Richness Curves for Evaluating Market Segmentation
Author(s): Thomas P. Novak, Jan de Leeuw, Bruce MacEvoy
Source: *Journal of Marketing Research*, Vol. 29, No. 2 (May, 1992), pp. 254-267
Published by: American Marketing Association
Stable URL: http://www.jstor.org/stable/3172574
Accessed: 23/04/2009 20:43

THOMAS P. NOVAK, JAN de LEEUW, and BRUCE MacEVOY*

Overall evaluation of market segmentation often is based on variance-explained measures, such as R^2. However, variance explained measures can be misleading if applied to the managerial problem of selecting target segments. Richness curves are proposed as an alternative way of evaluating market segmentation. Statistical considerations (bias and stability) are addressed.

Richness Curves for Evaluating Market Segmentation

The managerial utility of market segmentation schemes typically is evaluated on the basis of the degree to which the scheme maximizes within-segment homogeneity and among-segment heterogeneity (e.g., Frank, Massy, and Wind 1972; Kotler 1988). In the marketing literature, R^2 commonly is used to evaluate market segmentation because it assesses this property (see, e.g., Belk 1974; Frank, Massy, and Wind 1972; Green 1973; Kahle, Beatty, and Homer 1986; Kamakura and Mazzon 1991; Kamakura and Novak 1992; Lutz and Kakkar 1974; Novak and MacEvoy 1990; Rosekrans 1969; Sawyer and Ball 1981). However, though variance explained measures such as R^2 do provide an overall index of differences among market segments, they do not address the managerial issue of selecting target segments on which to focus marketing activity.

We propose a method for evaluating market segmentation for the strategic purpose of forming target segments from a current segmentation scheme. The target segment problem implies an *asymmetric* evaluation. For example, an attractive target segment could be constructed by combining segments containing high proportions of product users. Segments containing nonusers are simply not of interest in constructing the target. In contrast, variance explained measures, such as R^2, provide a *symmetric* evaluation because both users (positive deviations about the mean) and nonusers (negative deviations about the mean) contribute to variance explained. Though they are useful for quantifying the degree to which segments differ from each other, symmetric measures of evaluation are not useful for quantifying the degree to which target segments can be identified.

We propose a method of evaluation based on the concept of "richness" (Christen 1987; MacEvoy 1989). In a marketing context, richness is simply the proportion of individuals in a market segment who are "consumers." If a segmentation scheme can be used effectively to target consumers, the richness of a target segment will be substantially higher than the richness of the unsegmented market taken as a whole.

Because segmentation schemes divide the population into several segments, the evaluative criterion must combine information about the richness of each segment into information about the scheme as a whole. This is done by means of a richness curve, which is computed as a running average of the richness of each segment in the scheme as segments are added in descending rank order of segment richness. Richness curves, and statistics derived from them, form a simple and direct way of evaluating segmentation schemes. However, because richness curves are based on a rank ordering of sample values of segment richness, the front end of the richness curve will tend to be loaded with a positive bias due to sampling error that inflates the estimated richness. Therefore, statistical adjustment is necessary, and we describe how it can be accomplished using bootstrap methodology.

In evaluating market segmentation, we primarily consider consumption behaviors, broadly defined in terms of product or media use, frequency of purchase, dollars

*Thomas P. Novak is Assistant Professor, Edwin L. Cox School of Business, Southern Methodist University. Jan de Leeuw is Professor, Departments of Psychology and Mathematics, UCLA. Bruce MacEvoy is Senior Research Psychologist, SRI International.

The authors thank Lee Cooper, Donna Hoffman, Don Lehmann, Joel Steckel, three anonymous *JMR* reviewers, and the editor for their comments, Jin Han for computational assistance, and SRI International for providing the empirical data.

Journal of Marketing Research
Vol. XXIX (May 1992), 254–67

spent, and other factors, but many other measures of consumer behavior could be considered as well. In particular, we consider product usage behaviors, which are commonly used as criterion variables for evaluating segmentation schemes (e.g., Andreason 1966; Assael and Roscoe 1976; Bass, Tigert, and Lonsdale 1968; Frank, Massy, and Boyd 1967; Grover and Srinivasan 1987; Henry 1976; Kahle, Beatty, and Homer 1986; Kamakura and Mazzon 1991; Novak and MacEvoy 1990; Schaninger 1981).

We recognize that there are limitations to focusing on segment differences in product usage. Establishment of differences in consumption is a necessary, but not sufficient, condition for a complete evaluation of market segmentation. If the marketing problem is to increase the size of the market, directing resources to the segment containing most of the users may not be profitable because that segment may be close to its potential. In addition, if the segments containing the highest proportion of users are difficult to identify or access, targeting those segments will not be profitable. Numerous researchers have cautioned against defining target segments on the basis of consumption differences without considering response to marketing mix variables (Blattberg and Sen 1974; McCann 1974; Winter 1984). Because usage rates and segment sizes can be adjusted for acccessibility, response to marketing mix, or identifiability, many of these limitations can be overcome, but we do not explicitly address them.

In the following sections we define and discuss the richness curve and compare it with the closely related Lorenz curve (Singer 1968). Bias and stability in the richness curve are addressed and an empirical example is used to illustrate the application of richness curves.

RICHNESS CURVES

An Informal Definition of the Richness Curve

A richness curve is easily motivated and illustrated by example. Consider the data in Table 1, representing three hypothetical segmentation schemes. Each scheme specifies a different mutually exclusive partitioning of 400 consumers into eight segments of 50 consumers each. Given a pattern of segment proportions, as in Table 1, R^2 provides one way to compare overall magnitude of differences among segments. For these three schemes, we find that R^2 for each scheme equals .08, suggesting that all perform equally well. A quick inspection of the proportions, however, suggests that it is unrealistic to consider all three schemes equivalent. For example, a manager would likely find scheme A much more useful than scheme C, because the differentiation among scheme A segments is at the highest usage rates, whereas that for scheme C is among the lowest.

Table 1
THREE HYPOTHETICAL SEGMENTATION SCHEMES

	Proportion of consumers using brand x in each segment		
Segment	*Scheme A*	*Scheme B*	*Scheme C*
1	.74	.63	.54
2	.54	.57	.52
3	.40	.51	.50
4	.38	.45	.48
5	.36	.39	.46
6	.34	.33	.44
7	.32	.27	.30
8	.30	.21	.10
Total	.42	.42	.42
R^2 for scheme	.08	.08	.08

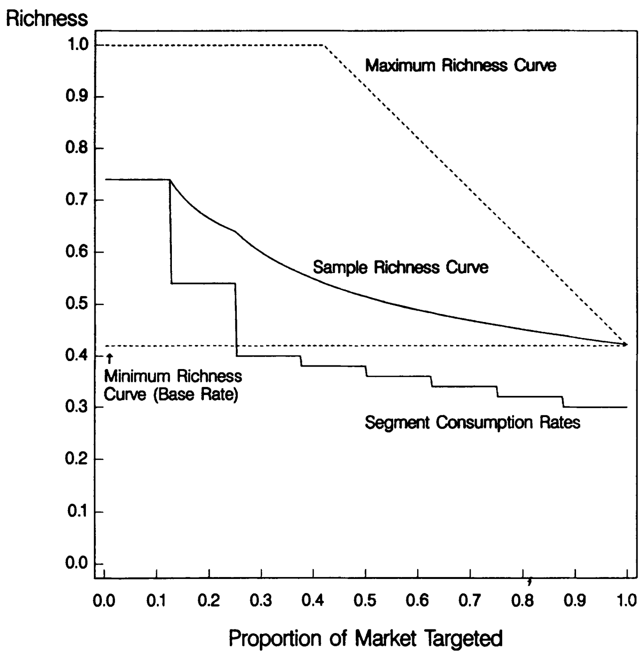
The difficulty in using R^2 for evaluating market segmentation schemes is that it does not reflect the marketing context. In obtaining R^2, one considers all deviations of segment means about the overall mean to be equally important. This is not the normal marketing context, however; marketers generally concentrate on only a few segments rather than all segments, and their concern is how effective targeting is within those few segments. If R^2 "fails" in this application, how else can we represent differences among segments in these three schemes?

Richness curves provide one solution. Figure 1 is the richness curve for scheme A.[1] The vertical axis represents richness, the percentage of consumers who use product x, and the horizontal axis the proportion of the market targeted. The richness curve shows how much above the base rate in the unsegmented market the marketer can expect to reach consumers by segmenting the market at a given size target market, given the optimal combination of segments in the segmentation scheme. For example, the marketer who targets the segment having the highest concentration of users in Figure 1 will get a return of 74% from the "richest" segment; as the marketer attempts to reach a larger part of the population by adding part or all of the second richest segment (and subsequent segments), the richness drops. The superiority of the first segment is clearly seen, and the richness for any target size can be seen at a glance.

Besides providing a context-contingent evaluation of market segmentation, richness curves afford a more interpretable evaluation than variance explained measures. As noticed by Bass, Tigert, and Lonsdale (1968) and Rosenthal and Rubin (1979, 1982), low R^2 values do not imply trivial differences among segment means. Richness curves measure scheme effectiveness in terms of proportion of market targeted, which is directly interpretable. Further, richness can be converted into profits and expenditures per capita across different shares of the total market. Thus, the curve provides what some re-

[1]We are not the first to apply richness curves to market segmentation. Sleight and Leventhal (1989) use richness curves to compare four geodemographic segmentation schemes, but mistakenly refer to the curves as "Lorenz" curves.

Figure 1
CONSTRUCTING THE RICHNESS CURVE



searchers have referred to as a "meaningful measure of effect" (Yeaton and Sechrest 1981).

The "base rate," r, indicating the proportion of the entire sample of $n = 400$ consumers that use the hypothetical product, is the richness of the unsegmented market and is indicated by a dashed horizontal line in Figure 1. This rate (42% in Figure 1) indicates the return the marketer can expect from not segmenting the market, but appealing to a random sample of consumers. The $S = 8$ segments in Figure 1 are arranged, left to right, in decreasing order of richness. The richness for each segment separately, u_s, is shown as a descending step function. The richness curve can be expressed as a weighted average of segment-level usage rates, added in decreasing rank order ($u_s \geq u_{s+1}$). The richness for a target containing a proportion t of the market, r_t, is informally defined as:

$$(1) \qquad r_t = \sum_{j=1,t} u_j^*/i$$

where:

$i = \text{RND}(nt)$,
$u_i^* = u_s$ if i^{th} consumer $\in s^{\text{th}}$ segment, and
RND is a rounding function.

Elements r_t specify the richness curve, where $r_{t'} \geq r_t$ if $t' > t$, so that r_t specify monotonically decreasing (i.e., nonincreasing) richness values of target segments containing $t = i/n$ of the sample. The richness curve is almost, but not quite, piecewise linear (actually it is piecewise concave). Quantities u_j^* are interpreted as the probability that the j^{th} respondent engages in a given consumer behavior, conditional on membership in segment s.[2] Appendix A provides a more formal definition of the richness curve as the solution to a linear programming problem.

RICHNESS CURVES AND LORENZ CURVES

Marketing researchers recently have expressed interest in applying measures of market concentration (Buchanan and Morrison 1987; Schmittlein 1986; Schmittlein, Cooper, and Morrison 1990) developed in industrial organization for problems involving measurement of customer concentration. In particular, these researchers have applied the Lorenz curve and Gini coefficient (see, e.g., Singer 1968). We therefore contrast the richness curve with the Lorenz curve, as the latter has also been discussed in a segmentation context and provides an alternate way to proceed. Whereas the richness curve specifies the proportion of respondents in a target of size t who are users of a product, the "modified" (Buchanan and Morrison 1987) Lorenz curve[3] specifies the percentage of *all* users that is contained in a target of size t, where targets are formed as in the richness curve by aggregating segments in descending order of u_s. The modified Lorenz curve is a simple function of the richness curve,

$$(2) \qquad L_t = tr_t/r,$$

where L_t is the proportion of all users who are in a target of size t.

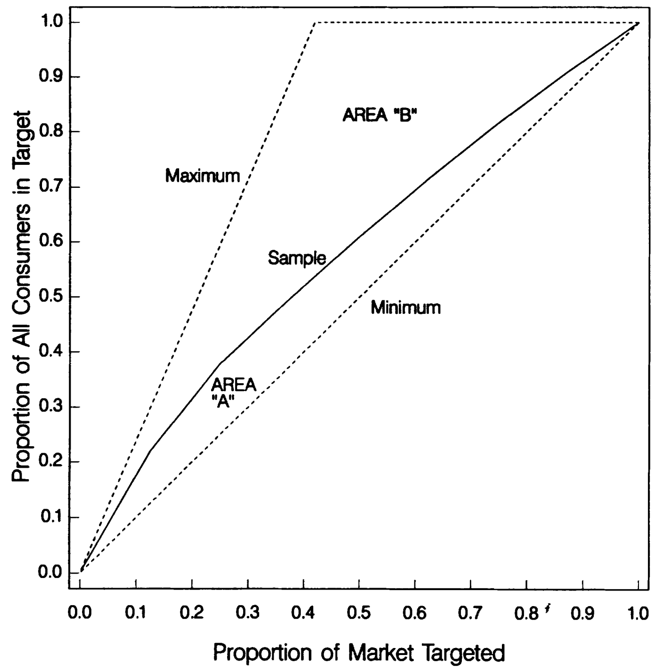
In the "worst case," all u_s are equal and the modified Lorenz curve is given by a straight line passing through (0,0) and (n,1). In the "best case," one segment (of size nr) captures 100% of the consumers; this defines a Lorenz curve from (0,0) to (nr,1) and from (nr,1) to (n,1). Figure 2 shows best, worst, and observed modified Lorenz curves for segmentation scheme A from Table 1.

The modified Lorenz curve, like the richness curve, is contingent on the size of market targeted and is easily

[2]The probabilities u_i^* are assumed to be constant within segment, as opposed to approaches taken by stochastic modelers, which specify a distribution of probabilities. In our case, a constant probability is reasonable. Our consumption variable is a fixed binary indicator of consumer status (user/nonuser) rather than an indicator of stochastic choice. The person either uses the product or does not. We do not have a true purchase probability estimated, say, from a string of scanner data. If we did have such data, we could model purchase probabilities with the beta-binomial distribution (Buchanan and Morrison 1987, 1988). Or, following Schmittlein, Cooper, and Morrison (1990), we could use Morrison's (1969) extension of the negative binomial to model number of purchases.

[3]The Lorenz curve typically orders consumption along the horizontal axis from lowest to highest; our modification orders consumption from highest to lowest for comparability with the richness curve.

Figure 2
MODIFIED LORENZ CURVE FOR SCHEME A



interpretable. However, the two curves are useful for different applications. The relationship presented in the richness curve is a simple series of "moving deviation scores"—that is, differences in proportions between target segments of gradually increasing size and the unsegmented total market. Thus, a marketing problem that can be formulated in terms of the degree of "deviance" of a target segment from the market as a whole suggests using the richness curve representation. Product positioning and advertising decisions center around documentation of such patterns of deviance. The modified Lorenz, in contrast, should be considered if attention is to be directed specifically to the problem of reaching or capturing a fixed proportion of all consumers. Because by observing the slope of the modified Lorenz we can identify a point of "diminishing return," the modified Lorenz is also useful for identifying an optimal target size (note that this can also be simply seen by comparing the segment proportions in Table 1 with the base rate). In Figure 2, diminishing returns are obtained in the sample after inclusion of the first two segments. However, comparing the two curves highlights some problems in the presentation of segmentation data. The Lorenz curve puts the visual emphasis at the "back end" of the curve, which is not usually of interest. Moreover, the superiority of segment 1 is not shown as clearly as in the richness curve.

Overall Measures for Comparing Segmentation Schemes

The Gini coefficient is an overall measure of concentration based on the Lorenz curve (Schmittlein 1986; Schmittlein, Cooper, and Morrison 1990) that has direct application as an evaluative index for market segmentation. The "modified Gini coefficient," g (Buchanan and Morrison 1987), is an overall measure that quantifies the extent to which the modified Lorenz curve deviates from the 45-degree line in Figure 2 that represents equal consumption rates across segments. The modified Gini coefficient is obtained as the ratio of the area B in Figure 2 to areas A + B,[4] and values for each of the three schemes are $g_A = .280$, $g_B = .323$, and $g_C = .281$. Thus, the modified Gini coefficient suggests that scheme A performs the worst. Again, inspection of Table 1 shows the practical utility of this conclusion to be questionable.

An alternate overall measure based on the richness curve is simply the area between the richness curve and the horizontal base rate line. It is directly interpreted as the "average richness gain"[5] (ARG) over the base rate across all t. For the three schemes in Table 1, average richness gains are $ARG_A = .127$, $ARG_B = .117$, and $ARG_C = .086$. These values fit with our intuition about Table 1.

The reason for the discrepancy between the Gini coefficient and ARG is easy to see if we express

$$\text{(3)} \qquad ARG = \int_0^1 (r_t - r)$$

$$g = \int_0^1 (t(r_t/r - 1)).$$

ARG is the mean increase in richness, whereas the modified Gini coefficient is the average proportionate increase in richness weighted by the target size. Though the Gini coefficient incorporates both target size and richness in an overall index, we believe target size is a separate strategic decision and should not influence the overall index.

However, one should question whether either overall index is practically meaningful. In a specific application, it is realistic to obtain an index only for values of t that

[4]The maximum Lorenz curve in Figure 2 assumes that individual probabilities are either 0 or 1. For product usage, a stochastic assumption that individual probabilities range between 0 and 1 may be more realistic (Buchanan and Morrison 1987), and the maximum curve will be attenuated. However, the relative across-scheme comparisons of the modified Gini coefficient will not be affected because the attenuated denominator term in the Gini coefficient will be constant for all schemes.

[5]The overall measure based on the richness curve can be normed by dividing by the area under the maximum richness curve. Given a base rate of r, the area under the maximum curve is $r(1 - r) + (1 - r)^2/2 = (1 - r^2)/2$. The normed measure should be used when one or more segmentation schemes are compared across a series of products with different base rates.

correspond to target sizes of managerial interest. As a manager is unlikely to target 80% or more of the marketplace, why should these target sizes affect the overall index? If overall indices are used, they should be conditional on a relevant range of target sizes.

BIAS AND STABILITY OF THE RICHNESS CURVE

Unfortunately, it is not possible simply to calculate richness curves, such as in Figure 1, from sample survey data and immediately use the curve and summary statistics to evaluate the effectiveness of one or more segmentation schemes. The estimates of richness, u_s, obtained in sample data contain sampling error and, because of the way richness curves are constructed, this error will overstate the richness that can be expected in the population. We first examine the problem and then describe bootstrap methods to correct for the bias.

Respondents in the sample who are classified into a given segment may overrepresent or underrepresent the actual segment richness in the population. However, because segments are ranked by apparent richness when richness curves are constructed, segments with a positive error in the estimated segment richness tend to be put first in the richness curve. As an illustration, suppose there are only two segments, A and B, and that our segmentation system is useless—on average, the population segment richness for both segments is equal to the base rate, which we assume is .50. To test the segmentation, we draw a sample and find, by chance, that one segment is "richer" than the other (say the sample richness values are A = .40 and B = .60). If we naïvely place the "richer" segment B first, construct a richness curve, and announce that we expect a lift of .10 over the base rate by targeting segment B, we are clearly wrong. However, with only one sample on which to base estimates, we cannot recognize this mistake.

This mistake means that the rank ordering of segments in the sample based on observed richness is also subject to error. Unless this possible transposition of segments and the inflation in richness caused by positive errors can be adjusted, the sample will give us misinformation as to which segments should be targeted first. Finally, without some sort of confidence interval about the (adjusted) richness curve, a manager cannot know how accurate results are or for which target sizes a richness curve is significantly different from the base rate. An overall test of the null hypothesis that the richness curve is equal to the base rate for all values of t is provided by the standard chi square test of the equality of segment proportions. However, if the chi square test is significant, we will not know for which specific target sizes the richness curve differs from the base rate. Confidence intervals based on the bootstrap provide such specific guidance.

Bootstrapped Richness Curves

The bootstrap (Efron 1982) is a nonparametric method for estimating bias and variability of a sample estimate, using replicated resampling from the empirical sample probability distribution. As such, the bootstrap provides a simple method of determining the extent of bias in a richness curve estimated from sample data, as well as a confidence interval for the richness curve.

Informally, the bootstrap adjusts for bias in the sample richness curve as follows. Let ρ_t specify the population value of the richness curve for target size t, and r_t the observed value of richness in the sample data. The true bias will be $(r_t - \rho_t)$. Unfortunately, we do not know ρ_t, so we cannot directly establish the degree of bias. However, we can estimate bias by considering a *different* population that has r_t as its true population richness. Then we can sample from this known population distribution and, by simulation over a series of J samples, estimate the average sample richness, $\hat{r}_t$, for this population. Then bias can be estimated indirectly as $(\hat{r}_t - r_t)$ and an estimate of the true population richness curve can be obtained as $\tilde{\rho}_t = r_t - (\hat{r}_t - r_t) = 2r_t - \hat{r}_t$.

Formally, let ϕ be the function that generates a sample richness curve, $\mathbf{r}$, from a joint probability vector, $\mathbf{p}$. Values p_i, where $i = 1, \ldots 2S$, specify the joint probability of the binary consumption variable and segment membership, as observed in the sample data. We can express the sample richness curve as

$$\mathbf{r} = \phi(\mathbf{p}). \tag{4}$$

We then generate $J = 200$ bias-adjusted "pseudo-richness curves" (Gifi 1990, p. 420), $\tilde{\mathbf{r}}_j$,

$$\begin{aligned} \tilde{\mathbf{r}}_j &= \phi(\mathbf{p}) - [\phi(\mathbf{v}_j/n) - \phi(\mathbf{p})] \\ &= 2\phi(\mathbf{p}) - \phi(\mathbf{v}_j/n) \\ &= 2\mathbf{r} - \hat{\mathbf{r}}_j, \end{aligned} \tag{5}$$

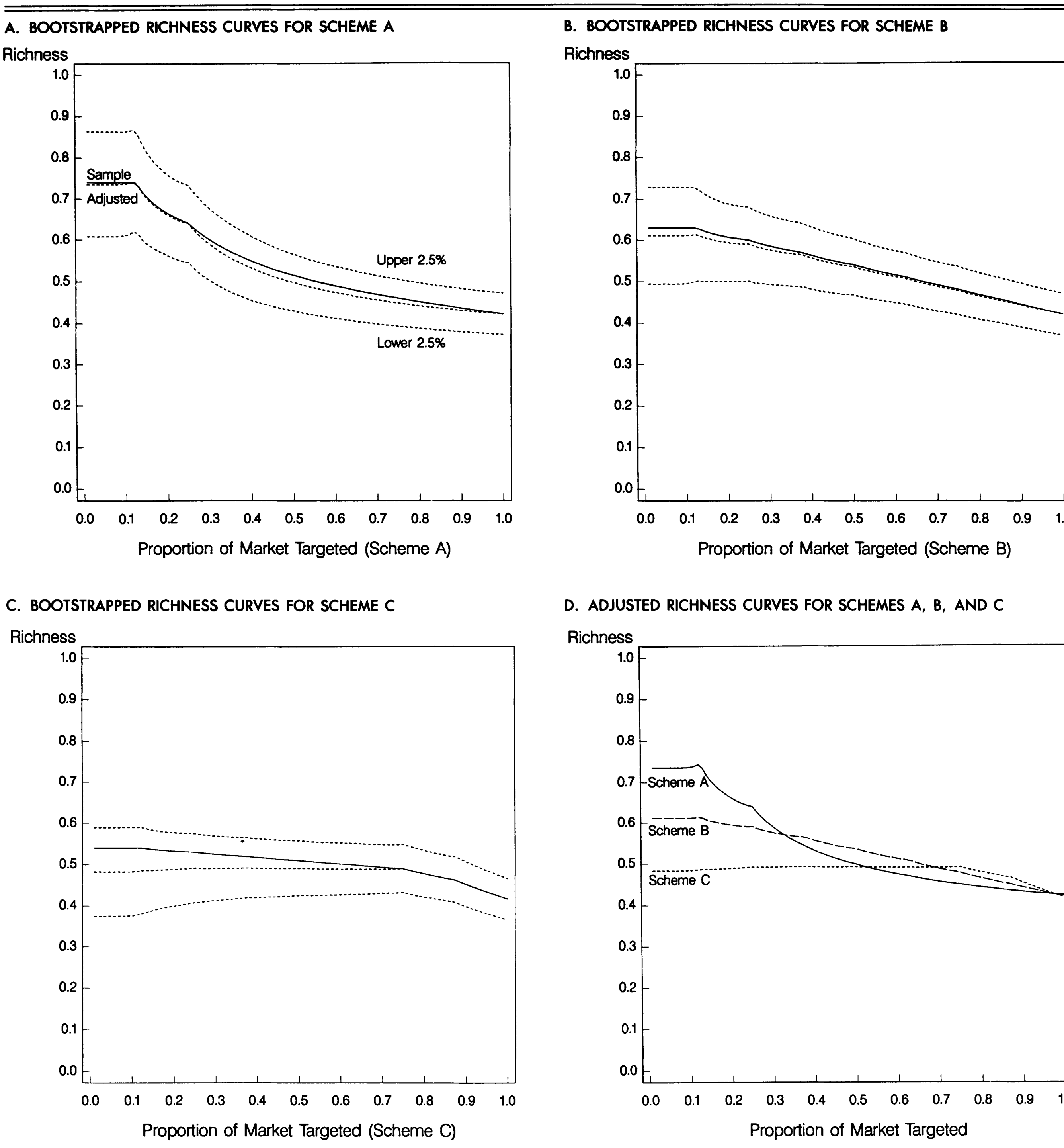
where $\mathbf{v}_j$ is one of J vectors whose elements specify n observations sampled with replacement from the multinomial distribution with probability vector $\mathbf{p}$, and where $\hat{\mathbf{r}}_j = \phi(\mathbf{v}_j/n)$. The population richness curve, $\boldsymbol{\rho} = \phi(\boldsymbol{\pi})$, with $\boldsymbol{\pi}$ the population probability vector, is then approximated as the average

$$\tilde{\boldsymbol{\rho}} = \Sigma\tilde{\mathbf{r}}_j/\mathbf{J}. \tag{6}$$

A confidence interval is defined, based on the middle 95% of $\tilde{\mathbf{r}}_{ij}$ values for each j in the J pseudo-richness curves.

Parts A through D of Figure 3 show sample (solid line) and mean adjusted, upper 2.5%, and lower 2.5% bootstrapped (dashed lines) richness curves. As expected, the confidence interval narrows for all three schemes as target size increases. Part D superimposes the adjusted curves for the three schemes. Clear differences among the three schemes can be seen. In parts A and B, curves for schemes A and B show minimal bias, whereas part C shows that bias in scheme C is particularly large, with an adjusted richness curve that is almost flat and a 95% confidence interval that generally includes the base rate for all target sizes. In part D, scheme A clearly outperforms scheme B for target sizes $t < .3$, whereas scheme B outperforms

Figure 3



A for $t > .3$. If a niche strategy is to be pursued, scheme A is preferable, but as the size of the target segment approaches the base rate, scheme B is preferred.

The richness curve makes clear that the superiority of a segmentation scheme is not effectively determined through (1) R^2 (which suggests all are the same), (2) modified Gini (which suggests scheme B is best), or (3) visual inspection of Table 1 (which suggests scheme A is the best). Richness curves, in contrast, facilitate a contingent evaluation of market segmentation schemes.

Factors That Contribute to Bias in Richness Curves

Parts A through C of Figure 3 show the extent of bias in a sample richness curve for our hypothetical example. In a broader context, however, how large is the bias in a sample richness curve and what factors affect the degree of bias and stability? A Monté Carlo analysis was performed to address these questions. Table 2 presents the levels of the seven factors investigated in the Monté Carlo analysis. The first three factors—number of segments, sample size, and base rate—are self-explanatory. Factor 4 specifies the range of segment sizes. Three conditions were used: all segments of equal size, two different segment sizes with the larger segments three times the size of the smaller, and two different sizes with the larger eight times the smaller. Factor 5 specifies whether the larger or smaller segments have relatively higher or lower usage rates and factor 6 specifies the range in segment usage rates. Finally, factor 7 specifies the degree of "skew" in the usage rates. For example, schemes A, B, and C in Table 1 provide illustrations, respectively, of top, equal, and bottom skew conditions.

Because the seven factors produce 2187 possible combinations, a 3^7 fractional factorial design (Addleman 1962) was used to generate a subset of 27 trials that allowed main effects to be estimated (see DeSarbo and Carroll

Table 2
FACTORS FOR MONTÉ CARLO ANALYSIS

Factor	*Levels*
1 Number of segments, S	A 4 B 8 C 12
2 Sample size, n	A 250 B 500 C 1000
3 Base rate, r (total sample usage rate)	A .10 B .20 C .30
4 Ratio of largest to smallest n_s, where half of the n_s are "large" and half are "small"	A 1:1 ($n_s = n/S$) B 3:1 (small $n_s = n/2S$, big $n_s = 3n/2S$) C 8:1 (small $n_s = 2n/9S$, big $n_s = 16n/9S$)
5 Size of n_s for segments ordered by usage rate	A Small n_s segments have highest usage B Small and large n_s segments alternate (half replications using *slsl...*, half *lsls...*) C Large n_s have highest usage rates
6 Range of segment usage rates, $d = \max(u_s) - \min(u_s)$	A .10 B .20 C .40
7 Evenness[a] of usage rates, u_s	A Top heavy B Equal differences C Bottom heavy

[a]After specification of r, n, and n_s from the above parameters, usage rates in these three conditions are solved for as follows:

A. Top skew condition. Bigger difference between largest usage rates. We have S equations in S unknowns, and solve for u_s:

$$\Sigma n_s u_s = rn$$
$$u_1 - u_2 = d(w_s/\Sigma w_s);\ w_s = 10$$
$$u_2 - u_3 = d(w_s/\Sigma w_s);\ w_s = 5$$
$$u_3 - u_4 = d(w_s/\Sigma w_s);\ w_s = 1$$
$$\vdots$$
$$u_{S-1} - u_S = d(w_s/\Sigma w_s)\ w_s = 1$$

B. Equal difference condition. Solve for u_s:

$$\Sigma n_s u_s = rn$$
$$u_1 - u_2 = d(w_s/\Sigma w_s);\ w_s = 1$$
$$u_2 - u_3 = d(w_s/\Sigma w_s);\ w_s = 1$$
$$\vdots$$
$$u_{S-1} - u_S = d(w_s/\Sigma w_s)\ w_s = 1$$

C. Bottom skew condition. Bigger difference between smallest usage rates. Solve for u_s:

$$\Sigma n_s u_s = rn$$
$$u_1 - u_2 = d(w_s/\Sigma w_s);\ w_s = 1$$
$$u_2 - u_3 = d(w_s/\Sigma w_s);\ w_s = 1$$
$$\vdots$$
$$u_{S-3} - u_{S-2} = d(w_s/\Sigma w_s);\ w_s = 1$$
$$u_{S-2} - u_{S-1} = d(w_s/\Sigma w_s);\ w_s = 5$$
$$u_{S-1} - u_S = d(w_s/\Sigma w_s)\ w_s = 10$$

1985; DeSarbo and Cho 1989 for similar analyses).[6] These 27 trials appear in Table 3.[7] Given the constraints in Table 2, each of the 27 trials uniquely specifies an $S \times 2$ contingency table. For each of the 27 trials, 200 sample contingency tables were randomly generated. The bias for five different target sizes was obtained as

$$b_{tij} = r_{tij} - \rho_{tj}, \qquad \text{for } t = .1, .2, .3, .4, .5, \tag{7}$$

where r_{tij} is the richness for the i^{th} replication of the j^{th} combination of design factors and ρ_{tj} is the "population" richness for the j^{th} combination. Standard errors, s_{tj}, also were obtained. Results for the 27 trials are reported in Table 3. We can see that some trials produce bias up to .10; clearly there are certain situations in which bias adjustments are crucial.

Results of main-effect ANOVAS for five mean bias dependent variables and five standard error variables, as well as the marginal means on these variables, are reported in Table 4. Across all five target sizes, bias increases with factor 1 (a higher number of segments), factor 2 (a smaller sample size), and factor 3 (a higher base rate). Factor 4 (the ratio of largest to smallest segment sizes) and factor 5 (the sample sizes of segments with relatively high or low usage rates) are not significantly related to bias. Thus, individual segment sizes do not significantly affect bias, but the number of segments and total sample size do. Market segmentation in small samples and with many segments will produce highly biased richness curves. For target size t equal to .1, factor 7 (a "bottom-heavy skew") contributes to bias, whereas for target sizes t equal to .1 and .5, factor 6 (a small range of segment usage rates) is related to bias. In practical

[6]Because of the modest scope of this Monté Carlo analysis, these results should be considered preliminary. Ideally, a full factorial design is preferred, possibly with additional levels of some factors. However, computational time and expense generally preclude that option in actual applications.

[7]Some of the potential combinations produced negative usage rates (e.g., 1A + 2A + 3A + 4C + 5C + 6C + 7C). Though the design in Table 3 was found by trial and error, the approach suggested by Steckel, DeSarbo, and Mahajan (1991) could be used to derive a design matrix more directly.

Table 3
BIAS AND STANDARD ERROR FOR 27 TRIALS

	Design factors[a]							*Bias measures*[b]					*Standard error measures*[c]				
Trial	*F1*	*F2*	*F3*	*F4*	*F5*	*F6*	*F7*	b_1	b_2	b_3	b_4	b_5	s_1	s_2	s_3	s_4	s_5
1	A	A	A	B	C	A	C	.0246	.0227	.0218	.0195	.0147	.0296	.0290	.0291	.0280	.0250
2	B	B	A	C	C	A	C	.0344	.0289	.0241	.0197	.0165	.0269	.0238	.0207	.0188	.0176
3	C	C	A	A	C	A	C	.0402	.0297	.0228	.0179	.0136	.0229	.0181	.0157	.0142	.0123
4	A	A	A	B	B	B	B	.0108	.0093	.0077	.0053	.0016	.0529	.0439	.0402	.0379	.0346
5	B	B	A	C	B	B	B	.0227	.0140	.0099	.0065	.0034	.0342	.0319	.0283	.0239	.0219
6	C	C	A	A	B	B	B	.0175	.0112	.0077	.0051	.0030	.0267	.0216	.0187	.0164	.0148
7	A	A	A	B	A	C	A	−.0083	−.0017	.0013	.0019	.0022	.0883	.0694	.0571	.0439	.0362
8	B	B	A	C	A	C	A	.0103	.0085	.0073	.0046	.0032	.0522	.0299	.0248	.0213	.0189
9	C	C	A	A	A	C	A	−.0005	.0039	.0055	.0036	.0018	.0439	.0295	.0232	.0193	.0168
10	A	B	B	A	B	A	A	.0019	.0019	.0029	.0048	.0054	.0326	.0326	.0290	.0252	.0236
11	B	C	B	B	B	A	A	.0118	.0100	.0127	.0128	.0111	.0306	.0269	.0223	.0190	.0168
12	C	A	B	C	B	A	A	.0813	.0647	.0622	.0451	.0421	.0531	.0446	.0405	.0373	.0343
13	A	B	B	A	A	B	C	.0171	.0171	.0115	.0053	.0003	.0334	.0334	.0300	.0271	.0255
14	B	C	B	B	A	B	C	.0336	.0214	.0166	.0115	.0064	.0322	.0245	.0209	.0191	.0176
15	C	A	B	C	A	B	C	.1089	.0812	.0616	.0472	.0356	.0611	.0456	.0393	.0365	.0346
16	A	B	B	A	C	C	B	−.0011	−.0011	−.0006	.0003	−.0014	.0468	.0468	.0408	.0352	.0334
17	B	C	B	B	C	C	B	.0175	.0161	.0123	.0097	.0077	.0270	.0262	.0220	.0201	.0182
18	C	A	B	C	C	C	B	.0658	.0533	.0418	.0339	.0278	.0530	.0459	.0398	.0358	.0331
19	A	C	C	C	A	A	B	.0123	.0064	.0043	.0033	.0024	.0419	.0257	.0209	.0193	.0185
20	B	A	C	A	A	A	B	.0823	.0669	.0549	.0449	.0357	.0584	.0481	.0434	.0394	.0359
21	C	B	C	B	A	A	B	.1041	.0815	.0673	.0559	.0459	.0441	.0342	.0305	.0278	.0262
22	A	C	C	C	C	B	A	−.0009	.0391	.0390	.0390	−.0001	.0233	.0233	.0233	.0233	.0210
23	B	A	C	A	C	B	A	.0301	.0329	.0363	.0374	.0326	.0694	.0511	.0432	.0390	.0357
24	C	B	C	B	C	B	A	.0180	.0251	.0287	.0297	.0272	.0472	.0366	.0301	.0267	.0246
25	A	C	C	C	B	C	C	.0070	.0018	.0000	−.0008	−.0043	.0277	.0221	.0216	.0217	.0212
26	B	A	C	A	B	C	C	.0728	.0579	.0462	.0363	.0273	.0606	.0507	.0460	.0429	.0405
27	C	B	C	B	B	C	C	.0518	.0350	.0255	.0200	.0147	.0462	.0395	.0348	.0316	.0293

[a]As identified in Table 2.
[b]Bias measures for five target sizes as defined in expression 7.
[c]Standard error in sample richness values for five target sizes.

Table 4
ANOVA RESULTS AND MARGINAL MEANS FOR BIAS AND STANDARD ERROR

Design factor	Factor level	Results for bias, b_t, where: $t = .1$	$t = .2$	$t = .3$	$t = .4$	$t = .5$	Results for S.E., s_t, where: $t = .1$	$t = .2$	$t = .3$	$t = .4$	$t = .5$
		****[a]	**	**	**	****	ns	ns	ns	ns	ns
F1	A	.0070	.0106	.0098	.0087	.0023	.0418	.0362	.0324	.0291	.0266
	B	.0351	.0285	.0245	.0204	.0160	.0435	.0348	.0302	.0271	.0248
	C	.0541	.0428	.0359	.0287	.0235	.0443	.0358	.0303	.0273	.0251
		***	**	**	**	****	****	****	****	****	****
F2	A	.0520	.0430	.0371	.0302	.0244	.0585	.0476	.0421	.0379	.0344
	B	.0288	.0234	.0196	.0163	.0128	.0404	.0343	.0299	.0264	.0246
	C	.0154	.0155	.0134	.0113	.0046	.0307	.0242	.0210	.0192	.0175
		**	*	**	**	****	ns	ns	ns	**	***
F3	A	.0169	.0141	.0120	.0094	.0067	.0420	.0330	.0287	.0249	.0220
	B	.0374	.0294	.0246	.0190	.0150	.0411	.0363	.0316	.0284	.0264
	C	.0419	.0385	.0336	.0295	.0202	.0465	.0368	.0326	.0302	.0281
		ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
F4	A	.0290	.0245	.0208	.0173	.0131	.0439	.0369	.0322	.0287	.0265
	B	.0293	.0244	.0216	.0185	.0146	.0442	.0367	.0319	.0282	.0254
	C	.0380	.0331	.0278	.0221	.0141	.0414	.0325	.0288	.0264	.0246
		ns	ns	ns	ns	ns	*	ns	ns	ns	ns
F5	A	.0400	.0317	.0256	.0198	.0148	.0506	.0378	.0322	.0281	.0256
	B	.0308	.0229	.0194	.0150	.0116	.0405	.0349	.0313	.0284	.0264
	C	.0254	.0274	.0251	.0230	.0154	.0385	.0334	.0294	.0268	.0245
		*	ns	ns	ns	***	*	*	*	*	*
F6	A	.0437	.0348	.0303	.0249	.0208	.0378	.0315	.0280	.0254	.0234
	B	.0286	.0279	.0243	.0208	.0122	.0423	.0347	.0305	.0278	.0256
	C	.0239	.0193	.0155	.0122	.0088	.0495	.0400	.0345	.0302	.0275
		**	ns	ns	ns	ns	*	ns	ns	ns	ns
F1	A	.0160	.0205	.0218	.0199	.0139	.0490	.0382	.0326	.0283	.0253
	B	.0369	.0286	.0228	.0183	.0140	.0428	.0360	.0316	.0284	.0263
	C	.0434	.0329	.0256	.0196	.0139	.0379	.0319	.0287	.0267	.0249

[a]Significance levels:
ns $p > .05$
*$p \leq .05$
**$p \leq .01$
***$p \leq .001$
****$p \leq .0001$
The p-values are for tests of the null hypothesis that the three bias measures (or standard errors) for the three factor levels are the same for each of the seven design factors.

terms, these results indicate that unless the characteristics of the data (few segments, high sample size, top skew, etc.) suggest otherwise, bias adjustment should be performed routinely.

Standard error, in contrast, is related significantly to factor 2 (sample size) across all target sizes, as one would expect. In addition, standard error differs by factor 6 (increases as the range of segment usage rates increases), also across all target sizes. For small target sizes ($t = .1$), standard error is reduced if the largest segments have the highest usage rates (factor 5) and if there is a "top-heavy" skew condition (factor 7). Finally, for large target sizes ($t = .4, .5$), standard error is smaller for lower base rates (factor 3). Though not as dramatic as the differences in mean bias, the results indicate that characteristics of the data other than sample size can affect the confidence interval about a richness curve, and that confidence intervals for a number of segmentation schemes should not be assumed to have equal widths.

Performance of the Bootstrap Adjustment

Nonmonotonicity of adjusted richness curves. On careful examination, we see that the bootstrapped richness curves in Figure 3 are *not* entirely monotone decreasing. Rather, they first increase and then decrease. Appendix B shows how the nonmonotonicity results from the bias adjustment formula in expression 5, in that a linear combination of monotone decreasing functions, $\phi(\mathbf{p})$ and $\phi(\mathbf{v}_j/n)$, may not be a monotone decreasing function. Thus, a monotone decreasing adjusted curve is not guaranteed by bootstrap methodology.

However, the nonmonoticity observed in a single sample disappears when we average adjusted curves obtained from a number of samples. We performed an additional analysis in which the usage rates in Table 1 were assumed to be *population* values for the three segmentation schemes, rather than observed sample tables. We then sampled from the population tables to generate $T = 50$ "observed sample tables." Further, for each of these 50 observed sample tables, we obtained a bootstrap-adjusted richness curve by using expressions 5 and 6, based on $J = 50$ resampled tables. The nonmonotonicity observed in Figure 3A–C canceled out when we aggregated over $T = 50$ observed samples. Hence, though increasing adjusted richness curves may appear in a given sample, the expected value of the adjusted curve does not appear to exhibit nonmonotonicities.

This assertion is supported by further Monté Carlo simulations with the same 27 trials defined by the factors in Table 2. For each of the 27 trials, 25 sample contingency tables were generated, and for each of the 25 sample tables an adjusted bootstrapped curve based on 25 resampled tables was obtained. Other than minor fluctuations likely due to sampling error, we found no evidence of nonmonotonicity in the average adjusted richness curves obtained for the 27 trials.

Adequacy of the bootstrap adjustment. Another issue is how much of the bias in the observed richness curve is corrected by the bootstrap adjustment. The 27-trial Monté Carlo simulation just described was used to address this issue. For each of the 25 sample contingency tables, we obtained the average deviation of the sample richness curve (r_t) from the known population curve (ρ_t) and the average deviation of the adjusted bootstrapped curve ($\tilde{\rho}_t$) from the known population curve. Results are reported in Table 5.

The five columns on the right side of Table 5 show the degree of bias in the observed sample richness curve ($r_t - \rho_t$), and correspond to the first five columns of Table 4.[8] The first five columns in Table 5 show the bias in the adjusted richness curve ($\tilde{\rho}_t - \rho_t$); ideally this bias would be zero. The bootstrap adjustment removes most but not all of the bias in the observed curve. For example, for 10% target sizes, the observed sample richness curve for simulation trials with 12 market segments (factor F1, level C) overestimated richness by .0635; the adjusted bootstrapped richness curve still overestimated richness, but by only .0291.

Why does the bootstrap not remove all bias? An informal argument follows. If the adjusted richness curve is biased, then $\tilde{\rho}_t - \rho_t > 0$, so the $2r_t - \hat{r}_t - \rho_t > 0$, so that $r_t - \rho_t > \hat{r}_t - r_t$. Thus, a sample richness curve r_t corresponding to a population curve ρ_t will have *greater*

[8]Though the results in Table 5 are based on 25 sample tables and those in Table 4 are based on 200 sample tables, the two tables agree closely. Because of the computational burden, we reduced the number of sample tables from 200 to 25 in Table 5 and used only 25 replications in the bootstrap.

Table 5
MARGINAL MEANS FOR BOOTSTRAP ADJUSTED BIAS AND SAMPLE BIAS

Design factor	Factor level	Bias[a] in bootstrap adjusted richness curve for target size					Bias[b] in observed sample richness curve for target size				
		$t = .1$	$t = .2$	$t = .3$	$t = .4$	$t = .5$	$t = .1$	$t = .2$	$t = .3$	$t = .4$	$t = .5$
F1	A	.0013	−.0011	−.0016	−.0021	−.0026	.0073	.0043	.0027	.0017	−.0000
	B	.0106	.0092	.0082	.0068	.0046	.0345	.0283	.0239	.0194	.0147
	C	.0291	.0221	.0202	.0177	.0149	.0635	.0492	.0411	.0341	.0275
F2	A	.0243	.0177	.0157	.0134	.0103	.0569	.0441	.0364	.0301	.0234
	B	.0140	.0126	.0115	.0100	.0084	.0335	.0280	.0238	.0199	.0157
	C	.0027	−.0001	−.0004	−.0010	−.0019	.0150	.0097	.0074	.0053	.0030
F3	A	−.0006	−.0028	−.0027	−.0027	−.0031	.0124	.0082	.0062	.0043	.0021
	B	.0179	.0138	.0126	.0102	.0085	.0424	.0328	.0273	.0221	.0171
	C	.0236	.0192	.0169	.0149	.0115	.0505	.0407	.0342	.0288	.0228
F4	A	.0135	.0109	.0097	.0081	.0061	.0332	.0277	.0235	.0191	.0145
	B	.0149	.0114	.0107	.0097	.0079	.0348	.0272	.0234	.0202	.0161
	C	.0125	.0080	.0064	.0045	.0029	.0373	.0268	.0208	.0159	.0115
F5	A	.0176	.0111	.0082	.0054	.0028	.0389	.0286	.0223	.0167	.0118
	B	.0163	.0123	.0116	.0105	.0089	.0389	.0295	.0247	.0207	.0163
	C	.0070	.0067	.0070	.0065	.0052	.0274	.0236	.0207	.0178	.0140
F6	A	.0242	.0172	.0147	.0119	.0093	.0505	.0382	.0318	.0263	.0211
	B	.0088	.0064	.0053	.0042	.0020	.0297	.0231	.0185	.0145	.0100
	C	.0079	.0066	.0068	.0062	.0056	.0250	.0204	.0174	.0144	.0110
F7	A	.0088	.0066	.0078	.0079	.0070	.0251	.0210	.0201	.0185	.0159
	B	.0124	.0081	.0063	.0048	.0029	.0356	.0261	.0203	.0158	.0111
	C	.0197	.0156	.0127	.0096	.0070	.0447	.0346	.0273	.0209	.0150

[a]Bias estimated as $\tilde{\rho}_t - \rho_t$.
[b]Bias estimated as $r_t - \rho_t$.

upward bias than a sample richness curve $\hat{r}_t$ corresponding to a population curve $r_t > \rho_t$. Because sample richness curves will have positive bias, $\hat{r}_t > r_t > \rho_t$. Note that we could extend this sequence by considering $\hat{r}_t$ to represent the population richness curve of still another population and simulating an average richness curve for that population as well. However, because the richness curve (and the underlying segment probabilities) has an upper bound, there is a limit to how extreme such a sequence of richness curves can become. Therefore, each richness curve in this sequence will tend to be less extreme than the previous (e.g., $\hat{r}_t > r_t > \rho_t$, with $r_t - \rho_t > \hat{r}_t - r_t$) and the adjusted richness curve will tend to have some bias.

EMPIRICAL EXAMPLE

We present a brief empirical example, using data from the 1987 Leading Edge survey of consumers conducted in a national probability sample of 2591 adults by Chilton Research (Novak and MacEvoy 1990). The consumer behavior in which we are interested is whether the respondent drinks wine with dinner (two or more times per month). The base rate of this behavior in the sample is $r = .143$. Four segmentation schemes are used to define consumer segments.

1. *VALS (Values and Lifestyles),* a proprietary segmentation system of SRI International (Mitchell 1983). Respondents are classified into eight psychographic groups on the basis of their responses to eight demographic or political questions and 22 social attitude questions (e.g., [agree or disagree] "Communists should be banned from running for mayor.") The eight segments range in size from 2.3 to 36.8%.
2. *VALS2,* a proprietary segmentation system of SRI International (MacEvoy 1989) based on the responses to four demographic items and 42 self-concept and motivational questions (e.g., [agree or disagree] "I like to try new things"). The eight segments range in size from 8.1 to 15.2%.
3. *LOV (List of Values),* a nonproprietary segmentation scheme developed at the University of Michigan Survey Research Center (Kahle 1983; Veroff, Douvan, and Kulka 1981), based on the ranking of nine value statements (e.g., [I value most] "a sense of belonging"). The seven segments in this sample range from 4.2 to 20.7%.
4. *DuoVALS2,* a scoring of VALS2 that combines the primary segment classification with a secondary segment score, based on answers to the same VALS2 questions. The 34 segments in this survey range from .12 to 6.7%.

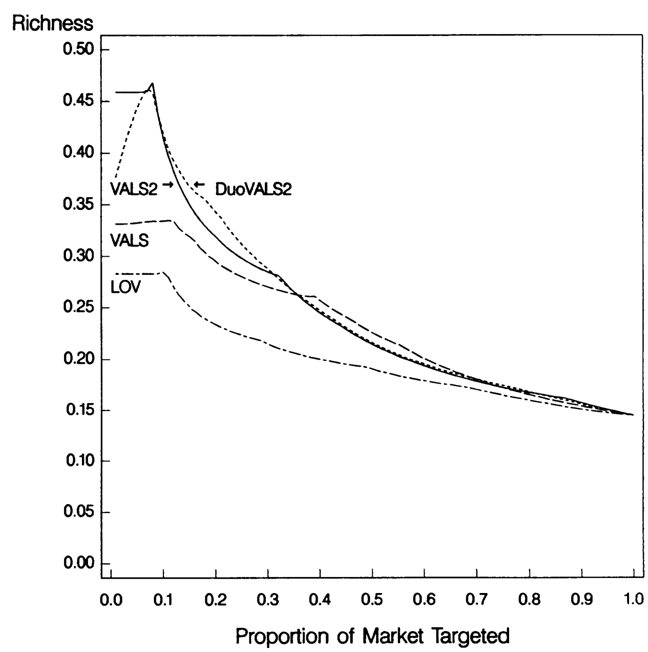
As a basis of comparison, adjusted R^2 values for the four schemes are: DuoVALS2 = .076, VALS2 = .085, VALS = .060, LOV = .013. Thus, R^2 suggests that (1) VALS2 is superior, (2) there is little difference between VALS2, DuoVALS2, and VALS, and (3) LOV performs poorly.

Our marketing objective is to form a target segment, based on each of the four schemes, that contains the greatest proportion of wine drinkers. Because none of the segments in any of the four schemes are directly accessible, we assume that customer self-selection is the operating principle. To the extent that we can identify a psychographic target segment that contains many wine drinkers, we can tailor a marketing communications program to that target. Which segmentation scheme can best identify such a target segment?

Adjusted bootstrapped richness curves for each of the four segmentation schemes are shown in Figure 4. Note that the nonmonotonicity of the adjusted DuoVALS2 curve is particularly evident in this example, likely because of the very large number of segments. In such situations, the pattern of nonmonotonicity indicates caution in interpreting the front end of the curve. This observation is reinforced by comparing standard deviations of the adjusted curves for DuoVALS2 and VALS2. For $t = .01$, the standard deviation is .087 for DuoVALS2 and .057 for VALS2; for $t = .10$ the values are .047 for DuoVALS2 and .050 for VALS2. The richness curve for DuoVALS2 is more variable than that for VALS2 for very small ($t < .10$) target sizes.

It is evident that for small target sizes ($t \leq .15$), VALS2 nearly doubles the increment over base rate provided by either LOV or VALS. However, note that even LOV does a respectable job of targeting consumer behavior for this range of target sizes. The adjusted R^2 (.013) disguises this fact and makes LOV appear nearly useless. The richness curves make explicit how each scheme performs at markets of different sizes. The superiority of VALS2 extends to a target segment of about 30% of the

Figure 4
WINE WITH DINNER: FOUR ADJUSTED RICHNESS CURVES



population; thereafter, the schemes are almost indistinguishable. Note that any errors in directing marketing communications to the target segment will attenuate the richness curve, so that the adjusted curve is best viewed as an upper bound. By straightforward multiplication of the total richness, or net increment over base rate, by the size of the market, one can convert differences among the schemes into upper bound estimates of the number of consumers for any projected market, and through this metric the marketer can project per capita profits, distribution costs, and advertising expenditures for each segmentation scheme.

DISCUSSION AND CONCLUSIONS

Though richness curves closely resemble Lorenz curves, our work is conceptually distinct from recent marketing applications of the Lorenz curve (Schmittlein 1986; Schmittlein, Cooper, and Morrison 1990). These applications are used to identify an individual-level consumption distribution for the purpose of specifying a cumulative concentration (i.e., Lorenz) curve. The resulting concentration curve implies the maximum possible richness curve, given specific stochastic assumptions about individual consumer behavior. Market segments are not identified, but the degree of concentration suggests the segmentability of the market. In contrast, we begin with predefined market segments and identify a segment-level richness curve as a way of quantifying the degree to which the segmentation can be used to target consumers effectively.

Though we emphasize that measures of variance explained can be misleading if used to evaluate market segmentation, that observation should not be taken to imply that variance explained measures should never be used. If interest is in a segment profile problem rather than a target segment problem, measures of variance explained provide a convenient, overall measure of magnitude of effect.

It is important to mention what we *have not* considered in evaluating market segmentation schemes. We focus on establishing the magnitude of actionable differences in segments, but other aspects of the segmentation also must be evaluated, such as segment stability and validity, segment responsiveness to the marketing mix, and segment accessibility. However, as these characteristics can be evaluated in relation to their impact on richness, richness curves provide a natural starting point for the evaluation process. For example, observed segment sizes could be replaced by "effective segment sizes" that adjust for reachability via media vehicles and cost of reaching.

Though we use binary criterion variables, the bootstrap can be used with continuous variables by resampling from the observed sample distribution. Continuous variables open the possibility of replacing a binary usage variable with an index that incorporates such considerations as segment accessibility and responsiveness. In addition, extensions to continuous (e.g., regression-based) rather than categorical segmentation variables, as well as the effect of measurement error on the segmentation and consumption variables, should be explored.

Last but not least, alternatives to the bootstrap that produce monotone decreasing adjusted curves, possibly based on the linear programming formulation in Appendix A, should be investigated. Because the bootstrap does not guarantee a monotone decreasing curve (though this problem disappears when one aggregates over multiple samples, and though the bootstrap does correct for most sample bias), other approaches to bias adjustment may ultimately prove preferable. Our results provide a benchmark against which future research can be assessed.

Besides evaluating segmentation, richness curves have other applications. They can be used to compare hierarchically nested market segments as a way of determining the number of segments to retain. The impact on richness of using four through 10 nested segments can be clearly seen. Further, effective graphic representation facilitates a quick and accurate comparison of segmentation schemes. Richness curves would be useful as part of a decision support system, allowing graphic analysis of large quantities of marketing data. Finally, richness can be used as a criterion for developing market segmentation schemes. Many cluster analysis algorithms implicitly attempt to maximize variance explained. However, if variance explained is not a reasonable evaluative criterion, it is not a reasonable developmental criterion either.

APPENDIX A
A FORMAL DEFINITION OF THE RICHNESS CURVE

Here we develop a rigorous definition of the richness curve. We have an $S \times 2$ table in which the rows indicate S present market segments and the columns the dichotomy user/nonuser. Now consider an arbitrary segment, T_t, containing the proportion t of users, representing a target segment that may or may not be one of the S present segments. The size of the arbitrary target is

(A1) $t = p(T_t) = \Sigma_{s=1,S}\, p(T_t|\text{segment} = s)p(\text{segment} = s)$,

and the proportion of users in the target is

(A2) $$p(\text{user}|T_t) = p(T_t|\text{user})p(\text{user})/p(T_t) = t^{-1}\{\Sigma_{s=1,S}\, p(T_t|\text{segment}=s \cap \text{user})\, p(\text{user}|\text{segment}=s)\, p(\text{segment}=s)\}.$$

Assume that selection into the target segment is unbiased within each of the S present segments, that is,

(A3) $p(T_t|\text{segment}=s \cap \text{user}) = p(T_t|\text{segment}=s)$,

or equivalently,

(A4) $$p(T_t \cap \text{user}|\text{segment}=s) = p(T_t|\text{segment}=s)p(\text{user}|\text{segment}=s).$$

The richness curve, $r_t = \max\{p(\text{user}|T_t)\}$, is defined as the set of target segments, $\{T_t\}$, with maximal usage rates. Substituting expression A3 into expression A2, and using the constraint implied by expression A1, we can formalize this interpretation of the richness curve by defining the richness curve as the (scaled) solution of a linear program,

(A5)

$$r_t = t^{-1} \max\{\Sigma_{s=1,S}\, \delta_s u_s p_s | (\Sigma_{s=1,S}\, \delta_s p_s = t) \cap (0 \le \delta_s \le 1)\}$$

where:

$\delta_s = p(T_t|\text{segment}=s)$,
$u_s = p(\text{user}|\text{segment}=s)$, and
$p_s = p(\text{segment}=s)$

Computing the richness at point t is thus a special case of a class of separable programs discussed, for instance, by Saaty (1959, p. 149–154).

To solve the linear program, we first reorder the usage rates u_s so they do not increase. Ties are broken arbitrarily. Now, let $u_{[1]}$ be the largest of the u_s, $u_{[2]}$ the second largest, and so on. Thus, the $u_{[s]}$ are the order statistics corresponding to the u_s. Order the p_s accordingly. Hence, $p_{[1]}$ is the size of the segment with the largest u_s, that is, the segment with usage rate $u_{[1]}$, and so on. Now,

(A6)
$$\begin{aligned}
r_t &= u_{[1]} && \text{if } t \le p_{[1]},\\
r_t &= t^{-1}\{p_{[1]}u_{[1]} + (t - p_{[1]})u_{[2]}\} && \text{if } p_{[1]} < t \le p_{[1]} + p_{[2]},\\
r_t &= t^{-1}\{p_{[1]}u_{[1]} + p_{[2]}u_{[2]} + (t - p_{[1]} - p_{[2]})u_{[3]}\} && \text{if } p_{[1]} + p_{[2]} < t \le p_{[1]} + p_{[2]} + p_{[3]},
\end{aligned}$$

and so on.

In addition to providing a rigorous definition of the richness curve, expression A5 is compact in that no complicated notation involving reordering is necessary to define the richness curve. The solution in expression A6 is also more convenient for obtaining selected values along the richness curve than expression 1, particularly for large sample sizes. Further, because the optimum of a linear program as a function of the parameters of the problem has been studied extensively, many results on continuity, convexity, and differentiability are available, so that statistical properties of richness curves can be studied in future work.

APPENDIX B
NONMONOTONICITY OF ADJUSTED RICHNESS CURVES

Figures 3 and 4 show nonmonotonic adjusted bootstrapped richness curves. First, note that the observed sample richness curve, $\phi(\mathbf{p})$, and the j^{th} bootstrap resampled richness curve, $\phi(\mathbf{v}_j/n)$, are both monotonically decreasing because the richness function, $\phi(\ )$, is by definition monotonically decreasing. However, the j^{th} bias-adjusted pseudocurve, $\tilde{\mathbf{r}}_j = 2\phi(\mathbf{p}) - \phi(\mathbf{v}_j/n)$, may not be monotonically decreasing because the linear combination of monotone decreasing functions is not necessarily a monotone decreasing function. It is easy to show how nonmonotonicities in the adjusted curve occur. Let $f(t)$ represent the observed richness curve, $g(t)$ the j^{th} bootstrap resampled curve, and $h(t)$ the j^{th} bias-adjusted pseudocurve, so the $h(t) = 2f(t) - g(t)$. Then, taking derivatives, we have

(B1) $$h'(t) = 2f'(t) - g'(t).$$

Then, $h'(t) > 0$ (the adjusted curve increases) when

(B2) $$2f'(t) > g'(t).$$

Because $f(t)$ and $g(t)$ are monotone decreasing, expression B2 implies that

(B3) $$g'(t) < 2f'(t) < 2f'(t) \le 0.$$

Expression B3 will almost certainly be true for values of t specifying a target size less than or equal to that of the richest segment. For such values of t, the observed richness curve will be flat and the derivative, $f'(t)$, will equal zero. If the j^{th} resampled curve, $g(t)$, is decreasing at this point, so that $g'(t) < 0$, the adjusted curve $h(t)$ will be increasing.

REFERENCES

Addleman, Sidney (1962), "Orthogonal Main-Effect Plans for Asymmetrical Factorial Experiments," *Technometrics*, 4 (February), 21–46.

Andreason, Alan R. (1966), "Geographic Mobility and Market Segmentation," *Journal of Marketing Research*, 3 (November), 341–50.

Assael, Henry and A. Marvin Roscoe, Jr. (1976), "Approaches to Market Segmentation Analysis," *Journal of Marketing*, 40 (October), 67–76.

Bass, Frank M., Douglas J. Tigert, and Ronald T. Lonsdale (1968), "Market Segmentation: Group Versus Individual Behavior," *Journal of Marketing Research*, 5 (August), 264–70.

Belk, Russell W. (1974), "An Exploratory Assessment of Situational Effects in Buyer Behavior," *Journal of Marketing Research*, 10 (May), 156–63.

Blattberg, Robert C. and Subrata K. Sen (1974), "Market Segmentation Using Models of Multidimensional Purchasing Behavior," *Journal of Marketing*, 38 (October), 17–28.

Buchanan, Bruce and Donald G. Morrison (1987), "Are Direct Marketers Really 98 Percent Wrong?" paper presented at the ORSA/TIMS Joint National Meeting, St. Louis (October 25–28).

——— and ——— (1988), "A Stochastic Model of List Falloff With Implications for Repeat Mailings," *Journal of Direct Marketing*, 2 (Summer), 7–15.

Christen, Francois G. (1987), "Richness: A Way to Evaluate Segmentation Systems," paper presented at the Attitude Research Conference, West Palm Beach, FL (May 29).

DeSarbo, Wayne S. and J. Douglas Carroll (1985), "Three-Way Metric Unfolding Via Alternating Weighted Least Squares," *Psychometrika*, 50 (September), 275–300.

——— and Jaewun Cho (1989), "A Stochastic Multidimensional Scaling Vector Threshold Model for the Spatial Representation of 'Pick Any/n' Data," *Psychometrika*, 54 (March), 105–29.

Efron, Bradley (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans*, CBMS Monograph #38. Philadelphia: Society for Industrial and Applied Mathematics.

Frank, Ronald E., William F. Massy, and Harper Boyd (1967), "Correlates of Grocery Product Consumption Rates," *Journal of Marketing Research,* 4 (May), 184–90.

———, ———, and Yoram Wind (1972), *Market Segmentation.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Gifi, Albert (1990), *Nonlinear Multivariate Analysis.* Chichester, England: John Wiley & Sons Ltd.

Green, Paul E. (1973), "On the Analysis of Interactions in Marketing Research Data," *Journal of Marketing Research,* 10 (November), 410–20.

Grover, Rajiv and V. Srinivasan (1987), "A Simultaneous Approach to Market Segmentation and Market Structuring," *Journal of Marketing Research,* 24 (May), 139–53.

Henry, Walter A. (1976), "Cultural Values Do Correlate With Consumer Behavior," *Journal of Marketing Research,* 13 (May), 121–7.

Kahle, Lynne R. ed. (1983). *Social Values and Social Change: Adaptation to Life in America.* New York: Praeger Publishers.

———, Sharon E. Beatty, and Pamela Homer (1986), "Alternative Measurement Approaches to Consumer Values: The List of Values (LOV) and Values and Life Style (VALS)," *Journal of Consumer Research,* 13 (December), 405–9.

Kamakura, Wagner and José Alphonso Mazzon (1991), "Values Segmentation: A Model for the Measurement of Values and Value Systems," *Journal of Consumer Research,* 18 (September), 208–18.

——— and Thomas P. Novak, (1992), "Value-System Segmentation: Exploring the Meaning of LOV," *Journal of Consumer Research,* 19 (June), forthcoming.

Kotler, Philip (1988), *Marketing Management,* 6th ed. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Lutz, Richard L. and P. Kakkar (1974), "The Psychological Situation as a Determinant of Consumer Behavior," in *Advances in Consumer Research,* Vol. 2, Mary J. Schlinger, ed. Chicago: Association for Consumer Research, 439–53.

MacEvoy, Bruce (1989), *Descriptive Materials for the VALS2 Segmentation System,* Values and Lifestyles Program. Menlo Park, CA: SRI International.

McCann, John M. (1974), "Market Segment Response to the Marketing Decision Variables," *Journal of Marketing Research,* 11 (November), 399–412.

Mitchell, Arnold (1983), *The Nine American Life Styles.* New York: Warner.

Morrison, Donald G. (1969), "Conditional Trend Analysis: A Model That Allows for Nonusers," *Journal of Marketing Research,* 6 (August), 342–6.

Novak, Thomas P. and Bruce MacEvoy (1990), "On Comparing Alternative Segmentation Schemes: The List of Values (LOV) and Values and Life Styles (VALS)," *Journal of Consumer Research,* 17 (June), 105–9.

Rosekrans, Frank M. (1969), "Statistical Significance and Reporting Test Results," *Journal of Marketing Research,* 6 (November), 451–5.

Rosenthal, Robert and Donald B. Rubin (1979), "A Note on Percent Variance Explained as a Measure of the Importance of Effects," *Journal of Applied Social Psychology,* 9 (September–October), 395–6.

——— and ——— (1982), "A Simple, General Purpose Display of Magnitude of Experimental Effect," *Journal of Educational Psychology,* 74 (April), 166–9.

Saaty, Thomas L. (1959), *Mathematical Methods of Operations Research.* New York: McGraw-Hill Book Company.

Sawyer, Alan G. and Dwayne Ball (1981), "Statistical Power and Effect Size in Marketing Research," *Journal of Marketing Research,* 18 (August), 275–90.

Schaninger, Charles M. (1981), "Social Class Versus Income Revisited: An Empirical Investigation," *Journal of Marketing Research,* 18 (May), 192–208.

Schmittlein, David C. (1986), "Issues in Measuring Market Concentration Among Firms, Suppliers, and Customers," working paper, University of Pennsylvania.

———, Lee G. Cooper, and Donald G. Morrison (1990), "Truth in Concentration in the Land of 80/20 Laws," working paper, UCLA.

Singer, Eugene M. (1968), *Antitrust Economics.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Sleight, Peter and Barry Leventhal (1989), "Applications of Geodemographics to Research and Marketing," *Journal of the Market Research Society,* 31 (1), 75–101.

Steckel, Joel H., Wayne S. DeSarbo, and Vijay Mahajan (1991), "On the Creation of Acceptable Conjoint Analysis Experimental Designs," *Decision Sciences,* 22 (Spring), 435–42.

Veroff, Joseph, Elizabeth Douvan, and Richard A. Kulka (1981), *The Inner American.* New York: Basic Books.

Winter, Frederick W. (1984), "Market Segmentation: A Tactical Approach," *Business Horizons,* 27 (January–February), 57–63.

Yeaton, William H. and Lee Sechrest (1981), "Meaningful Measures of Effect," *Journal of Consulting and Clinical Psychology,* 49 (October), 766–7.

Reprint No. JMR292108