

# Chapter 5

## Index Differences in Registering Area Group Proportions

My goal in this chapter is to help interested readers become more familiar with the residential outcomes for individuals and households that additively determine the scores of different indices of uneven distribution. To do so, I review the residential outcome scores that underlie segregation comparisons in the difference of means formulation looking in detailed at the segregation comparisons of Whites with Blacks, Latinos, and Asians in Houston, Texas in 2000. The data for these comparisons are taken from block group tabulations for families obtained from Summary File 3 of the 2000 census.<sup>1</sup> Table 5.1 presents the basic demographic information for the four groups and the three segregation comparisons considered here. The results for “overall” percentages document that Whites (non-Hispanic) are the largest group at 52.7% overall, followed by Latinos (34.8%), Blacks (16.5%), and Asians (4.8%). The results also document that the pairwise percentages for any group comparison are always higher than overall percentages for the obvious reason that groups outside the comparison are excluded from the denominator in the calculations.

Table 5.2 lists the values of G, D, R, H, and S obtained using standard computing formulas given in James and Taeuber (1985) for D, G, S and H, and a comparable formula for R adapted from Hutchens (2001) (reviewed in Appendices).

$$D = 100 \cdot \sum t_i |p_i - P| / 2TPQ$$

$$G = 100 \cdot \sum \sum t_i t_j |p_i - p_j| / 2T^2PQ$$

$$S = 100 \cdot \sum t_i (p_i - P)^2 / TPQ$$

---

<sup>1</sup> Specifically, I draw on Table 160 (A-I) which tabulates families by race, poverty status, family structure, and presence of related children under age 18 for block groups. In this tabulation White is Non-Hispanic White, Black and Asian counts include Hispanics, and Latinos are of any race. The information in the tabulations pertaining to social characteristics of poverty and family status are not used here, but are used in analyses presented in Chap. 9.

**Table 5.1** Group counts and overall and pairwise group percentages for Houston, Texas, 2000

Group	N of families	Percentage among all families	White-Black pairwise percentage	White-Latino pairwise percentage	White-Asian pairwise percentage
White	627,613	52.7	76.2	68.0	91.8
Black	195,928	16.4	23.8	–	–
Latino	294,931	24.8	–	32.0	–
Asian	55,746	4.7	–	–	8.2
Total	1,191,102	100.0	100.0	100.0	100.0

Source: US Census 2000, Summary File 3

**Table 5.2** Scores for White-Minority segregation indices obtained using standard computing formulas, Houston Texas, 2000

Group comparison	G	D	R	H	S
Computed using standard formulas					
White-Black segregation	87.07	70.97	47.02	53.59	57.39
White-Latino segregation	74.19	58.37	28.11	35.46	40.96
White-Asian segregation	76.28	58.22	34.96	31.31	23.88

Source: US Census 2000, Summary File 3

$$H = 100 \cdot \sum t_i (E - e_i) / ET$$

$$R = 100 \cdot \left( 1 - \sum t_i \cdot \sqrt{p_i q_i / PQ} / T \right)$$

Terms are defined as noted earlier (and also summarized in Appendices). In this particular analysis, the five segregation indices – G, D, R, H, and S – yield generally similar overall patterns of aggregate segregation between Whites and the three non-White groups. For example, all five indices show that substantial segregation is evident in each comparison. Similarly, all five indices show that White-Black segregation is the highest of the three segregation comparisons examined. There is one notable finding regarding how the different measures portray patterns of aggregate segregation. D, G, and R indicate that White-Latino segregation and White-Asian segregation are roughly similar. H and S indicate that White-Asian segregation is substantially lower than White-Latino segregation.

### 5.1 Segregation as Group Differences in Individual Residential Attainments

I next present results that demonstrate how the scores of the aggregate segregation indices can be obtained from simple differences of group means on residential attainments. Table 5.3 lists the values of D, G, S, H, and R calculated using the difference of means formulations introduced in this monograph. The three panels in

**Table 5.3** Details for obtaining scores for White-Minority segregation from difference of group means on residential outcomes, Houston, Texas, 2000

Residential outcome scored from index-specific scaling function $y=f(p)$	Mean for Whites	Mean for Minority	White-Minority difference
White-Black segregation			
y scored for G/2 ( $\times 100$ ) <sup>a</sup>	60.36	16.82	G=87.08
y scored for G ( $\times 200$ )	120.72	33.65	G=87.07
y scored for D/2 ( $\times 100$ ) <sup>a</sup>	58.44	22.96	D=70.96
y scored for D ( $\times 100$ )	87.73	16.75	D=70.98
y scored for R ( $\times 100$ )	72.98	25.96	R=47.02
y scored for H ( $\times 100$ )	82.57	28.98	H=53.59
y scored for S ( $\times 100$ )	89.86	32.48	S=57.38
White-Latino segregation			
y scored for G/2 ( $\times 100$ ) <sup>a</sup>	61.86	24.76	G=74.20
y scored for G ( $\times 200$ )	123.72	49.53	G=74.19
y scored for D/2 ( $\times 100$ ) <sup>a</sup>	59.33	30.14	D=58.38
y scored for D ( $\times 100$ )	81.49	23.12	D=58.37
y scored for R ( $\times 100$ )	63.37	35.26	R=28.11
y scored for H ( $\times 100$ )	72.95	37.50	H=35.45
y scored for S ( $\times 100$ )	81.12	40.17	S=40.95
White-Asian segregation			
y scored for G/2 ( $\times 100$ ) <sup>a</sup>	53.11	14.97	G=76.28
y scored for G ( $\times 200$ )	106.22	29.94	G=76.28
y scored for D/2 ( $\times 100$ ) <sup>a</sup>	52.38	23.27	D=58.22
y scored for D ( $\times 100$ )	75.15	16.93	D=58.22
y scored for R ( $\times 100$ )	70.70	35.74	R=34.96
y scored for H ( $\times 100$ )	83.47	52.16	H=31.31
y scored for S ( $\times 100$ )	93.79	69.91	S=23.88

Source: US Census 2000, Summary File 3

<sup>a</sup>For these scorings of y for G and D, the values of G and D are given by  $2 \cdot (Y_1 - Y_2)$

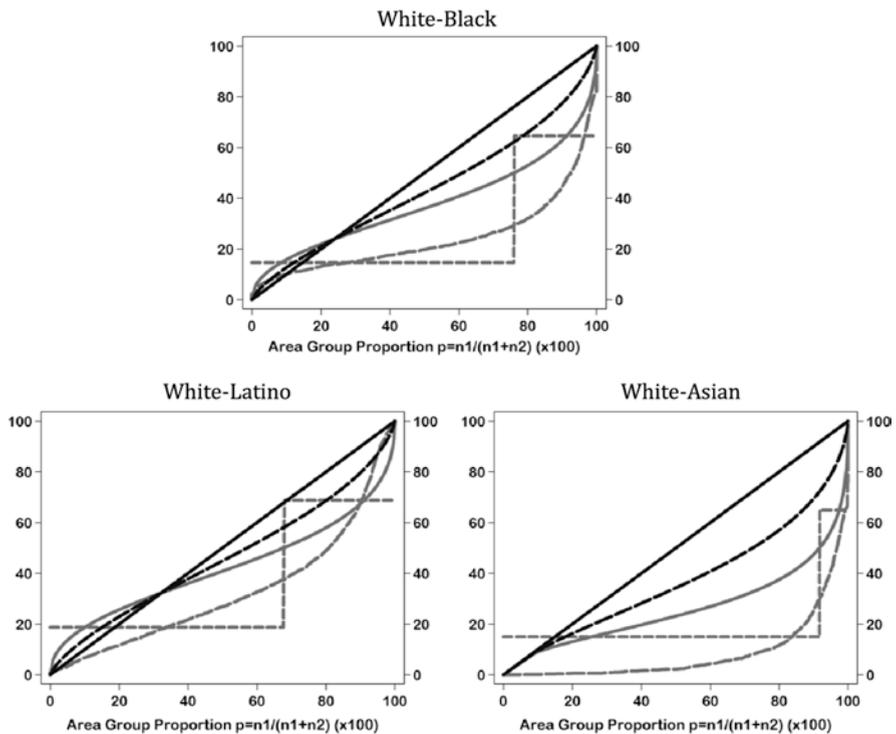
the table report results separately for the White-Black, White-Latino and White-Asian segregation comparisons, respectively. The first step in generating these results is to calculate the residential outcomes scores (y) at the block group level. I obtain these by applying the relevant index-specific scaling function  $y = f(p)$  to the value of pairwise proportion White (p) at the block group level. The second step is to calculate the group-specific means for scaled contact with Whites (y). The resulting values are reported in Table 5.3. The last step is to calculate the difference of the group-specific means which also are reported in Table 5.3.

The results are straightforward. The values of the differences of means equal the values of the index scores reported in Table 5.2. Any apparent differences reflect only rounding error and would disappear if the results were reported to greater precision. Of course, the index scores reported in Table 5.3 are redundant with the results already presented and do not themselves provide any new insights into

segregation patterns. But presenting the detailed results documents that the difference-of-means formulas yield the same results as the conventional formulas.

I noted in the previous chapter that the segregation-relevant residential outcomes ( $y$ ) that determine the group means ( $Y_1$  and  $Y_2$ ) are index-specific scores for scaled pairwise contact with Whites. The exact scoring of residential outcomes ( $y$ ) varies from index to index, but values of  $y$  always are a positive, monotonic function of pairwise proportion White ( $p$ ) for the household's or individual's area of residence. The minority group's average pairwise contact with Whites cannot exceed that observed for Whites and it can reach parity only under the condition of exact even distribution. When there is departure from uneven distribution, mean contact with Whites for Whites will diverge from mean contact with Whites for the minority group. The average magnitude of the difference will be reflected in the difference of means ( $Y_1 - Y_2$ ) which will yield the index score. Given this, it is instructive to consider how the different index-specific residential outcome scores compare to each other.

Figure 5.1 plots the values of residential attainment scores ( $y$ ) by values of pairwise proportion White ( $p$ ) for G, D, R, H, and S for the three White-Minority segre-



**Fig. 5.1** Scoring residential outcomes ( $y$ ) from pairwise proportion White ( $p$ ) to compute G, D, R, H, and S as a difference of means. Legend for index-specific curves:  $y$  scored for  $G/2$  – gray, long dashes;  $y$  scored for  $D/2$  – gray, short dashes;  $y$  scored for  $R$  – gray, solid line;  $y$  scored for  $H$  – dark, long dashes;  $y$  scored for  $S$  – dark, solid line

gations comparison presented in Table 5.3. These plots provide a basis for gaining insight into how each segregation index registers residential contact outcomes ( $y$ ) based on pairwise area racial mix ( $p$ ). I begin with the scores for the separation index ( $S$ ) because they are the easiest to describe. The scaling function  $y = f(p)$  for  $S$  maps  $y$  directly to the values of  $p$  producing a diagonal line rising from (0,0) to (100,100) in all three graphs. As a result, it is very easy to interpret the relationship between  $y$  and contact with Whites ( $p$ ); a one-point change in contact with Whites translates into a one-point change in  $y$ . Thus, the graph for the White-Black comparison indicates that a Black family that moves from a 20% White area to a 70% White area would experience an increase of 50 points on scaled contact with Whites. The graph for the White-Latino comparison shows that the same would be true for a Latino family moving from a 20% White area to a 70% White area and the graph for the White-Asian comparison shows that the same would be true for an Asian family moving from a 20% White area to a 70% White area. This similarity of change in  $y$  by change in  $p$  is not observed for the other indices because their scaling functions are nonlinear.

The scaling function  $y = f(p)$  for the Theil index ( $H$ ) converts values of  $p$  to values of  $y$  that fall on a smooth, ever-rising, backwards “S-curve”. In these graphs the departure from nonlinearity is not dramatic, especially in comparison to what will be seen for some other indices. Accordingly, the values of residential attainment scores ( $y$ ) relevant for  $H$  tend to be relatively close to residential attainment scores ( $y$ ) relevant for  $S$ . This provides a new insight to why scores for the separation index ( $S$ ) tend to correlate more closely with the scores of the Theil Index ( $H$ ) than with the scores of other indices. Looking across the three segregation comparisons one can see that the nonlinearity is most pronounced in the White-Asian comparison and least pronounced in the White-Latino comparison. This is because nonlinearity in the  $y$ - $p$  relationship for residential attainment scores ( $y$ ) relevant for  $H$  will be less pronounced when the two groups in the comparison are more equal in size and more pronounced when one group is substantially larger than the other. As a result, residential outcomes scores ( $y$ ) for  $H$  and  $S$  tend to track each other more closely when the two groups in the comparison are comparable in size and less closely when the groups are unequal in size.

The nonlinearity in the  $y$ - $p$  relationship for  $H$  just described has another implication. It means that a change of a fixed amount in contact with Whites ( $p$ ) will translate in different amounts of change in  $y$  for  $H$  depending on two factors; the initial starting value of  $p$  and relative size of the two groups. Thus, inspection of the three graphs in Fig. 5.1 indicates that a family that moves from an area that is 20% White area to an area that is 70% White area would experience an increase of 35.9 points on scaled contact with Whites ( $y$ ) in the White-Black comparison, 36.9 points in the White-Latino comparison, and 31.8 points in the White-Asian comparison. The change in scaled contact for the White-Asian comparison is smallest because the White-Asian group size comparison is the most imbalanced. This leads to greater nonlinearity in the  $y$ - $p$  relationship and smaller changes in  $y$  when moving from 20 to 70 on  $p$ . In contrast, the White-Latino group size comparison is the most balanced

of the three and leads to milder nonlinearity in the  $y$ - $p$  relationship and larger changes in  $y$  as  $p$  moves from 20 to 70.

In each group comparison the changes in  $y$  as  $p$  moves from 20 to 70 are smaller than the 50 point increase in  $y$  observed for  $S$  for the same group comparisons. This is because the  $y$ - $p$  relationship is linear for  $S$  and nonlinear for  $H$ . The nonlinearity in the  $y$ - $p$  relationship for  $H$  creates a large region in the middle portion of the range of  $p$  where the slope of the curve is less than 1.0 and thus changes in  $y$  are smaller than changes in  $p$ .<sup>2</sup> In addition, the degree to which changes in  $y$  are smaller than changes in  $p$  varies across the three segregation comparisons because the nonlinearity in the  $y$ - $p$  relationship varies; specifically, the departure from linearity is more pronounced when the two groups in the comparison are more unequal in size and thus changes in  $y$  over the middle range of  $p$  are smaller in these group comparisons.

The function  $y = f(p)$  for the Hutchens index ( $R$ ) also generates values of  $y$  that fall on a smooth, ever-rising, backwards “S-curve”. The curve is similar in form to the curve seen for the Theil index ( $H$ ). But the nonlinearity in the curve for  $R$  is noticeably more pronounced. Accordingly, the patterns for the scoring of  $y$  for  $R$  are similar to those just noted for  $H$ , but “amplified”. For example, as with  $H$ , changes of a fixed amount in contact with Whites ( $p$ ) translate into different impacts on  $y$  depending on the initial starting value of  $p$  and relative size of the two groups. Thus, the graphs in Fig. 5.1 indicate that a family that moves from an area that is 20% White area to an area that is 70% White area would experience an increase of 24.2 points on scaled contact with Whites ( $y$ ) in the White-Black comparison, 25.7 points in the White-Latino comparison and 18.3 points in the White-Asian comparison. The changes in  $y$  are even smaller than the changes in  $y$  noted for  $H$  because the departure from linearity in the  $y$ - $p$  relationship for  $R$  is greater. This “flattens” the  $y$ - $p$  curve over the middle range of  $p$  even more and causes changes in  $y$  to be smaller than changes in  $p$ . As seen with  $H$ , the changes in  $y$  vary across the different segregation comparisons; they are larger when groups are more equal in size and smaller when groups are more unequal in size.

The function  $y = f(p)$  for the gini index ( $G/2$ ) also produces an ever-rising, backwards “S-curve”. However, in contrast to the functions for  $H$  and  $R$ , this curve is irregular rather than smooth. This is because  $G$  tracks percentile scores for  $p$  and these depend not on the specific value of contact with Whites ( $p$ ) itself, but instead on how values of  $p$  translate into rank position on contact with Whites. In the case of White-Black segregation, for example, this is determined by the number of Whites and Blacks living in areas where  $p$  higher and the number of Whites and Blacks living in areas where  $p$  is lower. The nonlinearity of the function for  $G/2$  is more pronounced than that seen for the functions for  $H$  and  $R$  and this produces larger departures from the diagonal line for  $S$ . As a result, it is reasonable to say that scoring  $y$  as the percentile transformation of  $p$  is as the most “dramatic” rescaling of contact of those considered here. Thus, the graphs in Fig. 5.1 indicate that a family

---

<sup>2</sup>For example, from inspection of the figure for the White-Black comparison changes in  $y$  are smaller than changes in  $p$  for the portion of the curve between approximately 15 and 85 on  $p$ .

that moves from an area that is 20% White area to an area that is 70% White area would experience an increase of 13.2 points on scaled contact with Whites ( $y$ ) in the White-Black comparison, 27.7 points in the White-Latino comparison, and 6.4 points in the White-Asian comparison. In each case, the changes in  $y$  are even smaller than the changes in  $y$  seen for H and R because the pronounced nonlinearity in the  $y$ - $p$  relationship for G “flattens” the  $y$ - $p$  curve over the middle range of  $p$  quite dramatically causing changes in  $y$  to be much smaller than changes in  $p$ . As observed previously for H and R, the changes in  $y$  vary across the different segregation comparisons with changes being larger when groups are more similar in size and smaller when are more unequal in size. Thus, the change in  $y$  for the White-Latino comparison, where the two groups are more similar in size, is more than four times larger than the change in  $y$  for the White-Asian comparison where the two groups are more unequal in size.

In contrast to S, H, R, and G, the scoring of  $y$  for the index of dissimilarity (D) is not ever-rising as  $p$  increases. Instead, it follows a simple, two-value, monotonic step function. The scoring of  $y$  for  $D/2$  shown in the graphs draws on the formulation of D as a version of G computed from a two-category ranking scheme with areas where  $p \geq P$  being in the higher ranking category and all other areas being in the lower ranked category. For example, in the White-Black comparison,  $y$  is scored 14.6 when  $p < P$  and 64.6 when  $p \geq P$ .<sup>3</sup> The scoring of  $y$  for D could alternatively be shown as a step function where values of  $y$  are either at 0 or 100 depending on whether  $p$  is above  $P$  or not. But I present the  $D/2$  formulation here to facilitate the comparison of D with G.

The step function for  $D/2$  produces a rescaling of contact that responds to changes in  $p$  only when  $p$  crosses from being below  $P$  to equaling or exceed it. As the graph in Fig. 5.1 indicates, this does not occur when a family moves from an area that is 20% White area to an area that is 70% White area in the White-Black comparison. So a family making this move would experience no change in scaled contact with Whites ( $y$ );  $y$  is 14.6 when  $p$  is 20 and  $y$  remains at this value when  $p$  is 70. The same is true in the White-Asian comparison. In contrast, the change in  $y$  for a family making a comparable move in the White-Latino comparison would be 50.0 points (the maximum possible change under the  $D/2$  formulation).

These results highlight two things about D. They highlight that D responds to changes in  $p$  only when  $p$  crosses a specific value and otherwise D is insensitive to changes in  $p$ . The examples also highlight that the value of  $p$  that D responds to differs from one segregation comparison to another based on group size. Thus, when groups are equal in size, D responds to changes in  $p$  at 50% White and when the minority group is smaller in size, D responds to changes in  $p$  at increasingly higher levels. Thus, the 50 point change in  $y$  occurs when  $p$  crosses from below to above

---

<sup>3</sup>The value of 14.6 indicates that the areas in the lower ranking category contain 29.2% of the households in the analysis and thus have an average percentile score of 14.6. The value of 64.6 is based on the average percentile score for the 70.8% of households that are in the higher ranking category; that is,  $29.2 + 70.8 / 2$ .

68.0 in the White-Latino comparison, from below to above 76.2 in the White-Black comparison, and from below to above 91.8 in the White-Asian comparison.

## 5.2 Implications for Sensitivity to Separation and Polarization

The patterns just reviewed provide an intuitive basis for comparing indices of uneven distribution and placing them on a continuum. One end of the continuum is anchored by the separation index ( $S$ ). The  $y$ - $p$  relationship for  $S$  is linear. So it registers group differences in pairwise contact ( $p$ ) in its original metric. This is well-suited for measuring group separation and neighborhood polarization. If the group means on  $p$  differ by a large amount, it follows that groups live apart from each other with members of each group living in neighborhoods where their group predominates. If the group means on  $p$  are similar, it follows that the groups live together, not apart, and thus share similar neighborhood outcomes on pairwise racial mix ( $p$ ).

The other end of the continuum is anchored by the gini index ( $G$ ). The  $y$ - $p$  relationship for  $G$  is profoundly nonlinear. This is because it does not register group differences in pairwise contact ( $p$ ) in its original metric. Instead, the scoring function instead converts the level of actual contact into a score for rank order position via the percentile transformation. This is well-suited for measuring ordinal differences in group contact. But it is ill-suited for measuring group separation and neighborhood polarization. Accordingly, if group means on percentile scores ( $y$ ) based on pairwise group contact ( $p$ ) differ by a large amount in White-Minority comparisons, one can safely conclude that Whites consistently live in neighborhoods that rank higher on proportion White than do minorities. But, one cannot conclude that the minority group lives apart from Whites in neighborhoods where the minority group predominates. This is because percentile scores logically cannot provide reliable signals about underlying quantitative differences. As a result, percentile scoring of pairwise group contact cannot provide a reliable basis for assessing group residential separation and neighborhood polarization.

This is not an esoteric point. I will present empirical analyses in the next chapter that demonstrate that high scores on  $G$  can and do occur when group residential separation and neighborhood polarization is low, and in some cases even trivial. Ultimately, researchers should decide for themselves if they view this quality of  $G$  as desirable, undesirable, or irrelevant. But to decide, they first must become aware that  $G$  has this quality. In the main they are not aware and this is understandable because the issue receives little attention in methodological discussions in the literature. As a consequence, no one has set forth a well-articulated rationale for prioritizing group differences in rank order position on contact over the group differences in quantitative “raw score” standing on contact.

The remaining three indices of uneven distribution considered here – the index of dissimilarity ( $D$ ), the Hutchens square root index ( $R$ ), and the Theil entropy index

(H) – fall in intermediate positions on the continuum between the gini index (G) and the separation index (S). Not surprisingly, D is closest to G. R and H fall in between with R closer to D and H close to S. The basis for this ordering is suggested by the  $y$ - $p$  relationships for the indices depicted in the graphs in Fig. 5.1. G is at the opposite end of the continuum from S because its  $y$ - $p$  relationship is most profoundly nonlinear – resulting due to the fact that the percentile scoring of  $y$  from  $p$  often produces scores for  $y$  that depart dramatically from the original value of  $p$ . The dissimilarity index (D) is closest to the gini index (G) because D can be understood as a crude version of G based on a two-category ranking scheme. This is indicated visually by the fact that the step-function “curve” for the  $y$ - $p$  relationship for D overlays the “finer-grained” steps in the  $y$ - $p$  curve for G seen in the figures.

The Hutchens square root index (R) falls near the index of dissimilarity (D) based on the fact that the  $y$ - $p$  curve for R is closer to linear than the  $y$ - $p$  curve for G but is more nonlinear than the  $y$ - $p$  curve for the Theil entropy index (H). Perhaps this should not be surprising since Hutchens (2001) notes that R has the quality of ranking segregation comparisons in accord with the principle of segregation curve dominance. Since the segregation curve is a graphical depiction of rank order differences, it makes sense that R is more sensitive to group differences in rank order standing on group contact than to group differences in quantitative standing on contact.

The  $y$ - $p$  relationship for the Theil entropy index (H) displays only mild departure from linearity and thus produces curves that align more closely with the linear  $y$ - $p$  curve for the separation index (S). On this basis, one can infer that H is more sensitive to group residential separation and neighborhood polarization than every other index except S.

Each of the indices of uneven distribution considered here – G, D, R, H, and S – have been endorsed in methodological studies.<sup>4</sup> And each has been adopted by researchers who have seen the index as having qualities that are attractive for the purposes of the studies they were undertaking. The discussion here provides one additional basis for choosing among indices – sensitivity to group differences in rank order standing on group contact or sensitivity to group differences in contact measured in its “natural” metric. This can also be cast in terms of sensitivity to group residential separation and neighborhood polarization because this follows differences in actual contact, not differences in rank order position on contact.

If one is interested in identifying “prototypical segregation” as seen in traditional exemplars such as White-Black segregation in Chicago and White-Latino segregation in Los Angeles, the separation index (S) is a logical choice and the Theil entropy index (H) would be the next best choice. The basis for choosing S is this.

---

<sup>4</sup>Of these, D is viewed as most problematic on technical grounds, but it usually receives a “conditional pass” because its technical deficiencies often do not have important practical consequences. For example, in empirical studies it typically correlates very closely with G, its close and technically superior cousin.

*High values on S always signal a high level of group residential separation and neighborhood polarization of the kind featured in didactic discussions of examples of pronounced segregation.*

This is not strictly the case for high values on H. But the relatively close relationship of y scored for H with y scored for S (i.e., with p), dictates that high scores on H are very likely, albeit not necessarily guaranteed, to involve a high level of group separation and neighborhood polarization. In contrast, the other three indices – R, D, and G – are not reliable in signaling the presence of prototypical segregation that involves group separation and neighborhood polarization.

If one is interested in identifying segregation assessed strictly on rank-order standing on group contact (p) as registered by the segregation curve, S and H are not good choices. The gini index (G) and the Hutchens square root index (R) would be the superior choices on technical grounds and the dissimilarity index (D) would be an attractive choice based on past usage, ease of computation and interpretation, and related practical considerations.

A few simple questions can help frame the issues researchers confront when they choose to give priority to one index over others. One is “Do the theories and substantive concerns motivating analysis of segregation lead one to naturally focus on prototypical segregation which involves substantial area racial polarization and clear group differences in quantitative levels of contact or do they lead one to instead focus on group differences in rank order standing on contact?” If the substantive focus is on rank order standing, one should be able to explain why high scores of 76.3 and 58.2 on G and D, respectively, for the White-Asian comparison are sociologically important in light of the low score of 23.9 on S. The low score on S, as well as the component group means on contact with Whites that determine it, document that White-Asian segregation in Houston is not “prototypical” segregation. White-Asian segregation does not involve substantial group separation and neighborhood polarization; Asians are more than twice as likely to live with Whites (mean pairwise contact is 69.9%) as with Asians (mean pairwise contact is 31.1%).

In contrast, G and D for White-Latino segregation – at 74.2 and 58.4, respectively – take values comparable to those observed for White-Asian segregation, but White-Latino segregation is more in keeping with prototypical segregation. In contrast to Asians, Latinos are much less likely to live with Whites; Latino pairwise contact with Whites is only 40.2% while Latino pairwise contact with Latinos is 59.8%. As a result, the score of 40.1 on S for White-Latino segregation indicates that group separation and neighborhood polarization is nearly twice as high in the White-Latino comparison as in the White-Asian comparison. Similarly, G, D, and S are 87.1, 71.0, and 57.4, respectively, for the White-Black comparison. The values of G and D are only 10.8 and 13.4 points higher, respectively, than the values observed for the White-Asian comparison. But the value of S is some 33.5 points higher and is more than double the value of S for the White-Asian comparison. The component terms of S for the White-Black comparison indicate clearly that this is “prototypical” segregation involving substantial group separation and neighborhood racial polarization. Consistent with this, *both* Whites and Blacks live apart in

neighborhoods where their group predominates. White pairwise contact with Whites is 89.9% and Black pairwise contact with Blacks is 68.5%. The level of same group contact for Blacks is more than double the level of 31.1% seen for Asians. In sum, G and D suggest that all three segregation comparisons are fairly similar. S suggests White-Asian segregation is distinctively different from White-Latino and especially White-Black segregation.

Figure 5.1 clarifies why G and D yield high scores for White-Asian segregation when S does not. It is because G and D assign great importance to group differences on  $p$  that have minimal impact on S because they are quantitatively small. S takes a relatively low value of 23.8 because Asian pairwise contact with Whites, while not reaching the level of 93.8% seen for Whites, is nevertheless quite high at 69.9%. To calculate G, values of  $p$  are converted to percentile scores and the group difference is then doubled.<sup>5</sup> While the group means for  $p$  do not necessarily map exactly to the group means for percentile scores (because the percentile transformation is nonlinear), it is instructive to note that the values of 93.8 and 69.9 for  $p$  translate to percentile score values of 36.3 and 6.8, respectively. Taking twice the difference to obtain the implications for G yields the value of 59.0. Thus, the initial modest difference on  $p$  that produces a value of 23.8 points for S translates to an implied difference of 59.0 points for G. This is actually less than the observed value of G of 76.3 which means that the exaggeration of group differences on  $p$  is consistently larger than this particular calculation suggests.

Applying this same exercise to the group difference of medians also is “instructive.” The group medians on  $p$  are 97.5 for Whites and 76.7 for Asians. This yields a group difference at the medians of 20.8 (which is close to the difference in group means of 23.8). These values of  $p$  translate to 53.9 and 9.7, respectively, when converted to percentile scores. When this difference is doubled to obtain the implications for G specified as a difference of group medians, the result is 88.4. So the original quantitative difference in “typical” residential outcome of 20.8 when  $p$  is measured in its original metric grows to more than four times that size when  $p$  is rescaled by the percentile transformation curve shown in Fig. 5.1.

A similar pattern is observed when values of  $p$  are converted from their original metric to the 0 or 100 scoring scheme used for D. The values of 69.9 and 76.7, which represent the mean and median, respectively, for Asians on  $p$  become 0.0. In contrast, the values of 93.8 and 97.5, which represent the mean and median, respectively, for Whites on  $p$  become 100.0. Thus, the original group differences at these points of comparison – 23.8 points at the group means and 20.8 points at the group medians, expand to the maximum possible difference of 100.0.

The point to take away is simple, but important. The rescaling of  $p$  from its original metric, which determines S, to the scaled contact scores for  $y$  that determine G and D serves to exaggerate small quantitative differences on  $p$ . Accordingly, values of G and D are usually larger and are never smaller than values of S.<sup>6</sup> Furthermore,

<sup>5</sup>This is because the maximum possible group difference on percentile scores is 50.

<sup>6</sup>Additionally, since D is a crude version of G based on a three-point segregation curve instead of the full segregation curve, G is almost always higher and is never lower than D.

the degree to which the rescaling exaggerates quantitative differences on  $p$  is greater when groups are unequal in size as seen in the White-Asian comparison. Accordingly, the G-S and D-S discrepancies can be especially large in such comparisons.

This raises the question, “Why is it appropriate to score  $y$  in a way that dramatically amplifies group differences in contact with Whites as observed in this example?” Relatedly, “In what way is the exaggerated difference of 59.1 points on  $y$  scored for G and 100.0 points for  $y$  scored for D more sociological meaningful than the smaller difference of 23.8 points for  $y$  scored for S?” Perhaps compelling answers to these questions can be given. For now, however, the measurement literature does not provide a ready answer and I am skeptical that a compelling answer can be advanced. Regardless, it will remain the case that in these segregation comparisons examining S and its component terms reveals important information that would be missed if one looked only at G and D. Specifically, S documents that White-Asian segregation does not involve group residential separation and neighborhood polarization whereas White-Latino segregation and especially White-Black segregation do. The practical implication is straightforward; one cannot safely assume that high values of G and D indicate a prototypical pattern of segregation. One must also examine S to draw a safe conclusion on this issue.

## References

- Hutchens, R. (2001). Numerical measures of segregation and their desirable properties. *Mathematical Social Sciences*, 42, 13–29.
- James, D., & Taeuber, K. (1985). Measures of segregation. *Sociological Methodology*, 13, 1–32.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 2.5 International License (<http://creativecommons.org/licenses/by-nc/2.5/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

