UCLA Series #55
1990

371

RICHNESS CURVES FOR EVALUATING MARKET SEGMENTATION

Thomas P. Novak
Cox School of Business
Southern Methodist University

Jan de Leeuw
Departments of Psychology & Mathematics
UCLA

Bruce MacEvoy
SRI International

May, 1991

The authors thank Lee Cooper, Donna Hoffman, Don Lehmann, Joel Steckel, three anonymous Reviewers and the Editor for their comments, and thank Jin Han for computational assistance.

ABSTRACT

Overall evaluation of market segmentation is often based upon a variance-explained measure, such as R^2. However, variance explained measures can be misleading. For example, when considering product usage, deviations of users and nonusers about the overall mean contribute equally to variance explained, even though the marketer will typically be interested in using the segmentation scheme to target only users. In addition, widely different patterns of product usage can have identical amounts of variance explained. As an alternative to variance explained measures, this paper proposes and develops the richness curve as a more effective way of evaluating market segmentation. Statistical considerations involving the bias and stability of richness curves are also addressed.

INTRODUCTION

Managers often face the problem of comparing two or more market segmentation schemes. We define a market segmentation scheme as a mutually exclusive partitioning of consumers into one of S market segments, and assume that several alternative segmentation schemes have been identified. For example, consider comparing the predictive ability of two schemes - such as income and social class (Schaninger 1981), or Values and Lifestyles (VALS) and the List of Values (LOV) (Kahle, Beatty & Homer 1986; Novak & MacEvoy 1990) - with respect to a set of consumer behaviors. In such problems concern lies with evaluating the effectiveness of previously identified segmentation schemes, and not with developing segmentation schemes.

In this paper, we evaluate market segmentation for the strategic purpose of forming target segments from an existing segmentation scheme. Our attention is directed toward combining existing market segments to form a target segment which is optimal with respect to a given consumer behavior (e.g., product usage). The target segment problem implies an *asymmetric* evaluation, since we are only interested in the presence of a consumer behavior, such as product usage, and not its absence. Segments containing nonusers are simply not of interest. This is typically the case in applications of a controlled coverage strategy (Frank, Massy & Wind 1972), such as the use of geodemographics to identify product users (e.g. Sleight & Leventhal 1989) in direct marketing. In addition, a target segment orientation is often taken in marketing problems involving customer self-selection, since marketing communications must be tailored to the intended target segment.

The target segment orientation may be contrasted with profiling the pros and cons of each of the market segments, with respect to a series of consumer behaviors (e.g., Darden & Perreault 1976). The segment profile problem implies a *symmetric* evaluation, since both presence and absence of a series of consumer behaviors contribute to developing an overall understanding of a market segment. For example, the Achiever segment in the VALS-2 typology (MacEvoy 1989) is relatively more likely to belong to business organizations, to exercise, and to drink imported beer; however, this segment is relatively less likely to belong to civic and environmental organizations, to do risky sports, and drink coffee. A richer understanding of the Achiever segment is provided by understanding activities and behaviors that are both engaged in and not engaged in.

The extent of differences among market segments in consumer behavior in general, and product usage in specific, is often evaluated with symmetric measures of association or variance explained. Of these, the popular choice in the marketing literature has been R^2, which has seen extensive use as an evaluative criteria (see for example, Rosekrans 1969; Frank, Massy & Wind 1972; Green 1973; Belk 1974; Lutz & Kakkar 1974; Sawyer & Ball 1981; Kahle, Beatty & Homer 1986; Novak & MacEvoy 1990; Kamakura & Mazzon 1991). However, while variance explained measures such as R^2 do provide an overall index of the "goodness" of each segmentation scheme, they do not reflect the fact that a given segmentation scheme can be used in a variety of ways, each requiring a separate evaluation.

As an example, suppose that a firm wishes to identify target segments for two advertising campaigns for a market in which 10% of consumers regularly drink apple juice. The market has been segmented by product benefits (which explain 7.5% of the variance in apple juice usage) and by consumer lifestyles (which explain 2.3% of the variance). It appears that product benefits is the superior segmentation scheme, since the benefit segments explain more variance in product usage. However, this evaluation ignores the strategic purpose of the segmentation, which is to form target segments.

Suppose the first advertising campaign is directed toward a niche target of health enthusiasts, which is only 5% of the beverage market. Therefore, a relatively small target size will be desired. Further suppose that a segment defined by product benefits, containing 5% of the market, can be identified in which 37% of consumers regularly drink apple juice. With lifestyles, the best we can do is to create a 5% target in which 25% of consumers regularly drink apple juice. In this case, product benefits are superior to lifestyles for niche targeting. However, assume that the second advertising campaign is addressed toward a mass, family-oriented target of 45% of the market. Using product benefits, a target containing 45% of the sample can be constructed, of which 15% are users; a 45% lifestyles target has 19% users. The mass target favors the lifestyles segmentation scheme. Clearly, in this case the choice of a segmentation scheme is contingent upon the definition of the marketing problem.

Besides the specific measure of R^2, consider three broad approaches to evaluating market segmentation. First, some approaches focus upon evaluating target segments within schemes. Many standard managerial considerations for selecting target segments have been proposed (see for example Kotler 1988, Wilkie & Cohen 1977): segment homogeneity, parsimony, identifiability, reachability, size, spillover effects, responsiveness to marketing mix, nature of competitive activity within the segment, etc. Such considerations are useful for a multidimensional (although somewhat informal and subjective) evaluation of the pros and cons of alternative target segments. A more objective "segmentation coefficient" was proposed by McCann (1974). The segmentation coefficient is the product of relative segment size, relative level of demand, and relative response rate, is obtained separately for advertising, price and deal response rate, and is used to evaluate the potential usefulness of each segment within a scheme.

Second, at the other extreme, one can evaluate segmentation schemes as a whole. At the scheme level, Frank, Massy & Wind (1972) note that either behavioral or decision-oriented (normative) considerations may be of primary concern. The normative school focuses on the extent to which predictions from a segmentation model are effective and the degree to which resource allocation is improved by taking account of intersegment differences. By definition, normative evaluation is contingent upon the marketing problem. The behavioral school is more closely related to our particular problem, in that it attempts to identify & document generalizable differences among consumer groups. As an example, Schaninger (1981) uses significance tests in ANOVA & MANOVA to document differences among income and social class groups, and Currim (1981) and Gensch (1985) suggest that a segmentation scheme be evaluated based upon whether it increases the predictive accuracy of a choice model. And as noted, many authors have used R^2 to index segment heterogeneity.

A third approach is to evaluate a dichotomous selection criterion variable for its utility and costs to an institution (Cronbach & Gleser, 1965). While developed outside of marketing, this approach has direct relevance to market segmentation, in particular to problems involving target selection. For example, early research in the personnel testing literature (Taylor & Russell 1939; Thurstone 1931) examined the proportion of employees (read, "consumers") who were likely to be "satisfactory" (read, "purchasers"), before and after selection by means of a test (read, "in a target segment or in the unsegmented market"). This approach assumes that the utility of positive outcomes (purchasers) is greater than the utility of negative outcomes (non-purchasers). While this asymmetry should be reflected in evaluating target segments, symmetric measures such as R^2 weight each consumer equally and can provide misleading results.

Our approach develops a new evaluation tool that flexibly merges the specific emphases of these three traditions - a general tool for evaluating a segmentation scheme that permits assessment of one or more individual segments as target markets of a specific utility. We propose a method of evaluation based upon 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 effectively used to target consumers, then 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 richness curves, then form a simple and direct way of evaluating segmentation schemes. However, since richness curves are based upon 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 will describe how this 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 spent, etc., 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., Frank, Massy and Boyd 1966; Andreason 1966; Bass, Tigert and Lonsdale 1968; Assael and Roscoe 1976; Henry 1976; Schaninger 1981; Grover & Srinivasan 1982; Kahle, Beatty & Homer 1986; Novak

& MacEvoy 1990; Kamakura & Mazzon 1991). As noted by Schmittlein (1986), early work in the marketing literature (Church 1946; Twedt 1964) dealing with the degree of concentration of consumer purchases led to the search for the "heavy user" segment in behavior-based *a priori* market segmentation (Wind 1978). It is always worthwhile for a firm to "know its customers" - to be able to identify, using market segmentation, who the users are.

Of course, we recognize that there are limitations to focusing upon 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, then it may not be profitable to direct resources to the segment containing most of the users, since this 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 these segments also will not be profitable. Numerous researchers have cautioned against defining target segments based upon consumption differences, without considering response to marketing mix variables (McCann 1974; Blattberg & Sen 1974; Winter 1984). Since usage rates and segment sizes can be adjusted for accessibility, response to marketing mix, or identifiability, many of these limitations can be dealt with, although we do not explicitly deal with them in this paper.

In the following sections we define and discuss the richness curve, and compare it to the closely related Lorenz curve (Singer, 1968). Next, 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. However, 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*, since the differentiation among scheme *A* segments is at the highest usage rates, while scheme *C* is among the lowest.

----- Insert Table 1 here -----

The difficulty in using R^2 for evaluating market segmentation schemes is that it does not reflect the marketing context. In obtaining R^2, all deviations of segment means about the overall mean are considered to be equally important. This is not the normal marketing context, however: marketers will generally focus on only a few segments rather than all segments, and their concern is how effective targeting is within these 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 shows 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, if the marketer targets the segment having the highest concentration of users in Figure 1, he or she will get a return of 74% from the "richest" segment; as the marketer attempts to reach a larger part of the population by adding in 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.

----- Insert Figure 1 here -----

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

Besides providing a context-contingent evaluation of market segmentation, richness curves provide a more interpretable evaluation than variance explained measures. As noticed by Bass, Tigert & Lonsdale (1968) and Rosenthal & 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 have referred to as a "meaningful measure of effect" (Yeaton & Sechrest 1981).

The "base rate," r, indicating the proportion of the entire sample of $n=400$ consumers that uses the hypothetical product, is the richness of the unsegmented market, and is indicated by a dashed horizontal line in Figure 1. This figure (42% in Figure 1) indicates the return the marketer can expect from not segmenting the market, but appealing to a random sample of consumers. Each of $S=8$ 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) $$r_t = \Sigma_{j=1,i} u^*_j / i$$

where: $i = \mathrm{RND}(nt)$

$u^*_i = u_s$ if i^{th} consumer $\in$ s^{th} segment,

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

There has been recent interest in marketing (Schmittlein 1986; Buchanan & Morrison 1987; Schmittlein, Cooper & Morrison, 1990) in applications of measures of market concentration - in particular the Lorenz curve and Gini coefficient (see, for example, Singer 1968) - that were developed in industrial organization to problems involving measurement of customer concentration. We next contrast the richness curve with the Lorenz curve, since the latter has also been discussed in a segmentation context and provides an alternate way to proceed. While the richness curve specifies the proportion of respondents in a target of size t who are users of a product, the "modified" (Buchanan & Morrison, 1987) Lorenz curve[3] specifies the percent 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:

$$L_t = tr_t/r \quad . \tag{2}$$

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

[2]The probabilities u^*_j 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. Either the person uses the product or they do not. We do not have a true purchase probability, estimated, say, from a string of scanner data. If we *had* such data, we could model purchase probabilities with the beta-binomial distribution (Buchanan & Morrison, 1988; Buchanan & Morrison, 1987). Or, following Schmittlein, Cooper & 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 horizonal axis from lowest to highest; our modification orders consumption from highest to lowest, for comparability with the richness curve.

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.

----- Insert Figure 2 here -----

The modified Lorenz curve, like the richness curve, is contingent upon the size of market targeted and is easily 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" - i.e. 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, on the other hand, should be considered if attention is specifically to be directed to the problem of reaching or capturing a fixed proportion of all consumers. Since 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 to the base rate). In Figure 2, diminishing returns are obtained in the sample after including 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. At the same time, the superiority of segment 1 is not as clearly shown as in the richness curve.

Overall Measures for Comparing Segmentation Schemes.

The Gini coefficient is an overall measure of concentration based upon the Lorenz curve (Schmittlein, Cooper & Morrison, 1990; Schmittlein, 1986), which has direct application as an evaluative index for market segmentation. The "modified Gini coefficient," g, (Buchanan & Morrison, 1987) is an overall measure which 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 is $g_A = .280$, $g_B = .323$, and $g_C = .281$. Thus, the modified Gini coefficient suggests that scheme A performs the worst. Again, inspection of Table 1 makes the practical utility of this conclusion questionable.

An alternate overall measure based upon the richness curve is simply the area between the richness curve and the horizontal base rate line. This 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 regarding Table 1.

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

$$\text{(3)} \qquad \begin{aligned} \text{ARG} &= \int_0^1 (r_t - r) \\ g &= \int_0^1 (t(r_t/r - 1)) \end{aligned}$$

While ARG is the mean increase in richness, the modified Gini coefficient is the average proportionate increase in richness weighted by the target size. While 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 correspond to target sizes of managerial interest. Since it is unlikely that a manager will target 80% or more of the market

---

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

[5]The overall measure based upon 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.

place, why should these target sizes affect the overall index? If overall indices are used, they should be conditional upon 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 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. This section examines the problem, and describes bootstrap methods used 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, those 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 naively 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, this is clearly wrong. But, with only one sample to base estimates on, we cannot recognize this mistake.

This mistake means that the rank ordering of segments in the sample based upon 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 misguided information 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, and 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 upon 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.

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

$$(4) \qquad \underline{\mathbf{r}} = \phi(\underline{\mathbf{p}})$$

We then generate J=200 bias-adjusted "pseudo-richness curves" (Gifi, 1983), $\tilde{\underline{\mathbf{r}}}_j$:

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

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

$$(6) \qquad \tilde{\underline{\rho}} = \Sigma \tilde{\underline{\mathbf{r}}}_j/J \ .$$

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

Figures 3a through 3c 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. Figure 3d superimposes the adjusted curves for the three schemes. Clear differences among the three schemes can be seen. In Figures 3a and 3b, curves for schemes A and B exhibit minimal bias, while Figure 3c 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 Figure 3d, scheme A clearly outperforms scheme B for target sizes, $t<.3$, while scheme B outperforms A for $t>.3$. If a niche strategy is to be pursued, then scheme A is preferable, while as the size of the target segment approaches the base rate, scheme B is preferred.

----- Insert Figures 3a-3d here -----

It is important to note that the richness curve makes it 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, on the other hand, facilitate a contingent evaluation of market segmentation schemes.

Factors That Contribute To Bias in Richness Curves

Figures 4A through C show the extent of bias in a sample richness curve for our hypothetical example. But, taking a broader context, how large is the bias in a sample richness curve, and what factors affect the degree of bias and stability? A Monte Carlo analysis was performed to address these questions. Table 2 presents the levels of the seven factors that were investigated in the Monte Carlo analysis. The first three factors - 1) number of segments, 2) sample size, and 3) base rate -are self-explanatory. Factor 4 specifies the range of segment sizes. Three conditions were used: A) all segments have equal sizes, B) two different segment sizes, with the larger segments three times the size of the smaller, and C) two different sizes, with the larger eight times the smaller. Factor 5 specifies whether the larger or smaller segments have relatively higher or smaller 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.

----- Insert Table 2 here ---

Since 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 & Cho, 1989; DeSarbo & Carroll, 1985 for similar analyses)[6]. These 27 trials appear in Table 3.[7] Given the constraints in Table 2, each of the 27 trials uniquely specify 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:

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

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}, were also obtained. Results for the 27 trials appear in Table 3. As can be seen from Table 3, some trials produce bias up to .10; clearly there are certain situations where bias adjustments are crucial.

---

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

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

----- Insert Table 3 -----

Results of main-effect ANOVAS for five mean bias dependent variables and five standard error variables, and the marginal means on these variables appear 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, while 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, while for target sizes t equal to .1 and .5, Factor 6 (a small range of segment usage rates) is related to bias. Practically speaking, these results indicate that unless the characteristics of the data (few segments, high sample size, top skew, etc.) suggest otherwise, bias adjustment should be routinely performed.

----- Insert Table 4 -----

Standard error, on the other hand, is significantly related 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). While not as dramatic as the differences among 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. A careful eye will note 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 $\phi(\mathbf{y}_j/n)$, 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 shown 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 using expressions (5) and (6), based upon $J=50$ resampled tables. The nomonotonicity observed in Figures 3A-C canceled out when we aggregated over $T=50$ observed samples. Thus, while 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 Monte Carlo simulations, using 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 was obtained based upon 25 resampled tables. Other than minor fluctuations likely due to sampling error, there was no evidence of nonmonotonicity in the average adjusted richness curves obtained for the 27 trials.

Adequacy of the bootstrap adjustment. A further issue is how much of the bias in the observed richness curve is corrected by the bootstrap adjustment. The 27 trial Monte Carlo simulation just described was also used to address this issue. For each of the 25 sample contingency tables, the average deviation of the sample richness curve from the known population curve, and the average deviation of the adjusted bootstrapped curve from the known population curve were obtained. Results appear in Table 5.

----- Insert Table 5 here -----

The five columns on the right side of Table 5 show the degree of bias in the observed sample richness curve, 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; 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 only by .0291.

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 & MacEvoy, 1990). The consumer behavior we are interested in 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 developed at the University of Michigan Survey Research Center (Kahle 1983; Veroff, Douvan, and Kulka 1981), based upon 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%.

[8]While the results in Table 5 are based upon 25 sample tables, and those in Table 4 are based upon 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 only used 25 replications in the bootstrap.

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 ranged from .12% to 6.7%.

As a basis of comparison, adjusted R^2 values for the four schemes are as follows: DuoVALS2=.076, VALS2=.085, VALS=.060, LOV=.013. Thus, R^2 suggests that 1) VALS2 is superior, 2) there isn't that much difference between VALS2, DuoVALS2, and VALS, and 3) LOV performs poorly.

Our marketing objective is to form a target segment, based upon each of the four schemes, which contains the greatest proportion of wine drinkers. Since 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 this 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 due to the very large number of segments. In such situations, the pattern of nonmonotonicity indicates caution should be taken when interpreting the front end of the curve. This is reinforced by comparing standard deviations of the adjusted curves for DuoVALS2 and VALS2. For t=.01, the standard deviation for DuoVALS2 is .087, and for VALS2 is .057; for t=.10 the values are .047 for DuoVALS2 and .050 for VALS2. The richness curve for DuoVALS2 is more variable than for VALS2 for very small (t<.10) target sizes.

----- Insert Figure 4 here -----

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-square (.013) disguises this fact and makes LOV appear nearly useless. The richness curves make explicit how each scheme performs at markets of different size. The superiority of VALS2 extends to a target segment of about 30% of the 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 thought of as an upper bound. By straightforward multiplication of the total richness, or net increment over base rate, times the size of the market, differences among the schemes can be converted 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 using each segmentation scheme.

DISCUSSION & CONCLUSIONS

While richness curves closely resemble Lorenz curves, our work is conceptually distinct from recent marketing applications of the Lorenz curve (Schmittlein, Cooper & Morrison, 1990; Schmittlein, 1986). These applications focus upon identifying 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 pre-defined market segments and identify a segment-level richness curve as a way of quantifying the degree to which the segmentation can be used to effectively target consumers.

Although we emphasized that measures of variance explained can be misleading if used to evaluate market segmentation, this should not be taken to imply that variance explained measures should never be used. If interest lies with the "segment profile problem" rather than the "target segment problem," then measures of variance explained do provide a convenient, overall measure of magnitude of effect.

It is important to mention what we *have not* considered in evaluating market segmentation schemes. While we have focused upon establishing the magnitude of actionable differences in segments, other aspects of the segmentation must be evaluated as well, such as segment stability and validity, segment responsiveness to the marketing mix, and segment accessibility. However, as it is possible for these characteristics to be evaluated with respect to their impact upon 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.

While we have focused upon binary criterion variables, the bootstrap can be used with continuous variables by resampling from the observed sample distribution. Continuous variables open up the possibility of replacing a binary usage variable with an index that incorporates considerations such 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 upon the linear programming formulation in Appendix A, need to be investigated. Since the bootstrap does not guarantee a monotone decreasing curve (although this problem disappears when one aggregates over multiple samples, and although the bootstrap does correct for the majority of 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, there are further applications of richness curves. Richness curves can be used to compare hierarchically nested market segments, as a way of determining the number of segments to retain. The impact upon richness of using four through 10 nested segments can be clearly seen. Further, effective graphical representation facilitates a quick and accurate comparison of segmentation schemes. Richness curves would be useful as part of a decision support system, permitting graphical 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. But, 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

In this section we develop a rigorous definition of the richness curve. We have an $S\times 2$ table, with rows of the table indicating S existing market segments, and the columns the dichotomy user/non-user. Now consider an arbitrary segment, τ_t, containing the proportion t of users, representing a target segment which may or may not be one of the S existing segments. Then, the size of the arbitrary target is:

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

and the proportion of users in the target is:

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

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

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

or equivalently:

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

The richness curve, $r_t = \max\{p(\text{user}\,|\,\tau_t)\}$, is defined as the set of target segments, $\{\tau_t\}$, with maximal usage rates. Substituting expression (A3) into expression (A2), and using the constraint implied by expression (A1), this interpretation of the richness curve can be formalized 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: $\delta_s = \mathrm{p}(\tau_t\,|\,\text{segment}=s)$

$u_s = \mathrm{p}(\text{user}\,|\,\text{segment}=s)$

$p_s = \mathrm{p}(\text{segment}=s)$.

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

The solution to the linear program is as follows. We 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. Thus, $p_{[1]}$ is the size of the segment with the largest u_s, i.e., the segment with usage rate $u_{[1]}$, and so on. Now,

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

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, since 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 may be studied in future work.

Appendix B

Nonmonotonicity of adjusted richness curves

Figures 3 and 4 exhibit nonmonotonic adjusted bootstrapped richness curves. First, note that the observed sample richness curve, $\phi(\underline{p})$, and the j^{th} bootstrap resampled richness curve, $\phi(\underline{v}_j/n)$, will both be monotonically decreasing since the richness function, $\phi()$, is by definition monotonically decreasing. But the j^{th} bias-adjusted pseudocurve, $\tilde{\underline{r}}_j = 2\phi(\underline{p}) - \phi(\underline{v}_j/n)$, may not be monotonically decreasing, since 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 that $h(t) = 2f(t) - g(t)$. Then, taking derivatives:

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

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

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

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

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

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 $g'(t) < 0$, then the adjusted curve $h(t)$ will be increasing.

REFERENCES

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

Andreason, A.R. (1966), "Geographic Mobility and Market Segmentation," *Journal of Marketing Research*, 3 (November), 341-350.

Assael, H. & A.M. Roscoe, Jr. (1976), "Approaches to Market Segmentation Analysis," *Journal of Marketing*, *40* (October), 67-76.

Bass, F.M., D.J. Tigert, & R.T. Lonsdale (1968), "Market Segmentation: Group Versus Individual Behavior," *Journal of Marketing Research, 3(3).*

Belk, R.W. (1974), "An Exploratory Assessment of Situational Effects in Buyer Behavior," *Journal of Marketing Research, 10,* 421-427.

Blattberg, R.C. & S.K. Sen (1974), "Market Segmentation Using Models of Multidimensional Purchasing Behavior," *Journal of Marketing, 38,* 17-28.

Buchanan, B. & D.G. Morrison (1987), "Are Direct Marketers Really 98 Percent Wrong? Paper presented at the ORSA/TIMS Joint National Meeting, October 25-28, St. Louis.

Buchanan, B. & D.G. Morrison (1988), "A Stochastic Model of List Falloff With Implications for Repeat Mailings," *Journal of Direct Marketing, 2(3)*, 7-15.

Christen, F.C. (1987), "Richness: A Way to Evaluate Segmentation Systems," Paper presented at the *Attitude Research Conference,* West Palm Beach, Florida, May 29.

Church, D.E. (1946), "The Consumer Market for Canned Vegetables, Fruits and Juices," *Journal of Marketing, 1,* 44-54.

Cronbach, L.J. & G.C. Gleser (1965), *Psychological Tests and Personnel Decisions*, Urbana, IL: University of Illinois Press.

Currim, I.S. (1981), "Using Segmentation Approaches for Better Prediction and Understanding From Consumer Mode Choice Models. *Journal of Marketing Research*, *18*, 301-309.

Darden, W.R. & W.D. Perreault, Jr. (1976), "Identifying Interurban Shoppers: Multiproduct Purchase Patterns and Segmentation Profiles," *Journal of Marketing Research, 13* (February), 51-60.

DeSarbo, W.S. & & J.D. Carroll (1985), "Three-Way Metric Unfolding Via Alternating Weighted Least Squares. *Psychometrika, 50,* 275-300.

DeSarbo, W.S. & J. Cho (1989), "A Stochastic Multidimensional Scaling Vector Threshold Model for the Spatial Representation of 'Pick Any/n' Data," *Psychometrika, 54(1),* 105-129.

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

Frank, R.E., W.F. Massy & H. Boyd (1966), "Correlates of Grocery Product Consumption Rates," *Journal of Marketing Research, 4* (May), 184-190.

Frank, R.E., W.F. Massy & Y. Wind (1972), *Market Segmentation,* Englewood Cliffs, N.J.: Prentice-Hall.

Gensch, D.H. (1985), "Empirically Testing a Disaggregate Choice Model for Segments," *Journal of Marketing Research, 22,* 462-467.

Gifi, A. (1981), *Non-Linear Multivariate Analysis*, Leiden, The Netherlands: DSWO Press.

Green, P.E. (1973), "On the Analysis of Interactions in Marketing Research Data," *Journal of Marketing Research, 10,* 410-420.

Grover, R. & V. Srinivasan (1987), "A Simultaneous Approach to Market Segmentation and Market Structuring," *Journal of Marketing Research, 24* (May), 139-153.

Henry, W.A. (1976), "Cultural Values Do Correlate With Consumer Behavior," *Journal of Marketing Research, 13* (May), 121-127.

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

Kahle, L.R., S.E. Beatty & P. Homer (1986), "Alternative Measurement Approaches to Consumer Values: The List of Values (LOV) and Values and Life Style (VALS)," *Journal of Consumer Research, 13*, 405-409.

Kamakura, W. & J.A. Mazzon (1991), "Values Segmentation: A Model for the Measurement of Values and Value Systems," *Journal of Consumer Research,* (forthcoming).

Kamakura, W. & Novak, T.P. (1991), "Value-System Segmentation: Exploring the Value of LOV," *working paper.*

Kotler, P. (1988), *Marketing Management*, 6th Edition, Englewood Cliffs, NJ: Prentice Hall.

Lutz, R.L. & 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, 439-453.

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

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

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

Morrison, D.G. (1969), "Conditional Trend Analysis: A Model That Allows for Nonusers," *Journal of Marketing Research, 6,* 342-346.

Novak, T.P. & B. MacEvoy (1990), "On Comparing Alternative Segmentation Schemes: The List of Values (LOV) and Values and Life Styles (VALS)," *Journal of Consumer Research*, in press, June.

Rosekrans, F.M. (1969), "Statistical Significance and Reporting Test Results," *Journal of Marketing Research, 6,* 451-455.

Rosenthal, R. & D.B. Rubin (1979), "A Note on Percent Variance Explained as a Measure of the Importance of Effects," *Journal of Applied Psychology, 9(5),* 395-396.

Rosenthal, R. & D.B. Rubin (1982), "A Simple, General Purpose Display of Magnitude of Experimental Effect," *Journal of Educational Psychology, 72(2)*, 166-169.

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

Sawyer, A.G. & A.D. Ball (1981), "Statistical Power and Effect Size in Marketing Research," *Journal of Marketing Research, 18,* 275-290.

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

Schmittlein, D.C. (1986), "Issues in Measuring Market Concentration Among Firms, suppliers, and customers," Working paper.

Schmittlein, D.C., L.G. Cooper & D.G. Morrison (1990), "Truth in Concentration in the Land of 80/20 Laws," Working paper.

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

Sleight, P. & B. Leventhal (1989), "Applications of Geodemographics to Research and Marketing," *Journal of the Market Research Society, 31(1)*, 75-101.

Steckel, J.H, W.S. DeSarbo & V. Mahajan (1991), "On the Creation of Acceptable Conjoint Analysis Experimental Designs," *Decision Sciences*, forthcoming.

Taylor, H.C. & J.T. Russell (1939), "The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection: Discussion and Tables," *Journal of Applied Psychology, 23*, 565-578.

Thurstone, L.L. (1931), *The Reliability and Validity of Tests*. Ann Arbor, MI: Edwards Brothers, Inc.

Twedt, D.W. (1964), "How Important to Marketing Strategy is the 'Heavy User?'" *Journal of Marketing, 28(Jan)*, 71-72.

Veroff, J., E. Douvan, and R.A. Kulka (1981), *The Inner American*, New York: Basic Books.

Wilkie, W.L. & J.B. Cohen (1977), "An Overview of Market Segmentation: Behavioral concepts and research approaches," Marketing Science Institute Research Program Working Paper No. 77-105.

Wind, Y. (1978), "Issues and Advances in Segmentation Research," *Journal of Marketing Research, 15*, 317-337.

Winter, F.W. (1984), "Market Segmentation: A Tactical Approach," *Business Horizons*, 57-63.

Yeaton, W.H. & L. Sechrest (1981), "Meaningful Measures of Effect," *Journal of Consulting and Clinical Psychology, 49(5)*, 766-767.

Table 1
Three Hypothetical Segmentation Schemes

Segment:	Proportion of consumers using brand *x* in each segment of... 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

Table 2
Factors for Monte 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 specifying 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:

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

b) Equal difference condition. Solve for u_s:

$$\begin{aligned} &\Sigma n_s u_s = m \\ &u_1 - u_2 = d(w_s/\Sigma w_s);\ w_s=1 \\ &u_2 - u_3 = d(w_s/\Sigma w_s);\ w_s=1 \\ &\vdots \\ &u_{S-1} - u_S = d(w_s/\Sigma w_s)\ w_s=1 \end{aligned}$$

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

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

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
F7	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≤.05
**	p≤.01
***	p≤.001
****	p≤.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.

Table 5
Marginal Means for Bootstrap Adjusted Bias and Sample Bias

Design Factor:	Factor Level:	Bias in bootstrap adjusted richness curve for target size: t=.1	t=.2	t=.3	t=.4	t=.5	Bias in observed sample richness curve for target size: 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

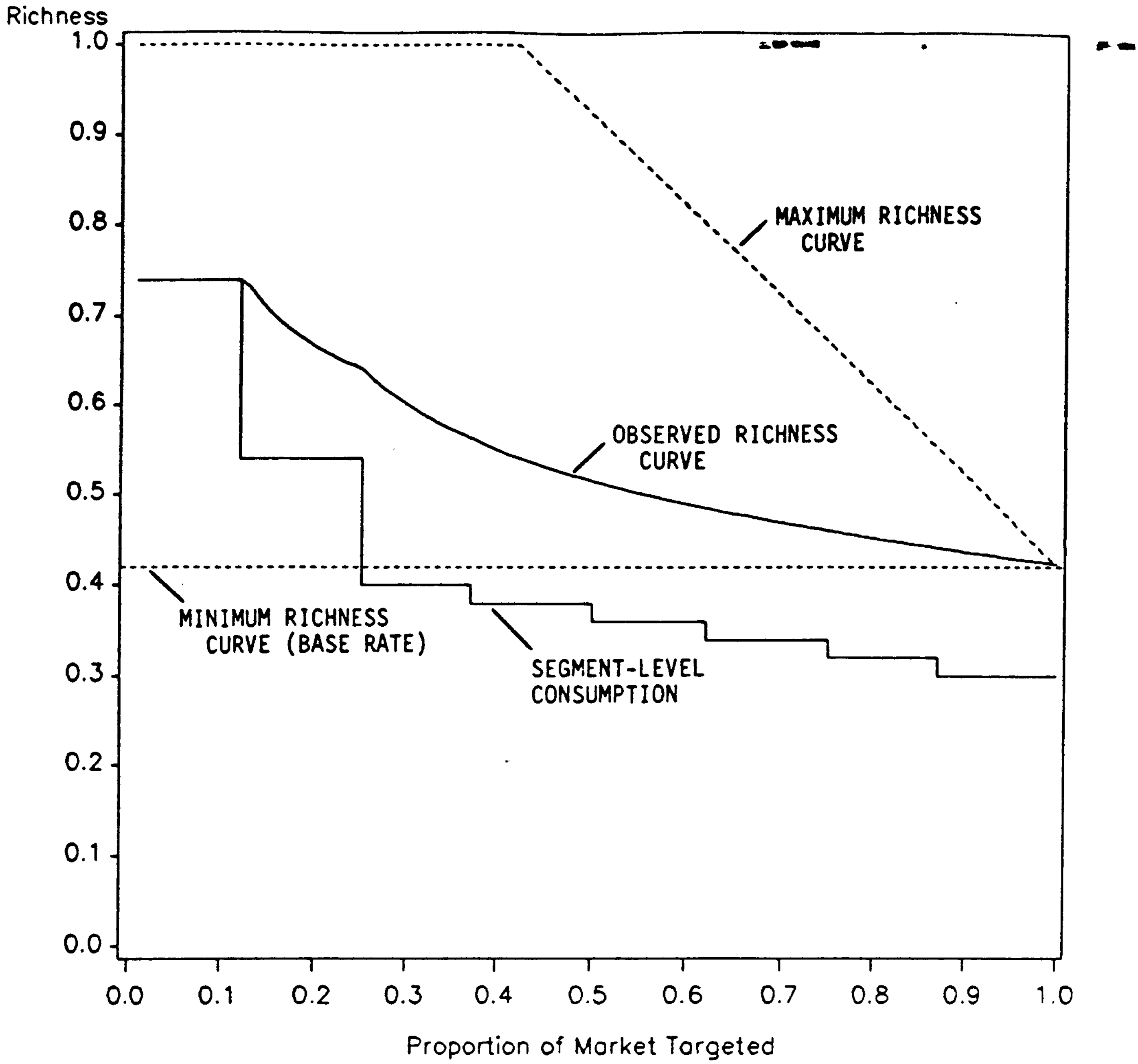


Figure 1

Observed Richness Curve for Scheme A

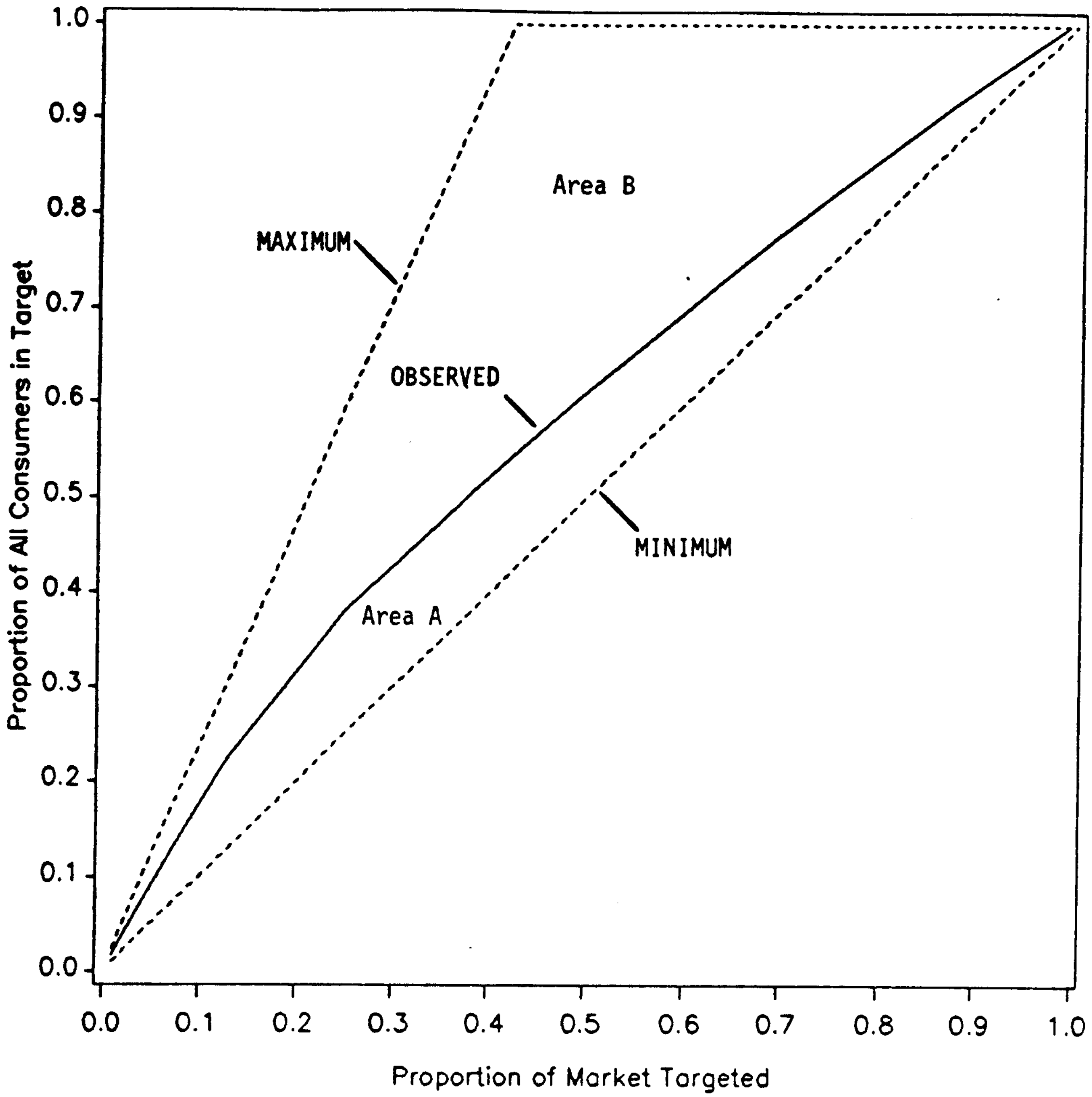
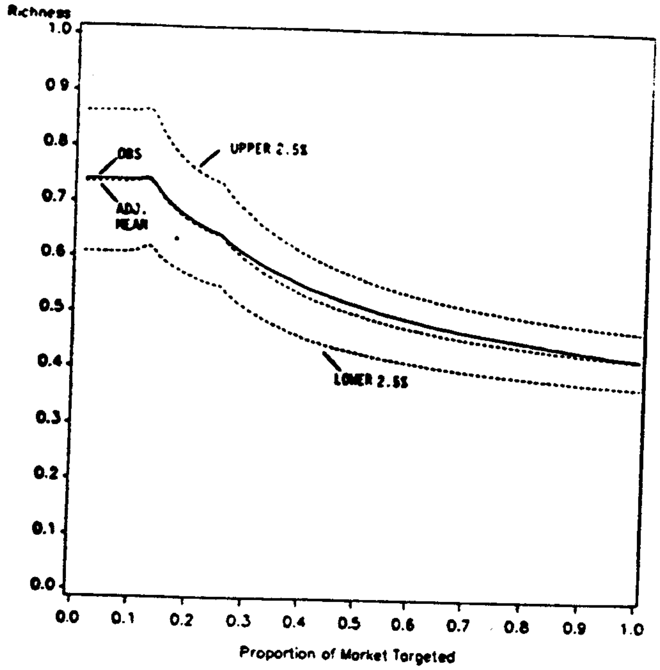
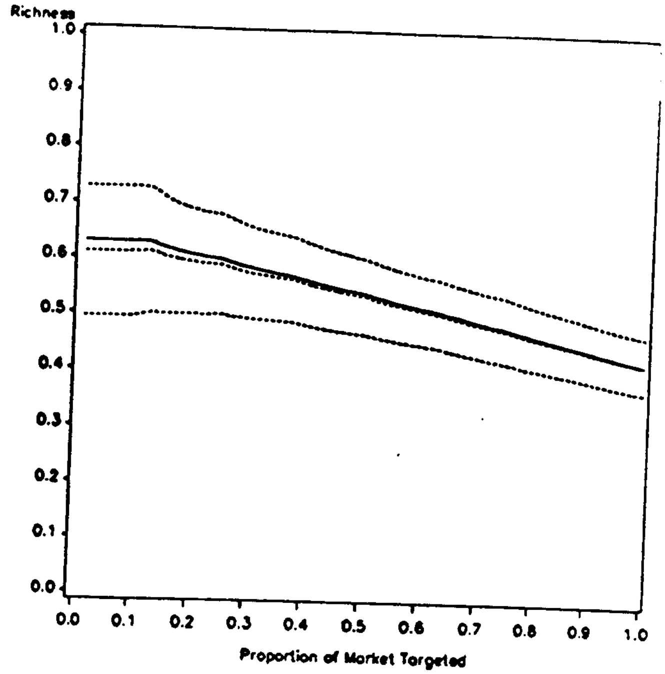


Figure 2

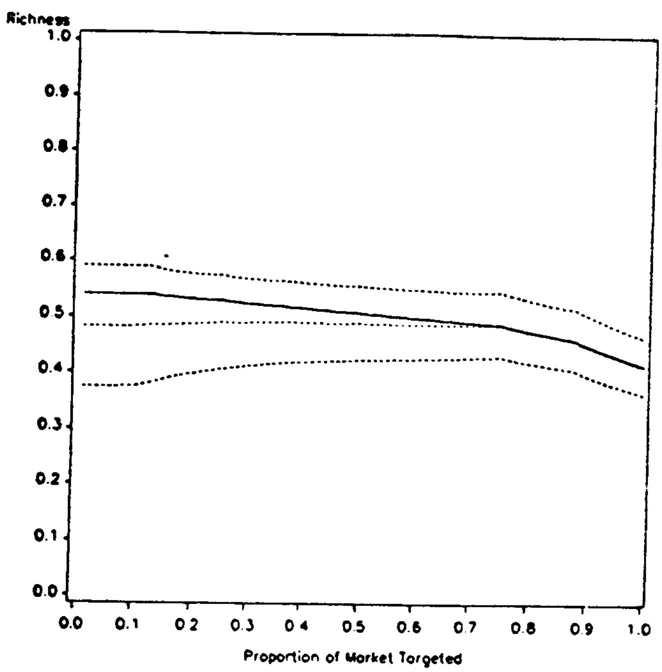
Observed Modified Lorenz Curve for Scheme A



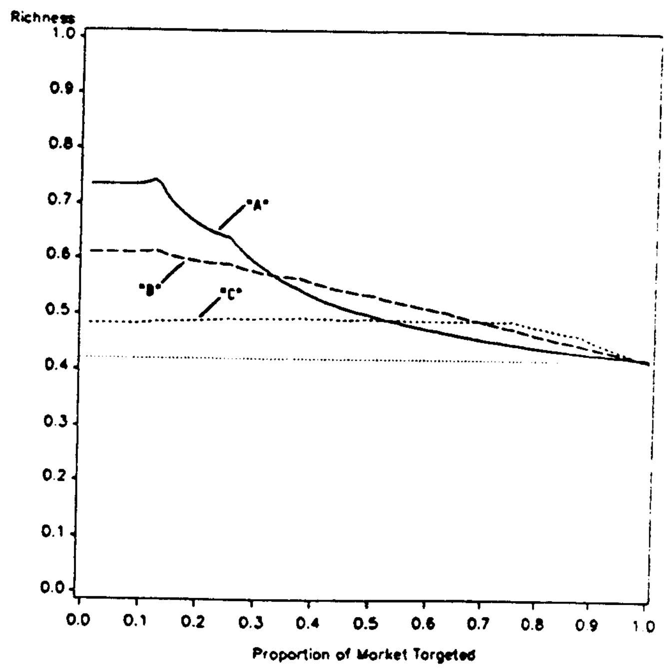
A) Bootstrapped Richness Curve for Scheme A



B) Bootstrapped Richness Curve for Scheme B



C) Bootstrapped Richness Curve for Scheme C



D) Adjusted Richness Curves for Schemes A, B, & C

Figure 3

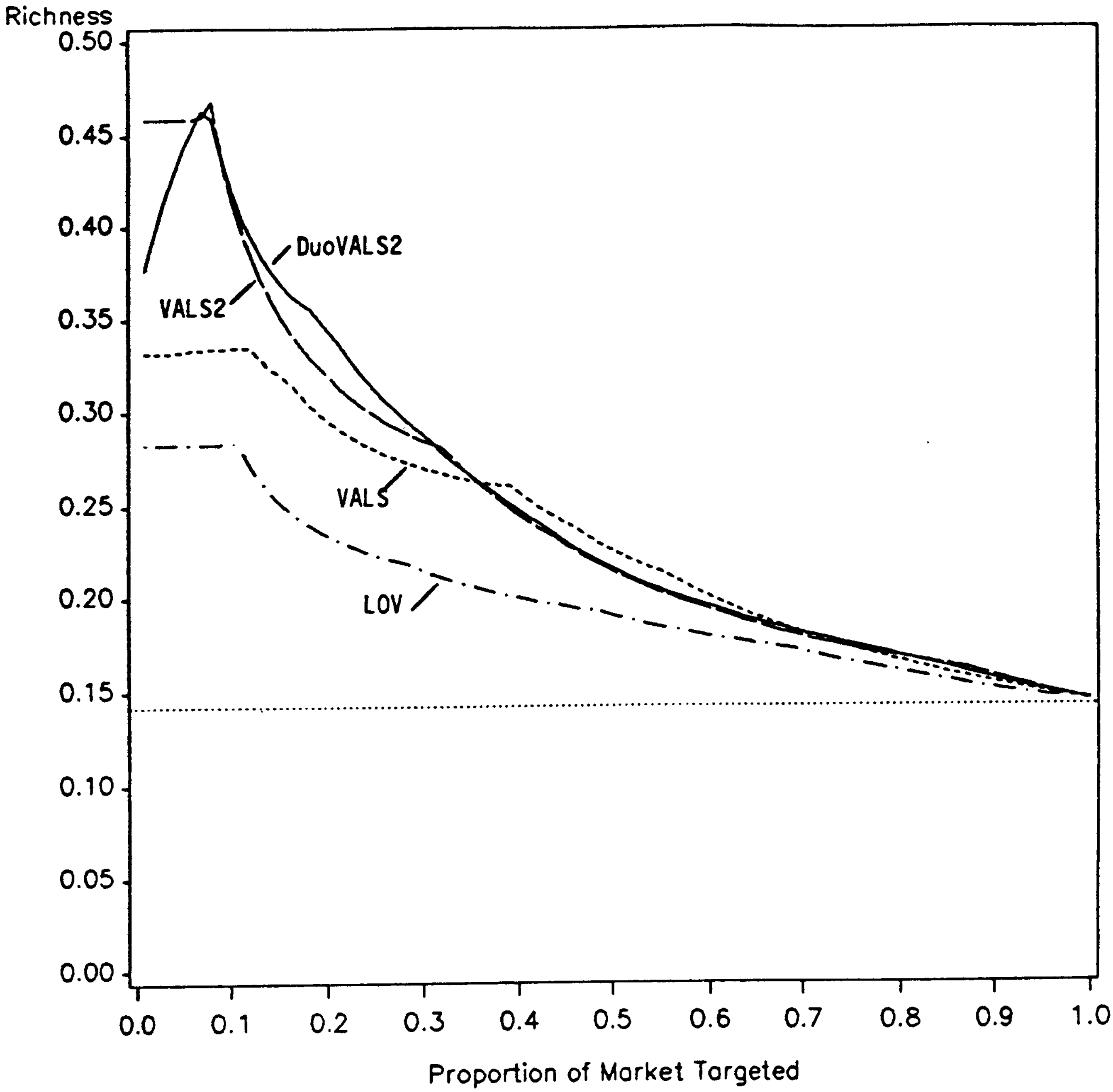


Figure 4

Wine With Dinner: Four Adjusted Richness Curves