On Measuring Segregation in Samples With Small Units

William J. Carrington
Charles River Association, Washington, DC 20004

Kenneth R. Troske
Department of Economics, University of Missouri, Columbia, MO 65211 (troske@BPA.Missouri.edu)

Standard indexes of segregation measure a sample's distance from *evenness*, which occurs when each sample unit (e.g., an occupation) has the population share of both the minority and majority groups. We show that random allocation of individuals to units generates substantial unevenness among small units and hence that standard segregation indexes reflect random allocation as well as systematic group segregation. We then modify two popular indexes so that they measure deviations from random allocation rather than deviations from evenness. An empirical example suggests that these modified indexes provide improved measures of the systematic component of group segregation.

KEY WORDS: Random allocation; Segregation; Segregation index.

Economists and other social scientists often assess the extent to which demographic groups are segregated from each other. Examples of such interest include the extent to which blacks and whites reside in the same area (Taeuber and Taeuber 1965; Massey and Denton 1988; Harrison and Weinberg 1992), the extent to which men and women share the same occupations and establishments (Bergmann 1986; Johnson and Solon 1986; Blau 1988; Groshen 1991; Carrington and Troske 1995), and the extent to which high- and low-skill workers share the same firms (Kremer and Maskin 1994). Analysts often summarize these patterns with segregation indexes such as the index of dissimilarity and the Gini coefficient of segregation. These indexes measure the extent to which the distribution of two *groups* (e.g., men and women, blacks and whites) across *units* (e.g., occupations or census tracts) deviates from an *even* distribution in which each group is proportionately represented in each unit. The indexes typically range from 0 to 1, with 0 representing complete evenness and 1 representing complete unevenness in which groups never share the same unit.

An important problem with these indexes is that they can be positive when workers are allocated randomly across units. This occurs for two reasons. First, there is sometimes a simple integer constraint in that each individual must be uniquely allocated to one unit. In a sample with 10 black workers and 20 firms, for example, evenness is unobtainable because it is impossible for each firm to get one half of a black worker. Second, the random allocation of individuals to units will typically generate some deviation from evenness. To see this clearly, consider a large sample of two-person firms that, in aggregate, employ a 50/50 mix of men and women. Random allocation of workers to firms will result in 25% of the firms employing two men, 50% of the firms employing one man and one woman, and 25% of the firms employing two women. Existing segregation indexes would report substantial segregation in this instance.

This is problematic for two reasons. First, analysts often interpret positive segregation indexes as evidence of discrimination or other systematic phenomena, though sometimes the observed patterns are completely consistent with random allocation. Second, samples with high segregation indexes are viewed as more systematically segregated than samples with low index values. Yet such conclusions are valid only if random unevenness is constant across samples. Unfortunately, this is often not the case. In samples with small units and small minority shares, random allocation implies substantial unevenness, and hence substantial segregation as measured by conventional indexes. In contrast, random allocation generates little unevenness in samples with large units and large minority shares. Because unit sizes and minority shares often vary, cross-sample comparisons of conventional indexes sometimes tell us little about the samples' relative degree of systematic segregation.

In an effort to address this problem, this article develops two indexes of systematic segregation that measure a sample's distance from randomness rather than its distance from evenness. Our emphasis on randomness as the appropriate baseline is consistent with the previous work of Blau (1977), Boisso, Hayes, Hirschberg, and Silber (1994), and Ransom (1995), all of whom developed statistical tests of the random-allocation hypothesis rather than the hypothesis of evenness. The contribution of our article is to develop methods that measure a sample's *economic* distance from randomness rather than just its statistical distance. Of course, both economic and statistical distances are important. In conjunction with the test procedures derived by the preceding authors, our indexes of systematic segregation allow a more informative view of empirical segregation patterns. In addition, our indexes impose little computational burden because they require no computations other than those required by the hypothesis tests proposed by Boisso et al. (1994).

Journal of Business & Economic Statistics
October 1997, Vol. 15, No. 4

The article proceeds as follows. Section 1 describes several popular segregation indexes and illustrates their small-unit properties. Section 2 reviews the literature on segregation and points out when small-unit issues are and are not important. Section 3 proposes new methods for measuring and interpreting segregation in small-unit samples. Section 4 then applies the new methods to the measurement of racial and gender workplace segregation. The results suggest that the new indexes produce different, and we believe better, interpretations of the data. More substantively, the results show that racial and gender workplace segregation is more prevalent than a random-allocation model would suggest.

1. RANDOM UNEVENNESS AND CONVENTIONAL SEGREGATION INDEXES

The following discussion will take "random allocation" or "randomness" to imply the urn model of statistical theory. Within units of any given size, the random allocation of a finite population leads to a hypergeometric distribution of minorities (and majorities) across units. In cases in which the population is large relative to each unit, the hypergeometric can be closely approximated by the binomial distribution. The binomial density function is of the form

$$B(m; s, p) = \binom{s}{m} p^m (1-p)^{(s-m)} \quad \text{for all } m = 0, 1, \ldots, s, \qquad (1)$$

where p = the minority's population share, s = unit size, and m = number of minorities in the unit. The binomial predicts that minority shares (m/s) will be distributed across units with a mean of p and a variance of $p(1-p)/s$. This implies that there will be substantial variance in small units' minority share but that the minority-share distribution collapses to p as units get large. This is important because all segregation indexes map interunit dispersion in minority shares into scalar indexes.

We now illustrate these properties in two popular segregation indexes. In interpreting these indexes, it is helpful to think of the *segregation curve* (Duncan and Duncan 1955; Hutchens 1991). If we first sort the units by increasing minority share, then the segregation curve plots the cumulative fraction of majority workers on the x axis and the cumulative fraction of minority workers on the y axis. Figure 1 plots a hypothetical segregation curve along with the 45-degree diagonal of evenness that results from a perfectly even distribution of minorities across units. We now use this construct to interpret two popular segregation indexes.

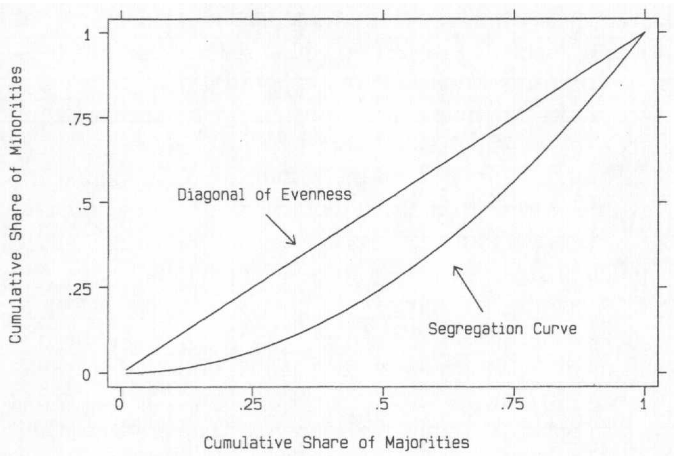

Figure 1. Segregation Curves: An Example.

The index of dissimilarity (also known as the displacement index, the segregation index, or the Duncan index) has long been the most widely used segregation index. If we let N = the total number of individuals in the sample, then the index of dissimilarity is

$$D = \sum \frac{1}{2} \left| \frac{s-m}{N(1-p)} - \frac{m}{Np} \right|, \qquad (2)$$

where the summation is taken over sample units and p, s, and m are now the sample analogs of the previously defined population parameters. If we let $\hat{f}$ denote the empirical joint distribution of s and m across units, then D: $\hat{f} \rightarrow [0, 1]$. The dissimilarity index may be interpreted as the maximum distance between the segregation curve and the diagonal of evenness or as the share of the minority (or majority) group that would have to move across units to achieve an even distribution.

Let $D^*(s, p, n)$ be the expected dissimilarity index implied by the random allocation of a population with minority share p to a sample of n units, each of which is of size s. The inclusion of sample size (n) as an argument in $D^*(s, p, n)$ recognizes the dependence of expected dissimilarity on sample size. Despite this sensitivity to sample size, it is useful to examine the expression for D as n tends to infinity:

$$\lim_{n\rightarrow\infty} D^*(s, p, n) = \sum_{m=0}^{s} \frac{1}{2} B(m; s, p) \left| \frac{(s-m)}{s(1-p)} - \frac{m}{sp} \right|. \qquad (3)$$

When units are small, the binomial heavily weights unrepresentative units, and dissimilarity is high. As unit size increases, however, the binomial puts increasing weight on even units and random dissimilarity falls. The implication is that, for large samples, random dissimilarity is highly dependent on the sample distribution of unit sizes. There is no comparably simple expression for $D^*(s, p, n)$ for finite n, but it is easy to simulate the distribution for fixed n, s, and p. As it turns out, the preceding asymptotic formula appears to be very close to the expectation of finite samples in all cases except those in which p, s, and n are all very small.

The preceding discussion suggests that random allocation implies some unevenness, but it does not indicate how the dissimilarity index maps this unevenness into the [0, 1] interval. To provide guidance on this issue, panel A of Table 1 reports $D^*(s, p, 100)$ for various unit sizes and minority shares. For finite samples, it is of course true that random allocation will generate some variation in dissimilarity across samples. Thus, the numbers in Table 1 are the means of the dissimilarity index computed from 500 randomly allocated samples. The table shows that expected dissimilarity is highly dependent on minority share and unit size. The index's sensitivity to unit size can be seen by looking across

Table 1. Segregation Under Random Allocation

Minority share of population	Number of individuals in each unit: 2	5	10	20	50	100	1,000
Panel A: Mean dissimilarity index under random allocation							
.01	.86	.96	.91	.83	.61	.38	.13
.02	.96	.92	.84	.69	.39	.28	.09
.05	.95	.81	.63	.39	.26	.18	.06
.10	.90	.65	.40	.29	.19	.13	.04
.20	.81	.42	.31	.22	.14	.10	.03
.30	.73	.41	.27	.19	.12	.09	.03
.40	.68	.35	.26	.18	.11	.08	.03
.50	.66	.37	.25	.18	.11	.08	.03
Panel B: Mean Gini coefficient under random allocation							
.01	.99	.96	.92	.84	.68	.52	.18
.02	.98	.92	.85	.73	.53	.39	.13
.05	.95	.83	.69	.54	.35	.25	.08
.10	.90	.72	.55	.41	.26	.18	.06
.20	.80	.59	.43	.31	.20	.14	.04
.30	.73	.53	.38	.27	.17	.12	.04
.40	.68	.49	.36	.25	.16	.11	.04
.50	.66	.49	.35	.25	.16	.11	.04

NOTE: For each combination of unit size and minority share, 500 random samples were drawn in which each sample had 100 units. The dissimilarity index and Gini coefficient were computed separately for each of the 500 samples, and the figures reported above are the means of these indexes. See the text for definitions of the dissimilarity index and the Gini coefficient.

the columns of any given row. For example, if we fix the minority population share at .05, we see that mean random dissimilarity ranges from .95 for 2-person units to .06 for 1,000-person units. The index's sensitivity to minority share can be seen by looking down the rows of any given column. For example, among samples of 20-person units, mean random dissimilarity ranges from .83 to .18 depending on the minority's sample share. In sum, Table 1 clearly shows that random allocation can generate substantial dissimilarity.

Table 1 illustrates the impact of random allocation in samples with uniform unit sizes, but what effect does random allocation have in samples with a mixture of sample sizes? To address this issue, let U_{sm} = the number of sample units with m minorities and s individuals, and let N_s = the number of sample individuals in units of size class s. D can then be expressed as

$$D = \sum_s \frac{N_s}{N} \sum_{m=0}^{s} \frac{1}{2} U_{sm} \left| \frac{(s-m)}{N_s(1-p)} - \frac{m}{N_s p} \right|, \quad (4)$$

where the outer summation is over size classes and the inner summation is over minority shares within size class s. Note that, if the sample minority share p is constant across size classes, then the interior summation is equivalent to the dissimilarity index for any given size class. In this case, therefore, the dissimilarity index is simply a weighted average of the index for each size class, where the weights are each size class's share of sample individuals. More generally, the dissimilarity index is the weighted average of within-size class dissimilarity, plus terms that account for unevenness across size classes. Thus, samples in which small units account for a large size-weighted fraction of the sample will have substantial random dissimilarity.

A second popular index is the Gini coefficient of segregation. If there are T units and if we first sort the units in ascending order of $(s-m)/s$, then the Gini coefficient G: $\hat{f} \rightarrow [0,1]$ can be expressed (Hutchens 1991) as

$$G = 1 - \sum_{i=1}^{T} \left(\frac{s-m}{s} \right) \left(\frac{m}{s} + 2 \sum_{j=i+1}^{T} \frac{m}{s} \right). \quad (5)$$

The Gini coefficient varies between 0 and 1, with 0 again corresponding to complete evenness and 1 representing complete segregation. The Gini coefficient may be interpreted as the area between the segregation curve and the diagonal of evenness, expressed as a proportion of the total area under the diagonal (Massey and Denton 1988). The Gini coefficient of segregation does not lend itself to a decomposition analogous to (4), but panel B of Table 1 shows that the same general points still apply. Random allocation implies relatively small index values for samples with large units and large minority shares because the mean Gini coefficient generated by random allocation is only .04 among samples with 1,000-person units and large minority shares. For small units, however, random allocation predicts a very high Gini coefficient. In a population of 10-person units with a minority share of .10, for example, random allocation implies a Gini coefficient of .55. Similarly, the Gini coefficient is expected to be above .1 even for 1,000-person units when the minority share is .02 or below. A comparison of panels A and B shows that the Gini coefficient is even more sensitive to random allocation than the dissimilarity index.

Table 1 has two important implications. First, analysts often interpret "large" segregation indexes as evidence of discrimination or other systematic phenomena. When units or minority shares are small, however, such conclusions are not always warranted because random allocation implies substantial unevenness. Second, analysts often compare index values across samples or subsamples and interpret higher index values as indicative of a greater role for systematic phenomena. Such comparisons are not generally valid, however, because Table 1 shows that the amount of unevenness generated by random allocation is quite sensitive to unit size and minority share. Thus, cross-sample comparisons of these indexes are misleading unless the samples have similar unit size distributions and minority shares. This requirement is often not met in practice.

Let us conclude this section with three points. First, these same criticisms would seem to apply to any other index of segregation-cum-unevenness. For example, we have computed tables analogous to Table 1 for Atkinson's index and Theil's entropy index, and the results are quite similar (tables are available from the authors on request). Second, note that the crucial issue here is each unit's *sample* size, not its population size. In terms of the unevenness implied by random allocation, a sample of 10 people from each of many large units is equivalent to having the entire population from each of many 10-person units. Finally, note that there are two ways in which a sample can get "large." The first is that the average size of units can increase. The preceding tabulations show that random allocation becomes a nonissue as units get arbitrarily large. The second is that the number of units can increase while the average unit size stays small. Movements in this direction have virtually no effect on the amount of unevenness implied by random allocation. Even a very large number of randomly allocated individuals will not be evenly distributed across units if the size of each unit stays small. Thus, simply increasing the number of sample units is no cure for these problems unless the average unit size also increases.

2. IS RANDOM UNEVENNESS EMPIRICALLY IMPORTANT?

Section 1 shows that random allocation generates substantial unevenness when units or minority shares are small. This section asks whether this is an important issue in the empirical literature. One difficulty with this assessment is that one can often only guess about the random unevenness expected in any given study. The problem is that, although the minority's population share is usually reported, authors rarely report the average unit size, let alone the full distribution of unit sizes required to make a precise statement about how much unevenness would be generated by random allocation. Nevertheless, it is apparent that random-allocation issues arise in a substantial subset of empirical studies of segregation.

Segregation indexes are perhaps most commonly used in studies of residential segregation in which the groups are blacks and whites and the units are census tracts. These tracts average about 4,000 residents, and Table 1 shows that random allocation generates little unevenness in units of this size. The fact that individuals are bunched into households, however, means that there may be substantially fewer than 4,000 independent locational decisions in any census tract. This implies that random unevenness may be substantial in cases in which the minority population is small. In addition, small-unit issues certainly arise when segregation is measured across more narrowly defined groups or geographical units. For example, Denton and Massey (1988) studied white/ethnic residential segregation within samples stratified by schooling, occupation, or income. There are often fewer than 100 individuals per tract within these strata, and our analysis suggests that 15% to 20% of their reported within-group segregation is due to random allocation. Tract-level analysis of minority–minority segregation (e.g., black/hispanic) is similarly sensitive to these issues, and small-unit issues will surely arise when segregation is measured in the block-level data that may soon be available (Harrison and Weinberg 1992).

Studies of gender occupational differences also use segregation indexes. Small-unit issues typically do not arise in aggregate studies of large samples such as the Decennial Census. Many authors, however, consider occupational segregation within much smaller portions of the economy or within much smaller samples. For example, Bielby and Baron (1984) studied occupational segregation within establishments in a sample in which there are often only a few people employed in any particular occupation. Table 1 implies that much of the segregation in their sample is potentially attributable to random allocation. Similarly, Fields and Wolf (1991) studied sex segregation across tens of thousands of industry-by-occupation cells. Even though their samples are quite large, the number of individuals per unit is small enough for random allocation to play an important role. Other studies of occupational segregation in which random allocation appears to be important are those of Wharton (1989), Groshen (1991), and Neuman (1991).

Small-unit issues also arise in studies of segregation across other dimensions. For example, Blau (1977) studied interfirm sex segregation within occupations in cases in which individual firms often employed only a handful of workers. More recently, Carrington and Troske (1995) studied sex segregation in establishments with fewer than 100 employees, and as much as 50% of their reported within-industry segregation can be attributed to random allocation. As another example, Hutchens (1988) studied age-group segregation in a sample of approximately 40,000 workers spread across over 500 industry-by-occupation cells. And as a final example, conclusions about racial segregation in professional sports are often drawn from samples in which the average unit size is less than 50 (see Kahn 1991 for a survey). Our analysis suggests that much of this segregation could be due to random allocation.

Some authors explicitly note the random unevenness generated by small-unit samples, and our crude calculations suggest that the qualitative nature of most conclusions is insensitive to random allocation. We will subsequently document a case, however, in which the traditional approach finds "systematic" segregation in a case that is completely consistent with random allocation, and there are surely other instances of this phenomena in the literature. In addition, even in cases in which the traditional approach appropriately reports systematic segregation, the *quantitative* nature of the departure from randomness is often quite sensitive to small-unit issues. Thus, the precise interpretation of small-unit studies as well as the usefulness of cross-study comparisons depend on the development of new methods that measure departures from randomness rather than evenness. Section 3 takes up this issue.

3. ALTERNATIVE APPROACHES TO MEASURING SEGREGATION

3.1 Preliminaries

This section describes alternative approaches to measuring segregation in small-unit samples. Each of these approaches requires the computation of the baseline index of segregation that would be expected under random allocation. This expectation may be calculated by shuffling the sample—that is, by randomly reallocating sample individuals to sample units, keeping the size of the sample units fixed. Each reshuffling produces a different distribution of minorities and majorities across units, and Gini coefficients and indexes of dissimilarity can be computed from each reshuffling. This in turn generates a range of index values that are consistent with random allocation. The means of these synthetic indexes provide a reasonable measure of what should be expected under random allocation, and we will refer to these as G^* (for the Gini coefficient) and D^*

(for the dissimilarity index). Of course, there will be variation around the mean across the shuffles, but, as Boisso et al. (1994) pointed out, this variation under the null provides a means of testing the hypothesis of random allocation.

3.2 Point Estimates

Our criticism of standard segregation indexes is that they measure the sample's distance from evenness rather than randomness. In this section we propose a set of modified segregation indexes that measure the extent to which a sample deviates from randomness. The advantage of these indexes is that their interpretation does not hinge on the sample share of minorities or on the size of units in the sample. Thus, these modified indexes provide a better means for comparing the extent to which samples are *systematically* segregated. In interpreting these indexes, it will be helpful to refer to Figure 2. Figure 2 is analogous to Figure 1 in that it graphs a hypothetical segregation curve and the diagonal of evenness. However, it also includes what we call the *curve of randomness,* which is the segregation curve that would be expected under random allocation.

Given D and D^* as defined earlier, our modified *index of systematic dissimilarity* $\hat{D}$: $\hat{f} \rightarrow [-1, 1]$ is simply

$$\hat{D} = \begin{cases} \frac{D - D^*}{1 - D^*} & \text{if } D \geq D^* \\ \frac{D - D^*}{D^*} & \text{if } D < D^*. \end{cases} \tag{6}$$

If there is excess unevenness—that is, $D > D^*$—then $\hat{D} > 0$ is simply the extent to which the sample is more dissimilar than random allocation would imply $(D - D^*)$, expressed as a fraction of the maximum amount of such excess dissimilarity that could possibly occur $(1 - D^*)$. $\hat{D} = 1$ is analogous to complete unevenness, as with the standard dissimilarity index, but $\hat{D} = 0$ now implies that the sample is equivalent to random allocation. If there is excess evenness—that is, $D < D^*$—then $\hat{D}$ is negative and represents excess evenness $(D - D^*)$, expressed as a fraction of the maximum amount of excess evenness that could possibly occur (D^*).

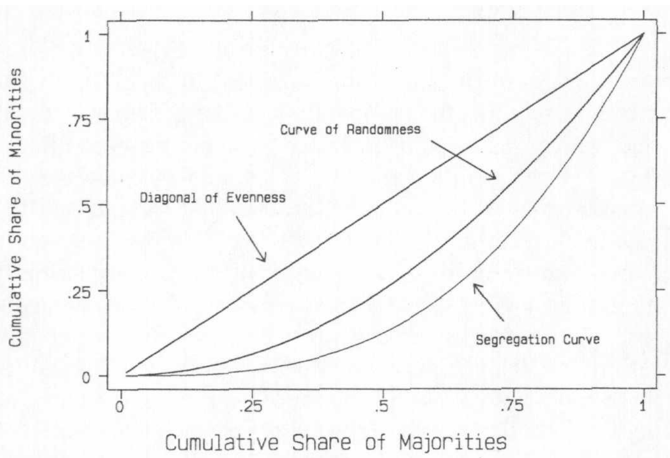

Figure 2. Curve of Randomness.

We also consider an analogously modified Gini coefficient of segregation, $\hat{G}$: $\hat{f} \rightarrow [-1, 1]$:

$$\hat{G} = \begin{cases} \frac{G - G^*}{1 - G^*} & \text{if } G \geq G^* \\ \frac{G - G^*}{G^*} & \text{if } G < G^*. \end{cases} \tag{7}$$

Recall that the standard Gini coefficient is interpretable as the area between the segregation curve and the diagonal of evenness as a fraction of the total area under the diagonal. Our modified Gini coefficient is then interpretable as the area between the segregation curve and the curve of randomness, expressed as a fraction of the total area under the curve of randomness.

Several features of these indexes deserve comment. We couch these points in terms of our modified dissimilarity index, but they apply equally to the modified Gini coefficient. First, $\hat{D} \rightarrow D$ as $D^* \rightarrow 0$. Thus, as unit size gets large, the modified index converges to the original index. This is appropriate because the problem that we are trying to correct goes away as units get large. Second, the range of $\hat{D}$ is $[-1, 1]$ rather than $[0, 1]$. In particular, if there is *more* evenness than predicted by random allocation, then $D < D^*$ and $\hat{D} < 0$. Negative values of $\hat{D}$ simply indicate excess evenness rather than excess segregation. Finally, note that our adjustment of these segregation indexes is similar in spirit to the index of concentration proposed by Ellison and Glaeser (1994).

3.3 The Measurement of Uncertainty

The vast majority of segregation studies report one or more segregation indexes without reporting any hypothesis tests and without any indication as to how much sampling variability might be associated with the estimates. This pattern arose because of the historical lack of analytical solutions for the sampling variability of these indexes. Let us here review some of the methods recently developed in the literature on segregation indexes.

We are aware of two methods for assessing the sampling variability of traditional segregation indexes. First, Ransom (1995) recently developed expressions for the asymptotic sampling distributions of the Gini coefficient and the dissimilarity index. Second, Boisso et al. (1994) applied standard bootstrap methods to the evaluation of segregation indexes. These approaches are, of course, asymptotically equivalent (Efron and Tibshirani 1993), but they will differ to an unknown extent in finite samples. Our empirical example will use the methods of Boisso et al. (1994) because they are more readily adapted to our modified indexes.

The most interesting null hypothesis is that the data are consistent with random allocation. Ransom's method and the bootstrap methods of Boisso et al. (1994) can be extended to test this or virtually any other hypothesis regarding the value of segregation indexes. There are in this case other methods available as well, however. Blau (1977) proposed a chi-squared test of the equality of the empirical and a hypothetical random distribution that is similar to the one from which our D^* is derived. This method is computationally simple, but it is less powerful than the other methods discussed here. Finally, Boisso et al. (1994) developed a randomization test along the lines of Fisher (1935), Edgington (1969), and Noreen (1989). In this test, many samples are generated under the null hypothesis of random allocation, and segregation indexes are computed for each sample. The extent to which the indexes computed from the empirical sample are outliers in the distribution of indexes computed from the hypothetical samples provides a means of testing the hypothesis of random allocation. This approach has the virtue of making use of the hypothetical samples that were previously used in the calculation of D^* and G^*. Thus, the additional computational burden is minimal.

4. AN APPLICATION TO WORKPLACE SEGREGATION

Section 3 proposed two new indexes of systematic segregation. In this section we apply these indexes to the measurement of workplace racial segregation among manufacturing plants in the Chicago metropolitan statistical area (MSA). Although our interest is primarily methodological, this exercise is of substantive interest because taste-based theories of labor-market discrimination (e.g., Becker 1957; Bergmann 1986) imply that workplace segregation is virtually a prerequisite for discrimination to induce group wage differences. As a result, precise measures of interfirm segregation are important for evaluating whether or not current labor-market discrimination is an important source of black/white income differences.

Our sample is drawn from the Worker Establishment Characteristics Database (WECD) more fully described by Troske (1994). The WECD is a sample of manufacturing workers from the 1990 Decennial Census of Population who have been linked to information on their establishment (i.e., their employer) drawn from the Longitudinal Research Datafile. From the current perspective, the important feature of these data is that they tell us which workers work in the same establishment. In the sample ultimately used in the analysis, the unit of observation is an establishment, and the information on each establishment is simply the number of blacks and whites it employs. We restricted the sample to a single MSA to avoid complications associated with the interpretation of cross-MSA segregation. We focused on the Chicago MSA in particular because it is large, well-represented in the WECD, and has a large black population.

Table 2 presents our analysis of racial segregation in these data. The five rows of Table 2 vary by the sample on which the statistics are computed. Row 1 analyzes segregation in the entire WECD sample, and the remaining rows analyze segregation among workers stratified by their educational attainment. In particular, row 2 analyzes segregation among workers with less than a high-school diploma, row 3 workers with a high-school diploma or GED, row 4 workers with some college education, and row 5 workers with a college or advanced degree. Within each row, we present statistics associated with the dissimilarity index (a) and the Gini coefficient (b).

The columns of Table 2 vary by the statistics reported. Column (1) reports the sample value of the traditional indexes. Column (2) reports the 5th and 95th percentile of the distribution of the traditional indexes computed from 500 bootstrap replications. In particular, each bootstrap replication is the segregation statistic computed from a pseudosample generated by drawing with replacement a sample of size N from the original sample of establishments. Thus, this column gives an indication of the sampling variability of the traditional indexes. Column (3) presents analogous statistics based on the distribution of 500 randomly reshuffled samples. The segregation statistics computed from these pseudosamples provide the means for implementing our index, as well as the randomization test proposed by Boisso et al. (1994). Finally, column (4) presents our index of systematic segregation.

Row 1 indicates that there is a substantial amount of unevenness in the distribution of black and white work-

Table 2. Systematic and Random Segregation Indexes: An Application to Interfirm Racial Segregation in Chicago

Sample description	Index	(1) Sample value of traditional index	(2) Segregation statistics generated by bootstrap replications: 5th percentile	95th percentile	(3) Segregation statistics generated by random allocation: 5th percentile	Mean	95th percentile	(4) Index of systematic segregation
1. Entire sample	a. Dissimilarity index	.504	.452	.553	.313	.337	.365	.251
	b. Gini coefficient	.664	.605	.714	.458	.488	.520	.344
2. Workers with less than 12 years of education	a. Dissimilarity index	.688	.633	.757	.513	.558	.615	.294
	b. Gini coefficient	.843	.798	.884	.688	.733	.780	.411
3. Workers with exactly 12 years of education	a. Dissimilarity index	.615	.548	.693	.480	.525	.572	.190
	b. Gini coefficient	.768	.700	.831	.658	.701	.741	.222
4. Workers with between 13 and 15 years of education	a. Dissimilarity index	.587	.514	.671	.501	.540	.585	.101
	b. Gini coefficient	.733	.658	.802	.676	.717	.756	.059
5. Workers with 16 years or more of education	a. Dissimilarity index	.666	.524	.821	.655	.727	.809	−.083
	b. Gini coefficient	.817	.687	.909	.813	.869	.923	−.061

ers among Chicago manufacturing establishments because the dissimilarity index is .504 and the Gini coefficient is .664. The bootstrap statistics of column (2) suggest that the 90% confidence interval for each index has a width of about .1. Thus, the hypothesis of evenness can easily be rejected. Column (3) indicates that a substantial amount of this unevenness is potentially attributable to random allocation. For example, random allocation of sample workers to sample establishments leads to an average Gini coefficient of .488. The 95th percentile of the distribution of Gini coefficients computed from randomly allocated samples (.520), however, lies well below the observed value (.664). Thus, we can readily reject the hypothesis that these data were generated by random allocation as well. This leads to the systematic Gini coefficient of .344 reported in column (4). This implies that actual excess unevenness is 34% of the maximum that could possibly be observed.

In this example, what advantage do our methods offer relative to the existing standard of reporting the traditional index, perhaps in conjunction with the results of a test of the hypothesis of random allocation? In terms of hypothesis testing, we offer very little because we use the test of Boisso et al. (1994). Our methods do lead to an important difference in economic interpretation, however. In particular, the traditional method reports that the dissimilarity index is .504, or 50% of the maximum dissimilarity that could be observed. This strikes the reader as being quite segregated. In contrast, our index reports that *excess* dissimilarity is only 25% (.251) of the maximum *excess* dissimilarity that could be observed. This 25% figure is likely to make a much weaker impression on the reader, and so it should. We strongly believe that excess dissimilarity is the interesting fact to know and therefore that our index provides a more informative description of the sample.

The remaining rows of Table 2 examine segregation within schooling groups. Column (1) indicates that there is substantial unevenness among all schooling groups. For example, the Gini coefficient for workers with less than 12 years of education is .843, and the Gini coefficient for college graduates is .817. Inspection of the Gini coefficients implied by random allocation, however, shows that the rough equality of the traditional indexes masks an important difference between these two groups. The hypothesis of random allocation is easily rejected for high-school dropouts, and the difference between the traditional index (.843) and the average index among randomly allocated samples (.733) is 41% of the maximum such excess unevenness that could possibly occur. Thus, among high-school dropouts, workers are substantially more segregated than random allocation would imply in both an economically and statistically meaningful sense. Contrast this with the results for college graduates. The college graduates' traditional Gini coefficient is similar to that of the high-school dropouts (.817 vs. .843), but for this group the high index is completely consistent with random allocation. Indeed, there is some evidence that college graduates are systematically *integrated*.

The comparison between high-school dropouts and college graduates clearly illustrates the benefits of our methods. Conventional indexes suggest that segregation is equally important within these two groups, but our methods show that systematic segregation is much more pervasive among the dropouts. More substantively, the reduced systematic segregation found among college graduates is consistent with the views of Smith and Welch (1984), who argued that federal affirmative action policies act more strongly on educated workers. In addition, the theories of Becker (1957) and Bergmann (1986) suggest that discrimination-induced group wage differences are usually accompanied by workplace segregation. Thus, these results suggest that current labor-market discrimination may be an important factor in the reduced earnings of black high-school dropouts; it is a less important factor in the black/white earnings gap among college graduates.

The following question may occur to some readers: Why not merely report the traditional index along with the results of the randomization test proposed by Boisso et al. (1994)? The answer is that, although the randomization test provides a useful means of measuring the *statistical* distance between the sample and random allocation, it does not provide a useful means of measuring the *economic* distance between the two. We believe that it is important to measure both distances carefully. Although the preceding example illustrates the real-world usefulness of our approach, a hypothetical example provides even clearer support for our view. Suppose that a very large sample of small units yields a traditional Gini coefficient of .5 and that random allocation implies an average Gini coefficient of .45 for this sample. If the sample is large enough, the hypothesis of random allocation will surely be rejected. Thus, under the prevailing methodology, this sample would be described as having a Gini coefficient of .5 and that random allocation is rejected. Is this a reasonable reading of the data? We think not. The sample is only slightly more uneven than implied by random allocation, and thus the sample is not *economically* different from random allocation in an important way. Our methods make the distinction between economic distance and statistical distance and, as McCloskey and Ziliak (1996) recently emphasized, it is at least as important to measure a sample's economic distance from a point of interest as it is to measure its statistical distance.

5. CONCLUSIONS

The sensitivity of indexes of segregation-cum-unevenness to random allocation was apparently first identified by Duncan and Duncan (1955), and Taeuber and Taeuber (1965) explicitly calculated the level of D implied by random allocation. Thus, it has long been recognized that random allocation implies nonzero segregation in conventional indexes. Yet Cortese, Falk, and Cohen (1976) appear to have made the only previous attempt to address this problem. Their proposal was to "standardize" D by computing

$$Z_D = \frac{D - D^*}{\sqrt{\mathrm{var}(D)}}, \qquad (8)$$

where D^* and $\mathrm{var}(D)$ are computed under the assumption of random allocation. Z_D is best viewed as a test statistic for the hypothesis of random allocation, however, and for this reason Cortese et al.'s index has been largely ignored (Massey 1978). Yet we believe that Cortese et al. were on the right track because it *is* necessary to control for random deviations from evenness. We believe that our indexes do this in a sensible way.

ACKNOWLEDGMENTS

We thank Bruce Hamilton, Robert Hutchens, Douglas Massey, and Finis Welch for helpful comments.

[Received July 1996. Revised October 1996.]

REFERENCES

Becker, G. S. (1957), *The Economics of Discrimination*, Chicago: University of Chicago Press.

Bergmann, B. (1986), *The Economic Emergence of Women*, New York: Basic Books.

Bielby, W., and Baron, J. (1984), "A Woman's Place Is With Other Women: Sex Segregation Within Organizations," in *Sex Segregation in the Workplace: Trends, Explanations, Remedies*, ed. B. Reskin, Washington, DC: National Academy Press, pp. 27–55.

Blau, F. (1977), *Equal Pay in the Office*, Lexington, MA: Lexington Books.

——— (1988), "Occupational Segregation by Gender: A Look at the 1980s," working paper, University of Illinois, Dept. of Economics.

Boisso, D., Hayes, K., Hirschberg, J., and Silber, J. (1994), "Occupational Segregation in the Multidimensional Case," *Journal of Econometrics*, 61, 161–171.

Carrington, W. J., and Troske, K. R. (1995), "Gender Segregation in Small Firms," *Journal of Human Resources*, 39, 503–533.

Cortese, C. F., Falk, R. F., and Cohen, J. K. (1976), "Further Considerations on the Methodological Analysis of Segregation Indices," *American Sociological Review*, 41, 630–637.

Denton, N. A., and Massey, D. S. (1988), "Residential Segregation of Blacks, Hispanics, and Asians by Socioeconomic Status and Generation," *Social Science Quarterly*, 69, 797–817.

Duncan, O. D., and Duncan, B. (1955), "A Methodological Analysis of Segregation Indexes," *American Sociological Review*, 20, 210–217.

Edgington, E. (1969), *Statistical Inference: The Distribution-Free Approach*, New York: McGraw-Hill.

Efron, B., and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, New York: Chapman and Hall.

Ellison, G., and Glaeser, E. L. (1994), "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," Working Paper 4840, National Bureau of Economic Research, Cambridge, MA.

Fields, J., and Wolff, E. N. (1991), "The Decline of Sex Segregation and the Wage Gap, 1970–80," *Journal of Human Resources*, 26, 608–622.

Fisher, R. A. (1935), *The Design of Experiments*, Edinburgh: Oliver and Boyd.

Groshen, E. (1991), "The Structure of the Female/Male Wage Differential," *Journal of Human Resources*, 26, 457–472.

Harrison, R. J., and Weinberg, D. H. (1992), "Racial and Ethnic Residential Segregation in 1990," unpublished paper presented at Population Association of America meetings, Denver, May.

Hutchens, R. M. (1988), "Do Job Opportunities Decline With Age?" *Industrial and Labor Relations Review*, 42, 89–99.

——— (1991), "Segregation Curves, Lorenz Curves, and Inequality in the Distribution of People Across Occupations," *Mathematical Social Sciences*, 21, 31–51.

Johnson, G., and Solon, G. (1986), "Estimate of the Direct Effects of Comparable Worth Policy," *American Economic Review*, 76, 1117–1125.

Kahn, L. M. (1991), "Discrimination in Professional Sports: A Survey of the Literature," *Industrial and Labor Relations Review*, 44, 395–418.

Kremer, M., and Maskin, E. (1994), "Segregation by Skill and the Rise in Inequality," working paper, Harvard University, Dept. of Economics.

Massey, D. S. (1978), "On the Measurement of Segregation as a Random Variable," *American Sociological Review*, 43, 587–590.

Massey, D. S., and Denton, N. A. (1988), "The Dimensions of Residential Segregation," *Social Forces*, 67, 281–315.

McCloskey, D. N., and Ziliak, S. T. (1996), "The Standard Error of Regression," *Journal of Economic Literature*, 34, 97–114.

Neuman, S. (1991), "Occupational Sex Segregation in the Kibbutz: Principles and Practice," *Kyklos*, 44, 203–219.

Noreen, E. W. (1989), *Computer-Intensive Methods for Testing Hypotheses: An Introduction*, New York: Wiley.

Ransom, M. R. (1995), "Sampling Distributions of Segregation Indices," working paper, Brigham Young University, Dept. of Economics.

Smith, J. P., and Welch, F. (1984), "Affirmative Action and Labor Markets," *Journal of Labor Economics*, 2, 269–301.

Taeuber, K. E., and Taeuber, A. F. (1965), *Negroes in Cities*, Chicago: Aldine.

Troske, K. R. (1994), "The Worker-Establishment Characteristics Database," Discussion Paper 95-10, Center for Economic Studies, U.S. Bureau of the Census.

Wharton, A. S. (1989), "Gender Segregation in Private-Sector, Public-Sector, and Self-Employed Occupations, 1950–1981," *Social Science Quarterly*, 70, 923–940.