index is obtained after replacing \( \theta \) by its U-statistic and \( \mu_x, \mu_y \) by their sample statistics \( \bar{x}, \bar{y} \) in equation (6) and its consistency follows from Slutsky’s theorem. By averaging the symmetric kernel of the functional \( \theta \) over all pairs of the sample units, we obtain the (unbiased) U-statistic:

\[
\hat{\theta} = \frac{1}{2n(n-1)} \sum_i \sum_j y_ix_jI(x_j < x_i) + x_iy_jI(x_j > x_i)
\]

Therefore, the Suits index estimator is a function of three U-statistics, \( \hat{\theta} \), with kernel of degree two and the means \( \bar{x}, \bar{y} \) with kernels of degree one. References to U-statistics can be found, for example, in Hoeffding (1948), Lee (1990), Kowalski and Tu (2008).

We shall use the plug-in estimator of the functional \( \theta \) instead of (7). Let us first consider the following conditional expectation \( \theta_c(x_i,y_i) = E[h(.) \mid x_i,y_i] \) where \( h(.) \) is the kernel of the

---

\(^3\) The kernel of a regular functional is a real-value function which is defined as the unbiased estimator with the smallest sample \( m<n \), being \( m \) the degree of the functional. In this case, \( m=2 \).
functional $\theta(F)$, the pair $(x_i, y_i)$ is fixed and the expected values are taken with respect to the random variable $X_2$ such that,

$$
\theta_i(x_i, y_i) = E[y_i, X_2|X_2 < x_i|x_i, y_i] \\
= y_i E[X_2|X_2 < x_i]
$$

The functional $\theta$ is defined then by iterated expectations as $\theta = E[\theta_i(x_i, y_i)]$. This random function $\theta_i(x_i, y_i)$, after imposing symmetry, gives rise to the first component of HD of the $U$-statistic $\hat{\theta}$. Notice that the $U$-statistic is a sequence of random variables which are not independent, whereas the random functions $\theta(x_i, y_i)$ are i.i.d. and their variance exists.

As the conditional expectation above is defined on the fixed pair $(x_i, y_i)$, its plug-in estimator excludes one observation and it is computed as $\hat{\theta}_i = \frac{1}{n-1} \sum X I(X < x_i)$ which, after symmetrising, leads to:

$$
\hat{\theta}_i = \frac{1}{2(n-1)} \left( \sum y_i X I(X < x_i) + x_i \left( \sum Y - \sum Y I(X \leq x_i) \right) \right)
$$

We can write (9) as

$$
\hat{\theta}_i = \frac{n}{2(n-1)} \left( \frac{1}{n} \sum X I(X \leq x_i) - \frac{1}{n} y_i x_i + x_i \mu - x_i \frac{1}{n} \sum Y I(X \leq x_i) \right)
$$

which is the consistent estimator of $\theta_i(x_i, y_i)$ where $\frac{n}{n-1} \to 1$ and $\frac{1}{n} y_i x_i \to 0$ for $n$ large.

Finally, $\hat{\theta} = n^{-1} \sum \hat{\theta}_i$ and the Suits index estimator is computed as

$$
\hat{S} = \frac{2}{\bar{x} \bar{y}} - 1
$$

Equation (10) can be rewritten, after deleting the terms which are asymptotically negligible, as $\hat{\theta}_i = \frac{1}{2} \left( y_i \hat{F}_n(x_i) \bar{x}_i + x_i \mu - x_i \hat{F}_n(x_i) \bar{y}_i \right)$ where $\hat{F}_n(x_i)$ is the empirical d.f. and $\bar{x}_i, \bar{y}_i$ are the natural estimates of $E(X|X \leq x_i)$ and $E(Y|X \leq x_i)$. These conditional means can be recursively
estimated as $x_i = x_{i-1}(1 - f_i) + x_i f_i$ with $f_i = w_i / \sum_{j=1}^n w_j$ and $w_i$ is the weight attached to each data point (which equals 1 for i.i.d. random samples).

To establish the asymptotic distribution of the Suits index estimator, we notice first that the distribution of the $U$-statistic $\hat{\theta}$ is approximated by a sum of independent random functions $\sqrt{n} (\hat{\theta} - \theta) \approx \left( \frac{1}{\sqrt{n}} \sum_{i=1}^m \theta_i (x_i, y_i) - \theta \right)$ (i.e., Lee (1990), Theorem 1, page 76) where $m=2$, in this particular case. As sample means converge to the population means, applying the multivariate CLT and the Delta method leads to the asymptotic distribution of the Suits index estimator (see Lemma 1 and Proposition 1 in Appendix 2.)

To make inferences it is necessary to estimate the sampling variance of $\hat{S}$. The asymptotic variance is presented in equation (A2.1) in Appendix 2. By substituting functionals in this equation by their corresponding plug-in estimators, we get a consistent estimator of the variance. Despite the fact that at a first glance the formula appears to be cumbersome, it is worthy of note that the sampling variance is easily estimated by constructing the auxiliary variable $W_i = (\hat{S} + 1) x_i + \frac{\hat{\theta}}{y} y_i - \frac{2}{y} \hat{\theta}$, such that:

$$\hat{\sigma}_S^2 = \frac{1}{n x^2} \hat{\sigma}_w^2$$

(12)

As a by-product of the formulae above, the Gini index estimator and its sampling variance can be obtained without modifying the computing-code: replace the $y$-variable by $y_i=1$, $i=1, \ldots, n$, in the equation (10) and also in the auxiliary variable $W_i$ above. This replacement gives rise a negative Gini coefficient (to be considered positive when presenting results) and the appropriate plug-in estimate of the Gini variance. All these calculations are carried out with just two passes through the data.
3. SIMULATION ANALYSIS.

In this section we analyze the finite sampling performance of alternative inference methods for the Suits index. In particular, we compare the asymptotic and bootstrap statistics in a simulation study. The experiment mimics real data on taxpayers’ taxable-income and the tax paid in the 2006 fiscal year. Taxable-income was generated from a log-normal distribution according to $x=\exp[\mu+\sigma U(0,1)]$ with $\mu=9.1171$, $\sigma=1.1021$ and tax liabilities were computed by applying the 2006 income-tax schedule to taxable-income values. The final file was simulated to have the same size of taxpayers’ population, eleven million units; and population figures for the Gini and Suits indices were calculated (see Table 1). The simulation runs one thousand replications for different sample sizes drawn from this parent file.\textsuperscript{4}

In this experiment, we compare confidence intervals based on the normal approximation with bootstrap intervals, namely, the percentile and the bootstrap-\textit{t} methods. We briefly describe the features of these two methods. Drawing with replacement $B$ independent samples with the same size $n$ as the original sample data (bootstrap resample), we obtain the bootstrap replication $\hat{S}^*$ of the Suits statistic $\hat{S}$. The two-sided \textit{percentile} interval is then defined $[\hat{F}_B^{-1}(\alpha/2), \hat{F}_B^{-1}(1-\alpha/2)]$ where $\hat{F}_B$ is the empirical d.f. of $\hat{S}^*$ and $\alpha$ the nominal size of the confidence interval. The studentized bootstrap or bootstrap-\textit{t} is based on the same logic that underlies the construction of the student’s \textit{t} interval. This method consists in computing:

$$t^*(b) = \frac{\hat{S}^*(b) - \hat{S}}{\hat{\sigma}(\hat{S}^*)}, \quad b=1,...,B.$$  \hspace{1cm} (12)

\textsuperscript{4}The log-norm parameters are weighted estimates obtained from the 2006 Income Tax Return file provided by Spanish Institute for Fiscal Studies (SIFS). This dataset contains one million observations stratified by taxable income. Stratification by income was done to overrepresent higher income-units in order to get accurate estimations of income and tax distributions. We have considered individual-tax returns with positive taxable-income. The Spanish Income Tax is levied on an individual basis but joint-tax returns are also permitted by the fiscal law, although its importance is numerically waning.
where $\hat{S}^*(b)$ is the bootstrap replication of the statistic and $\hat{\sigma}(\hat{S}^*)$ its standard error.

Subsequently, a number of $B$ iterations of $t^*(b)$ are generated in order to replicate the “bootstrap-$t$ table”. The bootstrap-$t$ interval is defined as $[\hat{S} - t_{(1-a/2)} \hat{\sigma}, \hat{S} - t_{(a/2)} \hat{\sigma}]$, where $t_{(.)}$ are the students’ percentiles and $\hat{\sigma}$ is the asymptotic standard error calculated in the original sample. An important consideration in this method is that the bootstrap estimation of the standard error incurs resampling each resample, which might be computationally expensive (this is the estimator used in Anderson et al. 2000). Instead of it, the standard error in (12) is computed according to the asymptotic formula in each $b$-replication. Bootstrap-$t$ is a pivotal statistics and theoretically, it is more asymptotically accurate than the normal approximation and the percentile bootstrap (Beran 1988). Figures 1 and 2 show the difference between the nominal and the real approximate coverages of the confidence intervals. Results regarding the empirical coverage, shape and length are found in table 1. This table also includes the results obtained for the Gini index.

{INSERT FIGURE 1}
{INSERT FIGURE 2}
{INSERT TABLE 1}

Differences between actual and nominal coverages are large in small samples ($n=100,500$) with coverage levels distant from the nominal size. For these sample sizes, the best results are shown by bootstrap-$t$ with $n=500$, with differences in coverage of 3.2% and -0.7% for the two Suits indices respectively. From sample sizes $n \geq 1000$ results show clearly that bootstrap-$t$ procedure has better sampling performance with respect to the others. For both Suits indices, bootstrap-$t$ coverage rates are very close to the nominal.

In particular, for $n=5000$, which is a reasonable size for datasets commonly used in the applied welfare literature (the Luxembourg Income Study datasets used by Bishop et al. (1998) in their empirical application range from ten to fifteen thousand observations), the
bootstrap-t produces small coverage errors (ranging from 0.4% and 0.8% for the two indices considered). Coverage errors are slightly greater for the asymptotic and percentile methods, 1.3% in the case of both asymptotic confidence intervals and 0.9% and 1.3% for the two bootstrap-percentile intervals. For \( n=10000 \) the three methods are very close to the nominal size of the interval.

According to our results, we should remark that Studentization improves the inference of the indices. Bootstrap-t intervals provide wider and asymmetric intervals that can better capture the characteristics of the index finite sampling distribution, leading to intervals which are longer on the right. We conclude that bootstrap-t with an analytical estimation of the index variance is the most accurate procedure in order to make inferences. However, for relatively large samples \( (n \geq 5000) \), the results based on the asymptotic standard normal approximation can yield acceptably inference as an alternative to the bootstrap procedures.

It is also interesting to stress that results for the Gini index are also good. Asymptotic and bootstrap-t confidence intervals provide remarkable coverage results for samples with \( n \geq 1000 \), including the case of \( n=500 \) for the latter method.

4. TESTING FOR CHANGES IN TAX-PROGRESSIVITY AND INCOME REDISTRIBUTION.

In this section, two empirical applications are used to illustrate how asymptotic and bootstrap-t tests are applied for testing changes in tax-progressivity and income-redistribution. First, we analyze the decline in tax-progressivity and income redistribution induced by the lack of inflation-adjustments in the Spanish Income Tax. Second, we provide a sensitivity analysis of the inference tests by comparing three different neutral-revenue income tax cuts in terms of their progressivity profiles. Let us suppose that the difference between two Suits indices is specified as \( D_8 = S_1 - S_2 \). We are interested in testing the null \( H_0 : D_8 = 0 \) against the
alternative $H_1 : D_S \neq 0$ or one-sided hypothesis $H_1 : D_S < 0$ ($D_S > 0$). To test these hypotheses, real data on income taxpayers are directly used rather than simulating income and tax payments from a parametric distribution. The microdata are weighted samples drawn from the Income Tax Returns files and data values were allocated proportional to the size of the original strata. The Suits index estimator and its sampling variance are reformulated by using weighted sample moments in their corresponding formulae to allow for non-identically by independent samples (i.e., Cameron and Trivedi 2005).

4.1 The fiscal drag effect on tax progressivity and income redistribution. 
During the years from 2004 to 2006 the Spanish Income-Tax remained basically unaltered, but income-tax progressivity – defined on the recipients’ net-tax liability and pre-tax income variables- decreased by 12%, according to SIFS figures (Kakwani index diminished from 0.30005 to 0.26338 during that period, see Picos et al. 2011). This decline in tax-progressivity was probably due to the effect of fiscal drag on the income tax function. It is well-known that no inflation-adjustments on income taxes, even when the level of inflation is low, will lead to a less progressive taxation. For a empirical analysis on this issue, see Immerwoll (2005). Before reviewing our results, it is worth to commenting on some aspects of the 2004 and 2006 Spanish Income Taxes. Tax schedules consisted in five income-brackets with marginal tax rates ranging from 15% (taxable-income ≤ 4000€) to 45% (taxable-income≥45000€). Income-limits were updated up to 4% in 2006. This adjustment was clearly insufficient to avoid the ‘bracket creep’ effect as long as the average taxable-income increased up to 15.5% between these years. No other components of the income-tax which could be potentially adjusted, e.g., family allowances, salary deductions and tax credits were modified. Finally, realized capital gains were taxed at 15% flat rate during the whole period.
To identify differences in tax-progressivity over the years under study, it is important to consider first that, in a progressive tax system, changes in progressivity can interact with changes in pre-tax income inequality. This fact makes comparisons of tax-progressivity over time more difficult. For example, the expected reduction on tax-progressivity due to the fiscal drag can be (partially) offset by increasing pre-tax inequality. Table 2 shows pre-tax income Gini coefficients. For testing the null $H_0: G_2 - G_1 = 0$, where the subscripts 1 and 2 refer to 2004 and 2006 respectively, asymptotic and bootstrap-$t$ two-tailed tests were performed. The test based on the normal approximation clearly indicates that we should not reject the null, and the bootstrap-$t$ is not conclusive ($\alpha=0.05$). If we are willing to accept that pre-tax income inequality did not change during that period of time, then the observed reduction in tax-progressivity from 2004 to 2006 should be imputed solely to the ‘bracket creep’ phenomenon and to the absence of non-inflation adjustments in other parameters of the tax function. The next step of the analysis is to test this change on tax-progressivity. To do this, a one-sided test would be appropriated. The null is specified as $D_8 = S_2 - S_1 = 0$ against the alternative $D_8 = S_2 - S_1 < 0$. The null establishes no change in tax progressivity during these years while the alternative implies that the difference is negative, meaning that the 2006 Income Tax is less progressive. The asymptotic test is defined as:

$$t = \frac{\hat{D}_8}{\hat{\sigma}(\hat{D}_8)} \quad (13)$$

where $\hat{D}_8 = \hat{S}_2 - \hat{S}_1$ and $\hat{\sigma}(\hat{D}_8)$ is the standard error of the difference between Suits’ indices $\hat{\sigma}(\hat{D}_8) = \sqrt{\hat{\sigma}^2_{S_1} + \hat{\sigma}^2_{S_2}}$. This $t$-statistic converges to a standard normal distribution. The studentized pivotal statistics or bootstrap-$t$ is defined as:

\footnote{For non-independent samples (i.e., two waves drawn from a balanced panel), the asymptotic standard error is estimated as $\hat{\sigma}(\hat{D}_8) = \sqrt{\hat{\sigma}^2_{S_1} + \hat{\sigma}^2_{S_2} - 2\hat{\sigma}_{S_1S_2}}$ with $\hat{\sigma}_{S_1S_2} = \frac{1}{n_{x^2}} S_{w_1w_2}$. In this case, when performing the bootstrap, resampling is done on the paired observations.}
\[ t^* = \frac{\hat{D}_S - \hat{D}_S}{\hat{\sigma}(\hat{D}_S)} \]  \hspace{1cm} (14)

This formula is obtained for each \( B \)-replication, giving rise to a series of \( t^* \)-values. The following steps are to be considered:

a) Draw a \( B \)-sample for 1 and 2 with same sizes as the original sample. Using the plug-in formulae, compute the Suits index estimator and its standard error in each \( B \)-sample.

b) Compute \( \hat{D}_S, \hat{\sigma}(\hat{D}_S), \hat{D}_S \) and the statistic \( t^* \). The difference \( \hat{D}_S \) is estimated using the original samples.

c) Replicate \( B \) times steps a) and b) to obtain a \( t^* \)-values set of size \( B \).

d) As the rejection region of the progressivity tests is on the left, calculate the critical point of the test as the \( \alpha \)100th percentile of the \( t^* \)-distribution. We reject the null if the \( t \)-statistic is less than this critical value. Bootstrap-\( t \) for the Gini and income-redistribution Suits indices are symmetrical two-tailed tests: the absolute values of \( t^* \) are ordered and the bootstrap critical value is the (1- \( \alpha \))th percentile of the ordered series. The null is rejected if \(|t|\) exceeds this critical value.

Table 2 presents the Gini and Suits index estimates for the years under study. Suits index values are computed on gross-tax, net-tax liabilities and post-tax income (columns 2, 3 and 4, respectively). This table also shows the differences \( \hat{D}_S \) and the asymptotic and bootstrap-\( t \) values for testing changes in tax-progressivity and income redistribution.

\{{\text{INSERT TABLE 2}}\}

The estimated differences \( \hat{D}_S \) for the gross and net tax-liabilities represented respectively a 19% and 22% reduction of tax-progressivity between these years\(^6\). Whichever statistic is used

\(^6\) The percentual reduction on net-tax liability progression using the Suits index (22%) is higher than the Kakwani index, estimated by SIFS (12%). There are two methodological aspects which might explain these differences. First, our samples do not include joint-tax returns as the SIFS files do and second, the Suits index assigns more weight on high-income units rather than the Kakwani index when measuring overall tax-progressivity.
results show strong statistical evidence in favour of the alternative, reinforcing the hypothesis of the loss in tax-progressivity as a consequence of fiscal drag.

Our results also show a 21% decline in income redistribution over time (third Suits index in Table 2). The income redistribution effect exerted by progressive taxes depends not only on the progression of the tax but also on the tax height and the reraking occurring in the post-tax income distribution. If these components of the redistributive effect act in different directions regarding progressivity, we cannot clearly anticipate the sign of the difference between the two income-redistribution indices, and hence, an asymptotic two-sided test would be more appropriate. According to these results, the $t$-statistic leads to reject the null of no change in income redistribution at the usual level of significance. The bootstrap-$t$ statistic also confirms this fact. In this case, a symmetrical two-tailed test is performed and the absolute value of the $t$-statistic exceeds the bootstrap critical value. The inference tests prove useful in demonstrating the effect that the fiscal drag had on income taxation as the loss of progressivity and income redistribution is statistically significant.

### 4.2 A sensitivity analysis for changes in tax-progressivity: testing ‘linear’ tax reforms.

As an additional assessment of the inference tests, a sensitivity analysis is carried out by comparing three different neutral-revenue income-tax cuts simulated from a baseline income-tax. These tax cuts are defined as (1) a fraction $a$ of tax liability, (2) a fraction $b$ of post-tax income, (3) a fraction $c$ of pre-tax income. For comprehensive income-taxes, these three tax cuts can be ordered according to: $L_{V_1} \prec_L L_{V_3} \prec_L L_{V_2}$ and $L_{T_1} \prec_L L_{T_3} \prec_L L_{T_2}$, where $\prec_L$ denotes the Lorenz Dominance criterion and $L_V$ (resp. $L_T$) denotes the Lorenz curve of post-tax income (resp. tax liability) distribution. For tax hikes, the order is reversed, see Pfähler (1984) for a detailed analysis. In the case of dual income taxes, Lorenz dominance among these tax cuts (or increases) requires imposing additional assumptions on labour and capital income
distributions, see Calonge and Tejada (2011) and Lambert (2012). The distributional consequences of applying such tax cuts, or increases, can also be consulted in these papers. In this application, tax cuts are applied to the 2007 Spanish Dual Income Tax. Tax cuts are deduced as linear transformations of the baseline tax schedule when applying \( a=2\% \) reduction on total tax revenue (see Calonge and Tejada for a description of the 2007 Dual Income Tax and the procedure used to deduce the simulated tax cuts). Table 3 shows marginal tax rates and the average revenue \( T \) for the three tax cut scenarios.

To evaluate differences in tax-progressivity, one-sided tests have been performed. For example, testing whether type-1 tax cut is less progressive than type-2 tax cut leads to reject the null \( H_1: D=S_1-S_2=0 \) in favour of the alternative \( H_1: D=S_1-S_2<0 \). As the income distribution is fixed (we are using a single sample to simulate the three different income taxes), the difference between Suits’ indices can be expressed as \( H_0: D=\frac{2}{\mu_X \mu_Y} (\theta_1 - \theta_2) \).

Assuming yield-equivalent tax cuts, it is true that \( \mu_{Y_1} = \mu_{Y_2} = \mu_Y \) and the null is then written as

\[
H_0: D=\frac{1}{\mu_X \mu_Y} (\theta_1 - \theta_2) = 0
\]

which is simplified to test \( H_0: \theta_1 - \theta_2 = 0 \).

The estimated null \( H_0: \hat{D} = \hat{\theta}_1 - \hat{\theta}_2 \) is a linear function of two U-statistics where estimators are calculated for each tax variable \( Y_2 \) and \( Y_1 \) and the sampling variance is estimated as

\[
\hat{\sigma}_D^2 = \frac{4}{n} \hat{\sigma}_W^2,
\]

where in this case \( W_i = \hat{\theta}_{1i} - \hat{\theta}_{2i} \).

The results for asymptotic tests are presented in table 3. These results agree with the theoretical statement that \( T_2 \) is the most progressive and redistributive tax cut and \( T_1 \) the least. The null hypotheses are strongly rejected in favour of the alternative in all cases.
5. CONCLUSION.

This article provides simple plug-in estimators for the Suits index and its sampling variance based on non-degenerate $U$-statistics theory. Inference based on first-order asymptotic approximation is compared with bootstrap methods using a simulation analysis. Focusing on confidence intervals and their real coverage probabilities, results reveal that the bootstrap-$t$ - which uses the plug-in variance estimator to form the pivotal $t$ statistic – has superior performance. However, asymptotic procedures provide good results for sample sizes which are common in this research area.

Using real data corresponding to the Spanish Income Tax, our results show that the asymptotic test and boot-$t$ methods perform quite well to detect changes in income-tax progressivity over time. Specifically, our results show strong evidence in favour of the fiscal drag hypothesis to explain the reduction in tax-progressivity and income redistribution experimented by the Spanish Income Tax in a relatively short period of time. Finally, an additional application proves the accuracy of the empirical tests to identify slight differences in progressivity among three different neutral-revenue income taxes. In both applications, empirical results clearly agree with our theoretical expectations.
REFERENCES


APPENDIX A1.

The functional $\theta$ can be rewritten as

$$\theta(F) = \int_0^{+\infty} \int_0^{+\infty} y_1 \{ \int_0^{+\infty} x_2 I(x_2 \leq x_1) dF(x_2) \} dF(x_1, y_1) = E(y F^*(x))$$

where $F^*(x) = \int_0^{+\infty} x_2 I(x_2 \leq x_1) dF(x_2)$ represents the cumulative income function. Second, from $E(Y, F^*(x)) = Cov(Y, F^*(x)) + E(Y)E(F^*(x))$, expression (4) in section 2.2 can be expressed as:

$$S = \frac{2Cov(Y, F^*(x))}{\mu_x \mu_y} + \frac{2E(F^*(x))}{\mu_x} - 1 \quad \text{A1.1}$$

Consider that:

1. $E(F^*(x)) = \int_0^{+\infty} \int_0^{+\infty} x_2 I(x_2 \leq x_1) dF(x_2) dF(x_1)$ can be rewritten after some slight rearrangement of the integration variables as $E(F^*(x)) = \int_0^{+\infty} x_2 \left\{ \int_0^{+\infty} I(x_2 \leq x_1) dF(x_1) \right\} dF(x_2)$. This expression leads to $E(F^*(x)) = E(x(1-F(x))$

2. $E(xF(x)) = Cov(xF(x)) + E(x)E(F(x))$ and $E(F(x)) = 1/2$

Substituting these elements in equation A1.1, it is relatively straightforward to obtain the ‘covariance’ formula of the Suits index as:

$$S = \frac{2Cov(Y, F^*(x))}{\mu_x \mu_y} - \frac{2Cov(x, F(x))}{\mu_x} = C_y - G_x \quad \text{A1.2}$$

APPENDIX A.2.

Lemma 1 applies Hoeffding’s Theorem 7.1 (Hoeffding 1984) to obtain the asymptotic joint-distribution of the estimators involved in the Suits formula.

**Lemma 1.** Let $F(x, y), F(x)$ and $F(y)$ be continuous d.f. with $E(x) < \infty, E(y) < \infty, E(x)^2 < \infty$

, $E(y)^2 < \infty$, $E(xy)^2 < \infty$, then the random vector

$$\sqrt{n}[(\bar{x} - \mu_x)(\bar{y} - \mu_y), (\hat{\theta} - \theta)]$$

20
tends, as \( n \to \infty \), to the 3-variate normal distribution with zero mean and variance-covariance matrix \( \Sigma \) where its \( i,j \)-element is given by \( m_i m_j \xi_{ij} \), being \( m \) the degree of the functionals. The term \( \xi_{ij} \) represents covariance between *conditional expectations* and they are defined as:

\[
\xi_{13}(\mu_1, \theta) = E\left[ \left[ \theta_1(x_1, y_1) - \theta \right] \left[ \theta_3(x_1, y_1) - \theta \right] \right] = \int \int x_1 y_1 \left\{ \int x_2 I(x_2 < x_1) dF(x_2) \right\} dF(x_1, y_1) - \mu_1 \theta
\]

\[
\xi_{23}(\mu_2, \theta) = E\left[ \left[ \theta_2(x_1, y_1) - \theta \right] \left[ \theta_3(x_1, y_1) - \theta \right] \right] = \int \int y_1^2 \left\{ \int x_2 I(x_2 < x_1) dF(x_2) \right\} dF(x_1, y_1) - \mu_2 \theta
\]

\[
\xi_{33}(\theta_1, \theta_2, \theta_3) = E\left[ \left[ \theta_3(x_1, y_1) - \theta \right]^2 \right] = \int \int y_1^2 \left\{ \int x_2 I(x_2 < x_1) dF(x_2) \right\}^2 dF(x_1, y_1) - \theta^2
\]

and

\[
\xi_{11}(\mu_1, \mu_2) = \sigma_{\xi_1}^2, \quad \xi_{12}(\mu_1, \mu_2) = \sigma_{\xi_1 \xi_2}^2, \quad \xi_{13}(\mu_1, \mu_2) = \sigma_{\xi_1 \xi_3}^2.
\]

Next, the application of the Delta method leads to the index asymptotic distribution of Suits index. Let now represent the following Jacobian-vector of partial derivatives of the index \( S(\mu_x, \mu_y, \theta) \) as

\[
J = \frac{1}{\mu_x} \left[ -\frac{S+1}{\mu_x} \frac{1}{\mu_y} \right]
\]

The following proposition leads to the asymptotic distribution of the Suits index.

**Proposition 1.** The same conditions as in *lemma 1* apply. The function \( S \) does not involves \( n \) and if it is assumed continuous together with its second order partial derivatives in some neighbourhood of the parameters, then the Suits index estimator has a limiting normal distribution:

\[
\sqrt{n} (\hat{S} - S) \xrightarrow{d} N(0, \sigma_s^2 \equiv J \Sigma J')
\]

where the asymptotic variance \( \sigma_s^2 \) is given by:
\[
\sigma_S^2 = \frac{1}{\mu_x^2} [(S+1)^2 \sigma_x^2 + \theta^2 \frac{\sigma^2}{\mu_y^2} + \frac{4 \xi_{3,3}}{\mu_y^2} + 2(S+1) \frac{\theta}{\mu_y^2} \sigma_{xy} - \frac{4(S+1)}{\mu_y^2} \xi_{1,3} - \frac{4\theta \xi}{\mu_y^2} \xi_{2,3}] \tag{A2.1}
\]

**Proof.** Equation (20) comes directly from applying Hoeffding’s theorem 7.5. Covariance functionals and the index \( S \) should be replaced by their consistent plug-in estimators to estimate the variance. For instance, the plug-in estimator of the functional \( \xi_{33} \) would be \( \hat{\xi}_{33} = \sum (\hat{\theta}_i - \bar{\theta})^2 \). However, the sampling variance is easier to estimate by using the sampling variance of a scalar (see section 2.3). ■
Figure 1. Relative concentration, Lorenz and Tax concentration curves.

Figure 2. Suits progressivity index: coverage errors of confidence intervals (α=0.05).
Figure 3. Suits redistribution index: coverage errors of confidence intervals ($\alpha=0.05$).
Table 1. Simulation analysis: real approximate coverage, shape and length of the Suits index confidence intervals ($\alpha=0.05$).

<table>
<thead>
<tr>
<th>Index concept</th>
<th>$n = 100$</th>
<th>$n = 500$</th>
<th>$n = 1000$</th>
<th>$n = 2000$</th>
<th>$n = 5000$</th>
<th>$n = 10000$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asymptotic confidence interval</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>coverage</td>
<td>.861</td>
<td>.917</td>
<td>.950</td>
<td>.929</td>
<td>.944</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.148</td>
<td>.078</td>
<td>.056</td>
<td>.040</td>
<td>.026</td>
</tr>
<tr>
<td>Suits</td>
<td>coverage</td>
<td>.753</td>
<td>.889</td>
<td>.910</td>
<td>.908</td>
<td>.937</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.045</td>
<td>.026</td>
<td>.019</td>
<td>.014</td>
<td>.009</td>
</tr>
<tr>
<td>Suits*</td>
<td>coverage</td>
<td>.720</td>
<td>.896</td>
<td>.909</td>
<td>.911</td>
<td>.937</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.029</td>
<td>.018</td>
<td>.013</td>
<td>.010</td>
<td>.007</td>
</tr>
<tr>
<td><strong>Bootstrap percentile</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>coverage</td>
<td>.827</td>
<td>.904</td>
<td>.945</td>
<td>.929</td>
<td>.940</td>
</tr>
<tr>
<td></td>
<td>shape</td>
<td>.722</td>
<td>.907</td>
<td>.953</td>
<td>.983</td>
<td>1.003</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.146</td>
<td>.077</td>
<td>.056</td>
<td>.041</td>
<td>.026</td>
</tr>
<tr>
<td>Suits</td>
<td>coverage</td>
<td>.671</td>
<td>.878</td>
<td>.899</td>
<td>.904</td>
<td>.941</td>
</tr>
<tr>
<td></td>
<td>shape</td>
<td>.435</td>
<td>.620</td>
<td>.722</td>
<td>.806</td>
<td>.888</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.051</td>
<td>.026</td>
<td>.019</td>
<td>.014</td>
<td>.010</td>
</tr>
<tr>
<td>Suits*</td>
<td>coverage</td>
<td>.705</td>
<td>.887</td>
<td>.908</td>
<td>.907</td>
<td>.937</td>
</tr>
<tr>
<td></td>
<td>shape</td>
<td>2.419</td>
<td>1.333</td>
<td>1.187</td>
<td>1.106</td>
<td>1.052</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.029</td>
<td>.018</td>
<td>.013</td>
<td>.010</td>
<td>.007</td>
</tr>
<tr>
<td><strong>Bootstrap-t</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>coverage</td>
<td>.915</td>
<td>.934</td>
<td>.948</td>
<td>.938</td>
<td>.943</td>
</tr>
<tr>
<td></td>
<td>shape</td>
<td>1.599</td>
<td>1.467</td>
<td>1.355</td>
<td>1.286</td>
<td>1.218</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.203</td>
<td>.091</td>
<td>.061</td>
<td>.043</td>
<td>.027</td>
</tr>
<tr>
<td>Suits</td>
<td>coverage</td>
<td>.840</td>
<td>.918</td>
<td>.941</td>
<td>.936</td>
<td>.946</td>
</tr>
<tr>
<td></td>
<td>shape</td>
<td>2.592</td>
<td>1.4201</td>
<td>1.281</td>
<td>1.262</td>
<td>1.225</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.076</td>
<td>.0344</td>
<td>.0231</td>
<td>.016</td>
<td>.010</td>
</tr>
<tr>
<td>Suits*</td>
<td>coverage</td>
<td>.884</td>
<td>.957</td>
<td>.946</td>
<td>.933</td>
<td>.942</td>
</tr>
<tr>
<td></td>
<td>shape</td>
<td>.540</td>
<td>.702</td>
<td>.743</td>
<td>.768</td>
<td>.808</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>.068</td>
<td>.025</td>
<td>.016</td>
<td>.011</td>
<td>.007</td>
</tr>
</tbody>
</table>

Values of the population parameters: Gini=0.56413, Suits=0.15730, Suits* =0.06017. Number of bootstrap replications $B=999$. The statistic shape is a measure of the interval asymmetry and it is defined as the ratio between the left-side length of the interval over its right-side length.
Table 2. Tests for changes in tax-progressivity and income-redistribution ($\alpha=0.05$).

<table>
<thead>
<tr>
<th>Period</th>
<th>Gini Index</th>
<th>Suits index</th>
<th>Suits index</th>
<th>Suits index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (2004)</td>
<td>.42757</td>
<td>.28316</td>
<td>.32785</td>
<td>.058929</td>
</tr>
<tr>
<td></td>
<td>(.012439)</td>
<td>(.012495)</td>
<td>(.013041)</td>
<td>(.0033052)</td>
</tr>
<tr>
<td>2 (2006)</td>
<td>.45938</td>
<td>.22768</td>
<td>.25558</td>
<td>.046408</td>
</tr>
<tr>
<td></td>
<td>(.026167)</td>
<td>(.026861)</td>
<td>(.025413)</td>
<td>(.0048915)</td>
</tr>
<tr>
<td>$D_s$</td>
<td>.03168</td>
<td>-.05548</td>
<td>-.07227</td>
<td>-.012521</td>
</tr>
<tr>
<td>$t$-value</td>
<td>1.09791</td>
<td>-1.87256</td>
<td>-2.53035</td>
<td>-2.12104</td>
</tr>
<tr>
<td>$t^*$-value</td>
<td>1.09315</td>
<td>-.74961</td>
<td>-.79243</td>
<td>1.74850</td>
</tr>
</tbody>
</table>

Data are two independent samples (of sizes 6006 and 6341) drawn from the 2004 and 2006 SIFS Income Tax Return Files. Standard errors are in parenthesis. It is interesting to note that the application of the bootstrap-$t$ to test each hypothesis is not computationally costly (54 seconds long using an i7-CPU processor, 8 GB RAM). A SAS programme is available on request to the authors.

Table 3. Sensitivity analysis.

| Baseline tax-schedule | Alternative Tax cut rules |  |  |
|-----------------------|---------------------------|  |  |
|                       | 1            | 2            | 3            |
| 24                    | 23.520%      | 23.394%      | 23.430%      |
| 28                    | 27.440%      | 27.426%      | 27.430%      |
| 37                    | 36.260%      | 36.498%      | 36.430%      |
| 43                    | 42.140%      | 42.545%      | 42.430%      |
| $\bar{T}(\€)$         | 3964         | 3969         | 3967         |
| Tax-progressivity Suits index | .29057 | .29277 | .29214 |
| $D_{\theta} = \theta_1 - \theta_2$ | -182875 (24502) | -52031 (6966) |
| $t$-value             | -7.643       | -7.469       |

Income taxes paid under the three scenarios are simulated using a fixed sample of taxpayers of size 6000. Standard errors are in parenthesis.