Económica, La Plata, Vol. XLVI , Nro.1, Año 2000

A NOTE ON THE STATISTICAL SIGNIFICANCE OF CHANGES IN INEQUALITY[1]

WALTER SOSA ESCUDERO y LEONARDO GASPARINI[2]

1. Introduction

Between 1993 and 1994 the Gini coefficient for Greater Buenos Aires rose from 0.443 to 0.457 based on information from the Permanent Household Survey (EPH).[3] If the interviewed households had been the same in both years and if their income were measured without error, this observed difference in the Gini coefficients should be unambiguously interpreted as an increase in inequality in the distribution of income. But if, as it is the case of the EPH, interviewed households differ between surveys, this increase in the Gini coefficient could have simply been caused by sampling variation rather than an increase in inequality.

In spite of this observation, most of the applied work on the subject is based on point estimates of inequality measures, ignoring the sampling variability problem. This issue is far to be minor since, as this note clearly shows, many of the observed changes in inequality for several regions in Argentina turn out to be statistically insignificant. This note utilizes the bootstrap to provide a simple and efficient way to compute interval estimates and standard errors for inequality measures for the case of Argentina, where income inequality presented sudden changes in both directions in the last two decades, making crucial the issue of distinguishing sampling variability from true changes in the distribution of income. The same procedure is also used to formally test the null hypothesis of no change in inequality between any pair of periods.

The rest of the note is organized as follows. In section 2 we discuss the use of bootstrap methods to deal with the problem of sample variability in

---

[1] JEL Classification: D3, D6, C4

[2] Departamento de Economía, Facultad de Ciencias Económicas, Universidad Nacional de La Plata, calle 6 entre 47 y 48, oficina 516, 1900 La Plata, Argentina. Phone-fax: 54-221-4229383. Correo electrónico: wsosa@feedback.net.ar and lgmm@netverk.com.ar.

[3] These figures correspond to the Gini coefficient of the distribution of household per-capita income based on the October wave of the Permanent Household Survey conducted by the National Institute of Statistics and Census (INDEC).

inequality measures, particularly in the Gini coefficient. An application of these techniques to derive standard errors and confidence intervals for the case of Argentina is presented in section 3. Section 4 concludes with some final remarks.

2. Bootstrap methods for inequality measures

Let Y be a positive and continuous random variable which represents income. We will denote its cumulative distribution function (CDF) with $F(y)$. In general terms, an inequality measure will be understood as a characteristic of the distribution of income which represents how total income is dispersed between individuals.[4] For simplicity, the following discussion will be illustrated using only the popular Gini coefficient G, defined as (Maasoumi, 1997):

$$G = (2/\mu_Y)\int_0^\infty y[F(y) - 1/2]dF(y)$$

where μ_Y is the expected value of the distribution of incomes. Based on a random sample of N individuals with incomes Y_i, $i=1,\ldots,N$, estimation of G usually proceeds by using a standard estimator like:

$$G^* = \frac{1}{2N^2\bar{y}}\sum_{i=1}^{N}\sum_{j=1}^{N}|Y_i - Y_j|$$

where $\bar{y}$ is the mean of the distribution of incomes.[5] Even though the estimated Gini coefficient is a complicated function of the data, large sample variances can be derived from the theory of U-statistics (Cowell, 1989).

In practice there are two reasons to consider alternative routes to the use of asymptotic results. First, there is the usual caveat about the reliability of large sample based methods when applied to finite samples. Second, large sample methods can become quickly cumbersome (or maybe nonexistent) when the interest is in additional features like the statistical significance of temporal changes in the Gini coefficient, the sampling variability of a welfare measure (which involves the mean income and the Gini coefficient in a non-linear fashion), or, simply, other alternative measures of inequality.

[4] See Massoumi (1997) for a recent survey on inequality measurement.
[5] See Lambert (1993).

Under very general conditions, the bootstrap provides a simple and convenient framework to handle the problem of sampling variation of many inequality measures, including the Gini coefficient. The bootstrap has recently became a widely used technique in applied statistics and econometrics. One reason for this increasing popularity is that even though the formal aspects of why it works are quite involved, it is intuitively simple to understand and implement[6]. Bootstrap methods for inequality measures have recently been used by Mills and Zandvakili (1997) and Gasparini and Sosa Escudero (1999).

Denote with θ a characteristic of the distribution which is the subject of interest, like the Gini measure of inequality. Let $\theta_n(\mathbf{Y})$ be an estimator of θ based on a i.i.d sample of size N with $\mathbf{Y} = (Y_1, Y_2, ..., Y_N)$. $\theta_n(\mathbf{Y})$ will be a random variable whose distribution we will denote with $G_n(q; F) = \Pr[\theta_n(\mathbf{Y}) \leq q]$. This notation stresses the fact that this distribution is actually a transformation of the original CDF of the variable of interest (income, in our case).

The first goal is to obtain a variance for $\theta_n(\mathbf{Y})$ and a confidence interval for θ. Consider the case of the variance (denoted by S) first. This will be:

$$S = \mathrm{E}[\theta_n(\mathbf{Y})^2 - \mathrm{E}\,\theta_n(\mathbf{Y})^2\,]$$

The analytic evaluation of the expectation in this expression requires knowledge of the distribution $G_n(q, F)$. In practice, and except for very particular cases, there are two difficulties. The first one is that the original distribution of incomes (F) is not known. The second one is that even when F is known, the derivation of $G_n(q; F)$ and characteristics like S could be analytically very cumbersome, when not impossible. A common solution is to rely on asymptotic expansions, which provide an adequate representation of the statistic of interest (or of a transformation), valid when the sample size is large.

The bootstrap method consists in simply approximating $G_n(q, F)$ using the *empirical* distribution of the sample (F_n), that is, using $G_n(q, F_n)$. Even when in practice this solves the problem of not knowing F, it remains to solve the problem of computing $G_n(q, F_n)$ from F_n and any characteristic of interest

---

[6] Efron and Tibshirani (1993) and Davidson and Hinkley (1997) provide excellent introductions to the subject. Horowitz (1997), Jeong and Maddala (1997) and Veall (1998) are useful surveys oriented to econometric applications.

like S. Efron (1978), in his seminal paper, proposes computing $G_n(q, F_n)$ and S based on the following Monte Carlo experiment:

1. Obtain a random sample with replacement from the empirical distribution of Y of size N.
2. Compute the statistic of interest from the realizations of the sample obtained in the previous step.
3. Repeat the process B times.
4. From the previous replications we will obtain B (bootstrap) estimates of θ. The method consists in approximating $G_n(q, F_n)$ with the empirical distribution of θ_n based on the estimates obtained in the previous steps. For example, for the case of the variance of the Gini coefficient, the method suggests to simply compute the sampling variance error using the B bootstrap estimates of θ.

Intuitively, if $F(\ )$ were known and we could sample from it, when N goes to infinity the empirical distribution G_n converges uniformly to G by application of the Fundamental Theorem of Statistics (Davidson and MacKinnon, 1993). Instead, the bootstrap method actually samples from the original sample as if it were the population, and the Monte Carlo step is performed as a simple way to evaluate an integral like the one involved in the computation of the variance.

In order to construct a 95% confidence interval (G_L,G_U) for the Gini coefficient we use the percentile method, that is, we take as G_L and G_U the 0.025 and 0.975 quantiles of the empirical distribution of the bootstrapped coefficients. The procedure to evaluate the null hypothesis of no change in the Gini coefficients of two distributions at different periods is similar to the one described above. In this case the population of interest consists in all incomes for a given pair of years to be compared. The bootstrap proceeds by taking a sample with replacement for each of the years in the comparison, computing the Gini coefficient for each year and their difference. According to the duality between interval estimation and hypothesis testing, the test rejects the null of equality if a confidence interval for the difference in the Gini coefficients does not contain zero.

It is important to remark that the procedure described above applies, under general conditions, to any inequality measure. For example, Mills and

Zandvakili (1997) apply the bootstrap to the Theil coefficient and other decomposable measures of inequality, and Gasparini and Sosa Escudero (1999) use the method to evaluate the sampling variation of the Sen measure of welfare.

A numerical example

In order to gain some understanding of the reliability of the bootstrap method for the case of inequality, consider the following numerical example. Assume that income is known to be log-normally distributed with parameters μ and σ^2, that is, $\ln Y \sim N(\mu, \sigma^2)$. For the log-normal case, the Gini coefficient can be shown to be equal to $2\,\Phi(\sigma/\sqrt{2})-1$, where Φ is the cumulative normal standard distribution (Cowell, 1997).

We set $\mu = 0$ and $\sigma^2 = 0.95387$, so for this case the Gini coefficient will be equal to 0.5. We take a sample of 1000 individuals from this distribution and estimate the Gini coefficient using the formula above, which gives a value of 0.49291. We then perform the bootstrap to obtain a standard error for the estimated Gini coefficient[7]. We take 1000 samples of size 1000 with replacement from the original sample and compute the Gini coefficient for each of them. The standard error of the bootstrapped Gini coefficients gives a value of 0.01011.

Cowell (1997, p. 118) suggests approximating the standard error of the estimated coefficients using the formula $0.8086\, c / N^{1/2}$, where c is the sample coefficient of variation. This formula is valid when the underlying distribution is normal and the sample size is large. This method gives an estimated standard error of 0.0277, considerably larger than the value obtained using the bootstrap.

In order to explore the accuracy of the bootstrap, we performed a small Montecarlo experiment, extracting 1000 samples of size 1000, in this case from the original (known) distribution of income, compute in each case the Gini coefficient, and from the empirical distribution, compute the standard error, which gives a value of 0.01304, which is very similar to the one obtained by bootstrapping the original data.

---

[7] All computations were performed using the bootstrap module of Splus 4.0 for Windows. Specific routines for inequality measures are available by request to the authors.

This simple example illustrates how the bootstrap can provide reliable approximations to the characteristics of interest, specially in cases where the applicability of large sample approximations are questionable.

3. Applications to the case of Argentina's EPH

Inequality measures for Argentina are usually based on the Permanent Household Survey (EPH) collected and processed by the local National Statistical and Census Institute (INDEC) in 23 major urban areas.

First consider the case of Greater Buenos Aires (GBA), an area with nearly a third of the population of Argentina. The EPH gathers information on around 3,400 households (more than 11,000 individuals). The first panel of Table 1 and Figure 1 present the evolution of the Gini coefficient of the distribution of per capita household income for that urban area..

The computed Gini increased from 1992 to 1995 and then slightly fell. According to Table 1, inequality in 1997 was much higher than in 1992. However, as mentioned before, since surveyed households change period by period, these variations could be due to alterations in income distribution, or simply to the fact that the sample had changed, or to both factors. In order to quantify the sampling variability Table 1 accompanies Gini estimates with bootstrapped standard errors, the corresponding coefficient of variation and confidence intervals.[8] In all cases, the number of bootstrap replicates is B=300. Given the large size of the sample, we expect the Gini coefficients to be estimated with high precision. This is reflected in the low values of the standard errors (0.0072 on average) and the coefficients of variation (1.6% on average).

Although Gini coefficients are apparently estimated with high precision, researchers and policy makers are usually more concerned with changes in

[8] The use of the simple bootstrap requires independence of observations in the sample for a particular cross-section. This is clearly not the case if computations are based on individuals (where within-family effects are likely to be strong). But since the relevant income for each individual is the per-capita income (which is constant within the family), the problem is easily overcome by sampling *households* (instead of individuals), for which the independence assumption is more realistic while preserving the (trivial) within household dependence. Actually, this would conform to a special case of the *block bootstrap*, which is a valid method for dependent observations (see Davison and Hinckley (1997)). A simple formula to compute the Gini coefficient based on individual level per-capita income using sampling weights is given in Deaton (1997, pp. 154) and is the one used in this paper.

inequality. The top panel of Table 2 shows the changes in the Gini coefficients for all pairs of years between 1992 and 1997, for the case of Greater Buenos Aires. The third column shows the differences between the Gini coefficients for each pair of years. Columns 4 to 7 show the percentiles of the bootstrapped distribution of these differences. For example, the numbers in columns 5 and 6 correspond to a confidence interval of 90%. According to the previously described procedure, the null hypothesis of equality between the Gini coefficients is rejected if the confidence interval for this difference does not include the number zero. In each row it is indicated with a "*" whether the null hypothesis is rejected for a significance level of 0.95.[9]

For the case of GBA, out of 15 possible comparisons, 9 of them turn out to be statistically significant. As it can be observed, the cases in which equality can not be rejected correspond, in general, to comparisons between successive years. Except in one case, for the rest of the comparisons between consecutive years it is not possible to reject the null hypothesis of absence of changes in the Gini coefficient. This implies that according to the experience of the 90's and given the current sample size, it seems precipitated to state propositions about the evolution of inequality from the observation of the Gini coefficient for two consecutive years. Trend changes in inequality take some time to be evident. It is likely that the analysis of two consecutive years capture more sample variability (noise) than real changes (signal).

It is interesting to contrast the case of Buenos Aires with a smaller urban area like Neuquén, where the sample size is on average 748 households. The level and the evolution the Gini measure is similar to that of Buenos Aires, as can be seen in Table 1 and Figure 1. As expected, the smaller sample size traduces in larger standard errors and wider confidence intervals. Whereas in Buenos Aires the average coefficient of variation was 1.6%, for Neuquén this figure is 2.9%. In absolute terms this number may seem low, so in spite of the reduction in the sample size it is still possible to obtain precise estimates of the Gini coefficient. The problem arises when the focus is the change in inequality. The bottom panel of Table 2 shows that even though the evolution of the Gini measure is similar to that of Buenos Aires, only two of the 15 observed changes are statistically significant. Detailed results for several

[9] The EPH has a rotating panel structure, which creates temporal dependence between observations As Mills and Zandvakili (1997) stress, the dependence between time periods is automatically taken into account by the bootstrap since each subsample is conditional on the data observed in the previous period.

regions of Argentina emphasize the conclusions of this note and are not presented in the paper in order to save space.[10]

4. Final remarks

The usual practice in income distribution analysis is to draw conclusions on inequality changes from the comparison of indices computed from household surveys. This procedure ignores that changes in inequality measures can be partly driven by sample variability. As this note clearly illustrates, the bootstrap provides a simple and efficient way to compute standard errors and confidence intervals for inequality measures and their changes, allowing us to assess the effect of sample variability on the results.

The application to Greater Buenos Aires data in the nineties reveals that although there was a clear increasing trend in inequality, changes between two successive years have not been statistically significant, given the current sample size. The case of Neuquén, where almost all changes in the Gini coefficients are not statistically significant, clearly shows that researchers and policy makers should be very careful when evaluating distributional aspects in regions where the sample size is not large enough to distinguish true changes in inequality from sampling variability.

[10] They can be obtained by request from the authors.

REFERENCES

COWELL, F., 1989, "Sampling Variance and Decomposable Inequality Measures", *Journal of Econometrics*, 42, 27-41.

COWELL, F., 1997, *Measuring Inequality*, 2nd edition, Prentice Hall.

DAVISON, A. AND HINKLEY, D., 1997, *Bootstrap Methods and their Application*, Cambridge University Press.

DAVIDSON, R. AND MACKINNON, J., 1993, *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.

EFRON, B. , 1978, "Bootstrap Methods: Another Look at the Jacknife", *Annals of Statistics*, 7, 1-26.

EFRON, B. AND TIBSHIRANI, R., 1993, *An Introduction to the Bootstrap*. Chapman & Hall.

GASPARINI, L. AND SOSA ESCUDERO, W., 1999, "Assessing Aggregate Welfare: Growth and Inequality in Argentina", *Working paper*, Universidad Nacional de La Plata.

HOROWITZ, J., 1997, "Bootstrap Methods in Econometrics: Theory and Numerical Performance". In Kreps, D. and Wallis, K. (eds.), *Advances in Economics and Econometrics: Theory and Applications*, Cambridge University Press, Cambridge.

JEONG, J. AND MADDALA, G., 1993, "A Perspective on Application of Bootstrap Methods in Econometrics". In Maddala, G., Rao, C., and Vinod, H. (eds.), *Handbook of Statistics, Vol. 11*, North-Holland, Amsterdam.

LAMBERT, P., 1993, *The Distribution and Redistribution of Income. A Mathematical Analysis*. Manchester University Press.

MASSOUMI, E., 1997, "Empirical Analyses of Inequality and Welfare". In Pesaran, H. and Schmidt, P. (eds.), *Handbook of Applied Econometrics, vol II*: Microeconomics, Blackwell Publishers, Oxford.

MILLS, J., AND ZANDVAKILI, S., 1997, "Statistical Inference via Bootstrapping for Measures of Inequality". *Journal of Applied Econometrics* 12, 133-150.

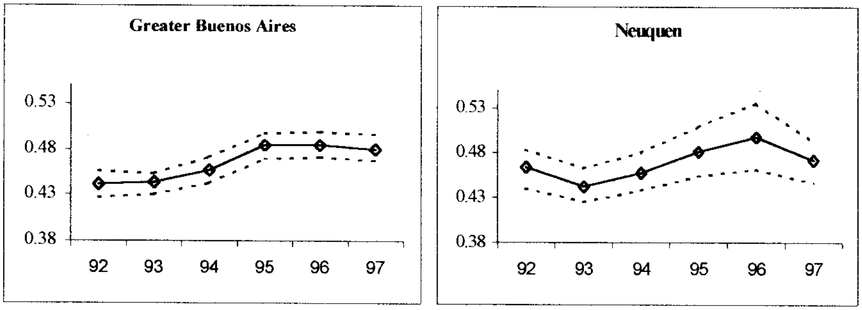
VEALL, M., 1998, "Applications of the Bootstrap in Econometrics and Economic Statistics", in Ullah, A. and Giles, D. (eds.), *Handbook of Economic Statistics*, Marcel Dekker Inc., New York, NY.

Table 1
Gini coefficients and bootstrapped confidence intervals and standard errors
Greater Buenos Aires and Neuquén, 1992-1997

Region	Year	Gini	Confidence Interval				Standard	Coef. Of
			0.025	0.05	0.95	0.975	Error	Variation
GBA	92	0.4415	0.4264	0.4285	0.4529	0.4548	0.0072	1.6%
	93	0.4430	0.4298	0.4308	0.4520	0.4529	0.0062	1.4%
	94	0.4570	0.4418	0.4440	0.4691	0.4708	0.0077	1.7%
	95	0.4843	0.4687	0.4704	0.4955	0.4971	0.0074	1.5%
	96	0.4840	0.4705	0.4721	0.4956	0.4982	0.0073	1.5%
	97	0.4797	0.4670	0.4686	0.4938	0.4958	0.0074	1.6%
Neuquén	92	0.4632	0.4393	0.4423	0.4798	0.4830	0.0113	2.4%
	93	0.4422	0.4251	0.4266	0.4608	0.4619	0.0109	2.5%
	94	0.4574	0.4374	0.4406	0.4759	0.4801	0.0110	2.4%
	95	0.4813	0.4534	0.4582	0.5047	0.5087	0.0148	3.1%
	96	0.4973	0.4608	0.4651	0.5297	0.5357	0.0195	3.9%
	97	0.4718	0.4458	0.4482	0.4910	0.4929	0.0137	2.9%

Source: Authors' calculations based on the EPH, October 1992-1997.

Figure 1
Gini coefficients and confidence intervals
Greater Buenos Aires and Neuquén, 1992-1997



Source: Authors' calculations based on the EPH, October 1992-1997.

Table 2
Differences in Gini coefficients and bootstrapped confidence intervals and standard errors
Greater Buenos Aires and Neuquén, 1992-1997

Region	Years		Difference	Confidence Interval				Standard Error	
				0.025	0.05	0.95	0.975		
GBA	92	93	-0.0015	-0.0201	-0.0188	0.0172	0.0184	0.0103	
	92	94	-0.0155	-0.0341	-0.0327	0.0021	0.0063	0.0108	
	92	95	-0.0428	-0.0629	-0.0602	-0.0264	-0.0237	0.0103	*
	92	96	-0.0425	-0.0653	-0.0615	-0.0252	-0.0212	0.0112	*
	92	97	-0.0382	-0.0569	-0.0549	-0.0223	-0.0193	0.0099	*
	93	94	-0.0140	-0.0338	-0.0317	0.0028	0.0064	0.0107	
	93	95	-0.0413	-0.0583	-0.0559	-0.0236	-0.0213	0.0104	*
	93	96	-0.0410	-0.0588	-0.0557	-0.0242	-0.0220	0.0100	*
	93	97	-0.0367	-0.0570	-0.0546	-0.0204	-0.0177	0.0108	*
	94	95	-0.0274	-0.0480	-0.0451	-0.0108	-0.0090	0.0106	*
	94	96	-0.0270	-0.0493	-0.0446	-0.0069	-0.0051	0.0111	*
	94	97	-0.0228	-0.0412	-0.0379	-0.0048	-0.0008	0.0100	*
	95	96	0.0004	-0.0204	-0.0187	0.0168	0.0211	0.0113	
	95	97	0.0046	-0.0162	-0.0141	0.0232	0.0258	0.0112	
	96	97	0.0042	-0.0146	-0.0125	0.0199	0.0267	0.0103	
Neuquén	92	93	0.0209	-0.0109	-0.0058	0.0408	0.0448	0.0150	
	92	94	0.0058	-0.0271	-0.0230	0.0338	0.0358	0.0158	
	92	95	-0.0182	-0.0507	-0.0454	0.0067	0.0129	0.0164	
	92	96	-0.0342	-0.0735	-0.0673	-0.0018	0.0073	0.0213	
	92	97	-0.0086	-0.0382	-0.0350	0.0207	0.0264	0.0171	
	93	94	-0.0152	-0.0457	-0.0402	0.0125	0.0158	0.0169	
	93	95	-0.0391	-0.0703	-0.0664	-0.0124	-0.0100	0.0165	*
	93	96	-0.0551	-0.1011	-0.0936	-0.0266	-0.0198	0.0217	*
	93	97	-0.0295	-0.0637	-0.0595	-0.0028	0.0066	0.0168	
	94	95	-0.0239	-0.0576	-0.0499	0.0062	0.0122	0.0165	
	94	96	-0.0399	-0.0943	-0.0824	-0.0006	0.0040	0.0256	
	94	97	-0.0144	-0.0419	-0.0395	0.0166	0.0222	0.0170	
	95	96	-0.0160	-0.0655	-0.0575	0.0186	0.0228	0.0235	
	95	97	0.0096	-0.0294	-0.0213	0.0383	0.0413	0.0190	
	96	97	0.0256	-0.0139	-0.0091	0.0605	0.0644	0.0218	

Note: differences significant at 5% are indicated with *.
Source: Authors' calculations based on the EPH. October 1992-1997.

UNA NOTA SOBRE LA SIGNIFICATIVIDAD ESTADÍSTICA DE LOS CAMBIOS EN LA DESIGUALDAD

WALTER SOSA ESCUDERO y LEONARDO GASPARINI

RESUMEN

Clasificación JEL: D3, D6, C4
Esta nota utiliza técnicas computacionales como el bootstrap para obtener estimaciones por intervalos y errores estándar para las medidas de desigualdad. Adicionalmente, esta metodología es utilizada para implementar un test formal de la hipótesis nula de ausencia de cambios en la desigualdad entre dos períodos. Los resultados se aplican al caso de Argentina, en donde la desigualdad varió sustancialmente en la última década, por lo que resulta crucial distinguir la variabilidad muestral de los verdaderos cambios en la distribución del ingreso. Los resultados muestran que este problema no es menor, dado que para varias regiones de Argentina los cambios observados en el coeficiente de Gini no son estadísticamente significativos.

A NOTE ON THE STATISTICAL SIGNIFICANCE OF CHANGES IN INEQUALITY

WALTER SOSA ESCUDERO and LEONARDO GASPARINI

SUMMARY

JEL Classification: D3, D6, C4
This note illustrates how modern computer intensive tools like the bootstrap provide a simple and efficient way to compute interval estimates and standard errors for inequality measures. Additionally, the same methodology is used to implement a formal test of the null hypothesis of no changes in income inequality between two periods. Results are applied to the case or Argentina, where inequality varied substantially in the last decade, making crucial the issue of distinguishing sampling variability from true changes in the distribution of income. Our results show that the problem is not minor, since the observed changes in the Gini coefficients for several regions in Argentina are not statistically significant.