A Theory of Market Segmentation
Author(s): Henry J. Claycamp and William F. Massy
Source: *Journal of Marketing Research*, Vol. 5, No. 4 (Nov., 1968), pp. 388-394
Published by: American Marketing Association
Stable URL: http://www.jstor.org/stable/3150263

HENRY J. CLAYCAMP and WILLIAM F. MASSY*

In this article a normative theory of market segmentation is developed that takes account of major implementation problems. The theory is presented as a multistage mathematical model of the full range of segmentation possibilities from the perfectly discriminating monopolist to the mass marketer. The theory's major implications for the philosophy and application of the market segmentation strategy are discussed in detail.

A Theory of Market Segmentation

INTRODUCTION

One of the most striking developments in marketing is the amount of interest shown in market segmentation strategy. Nearly every issue of major marketing journals includes discussions of it and its implications for marketing management.

Despite the attention devoted to this topic, little progress has been made in developing a normative theory of the segmentation process. Most articles on segmentation tend to be either general discussions of the basic concept or research reports showing differences in consumption patterns among specific consumer groups. The *strategy* of segmentation often seems to be roughly equated with the *act* of defining subparts of some total market. As a result, considerable controversy—and perhaps some confusion—exist about the strategy's implications for the optimal allocation of scarce marketing resources and the requirements for effective implementation.

This article takes a fresh look at the theory of market segmentation and its implications for marketing management. Although the article is intended as a contribution to marketing theory, we believe it also has implications for the practice of market segmentation. In the following sections we will (a) briefly review the theoretical foundations of the segmentation strategy and identify the major barriers to effective implementation, (b) develop a normative theory of the segmentation process that recognizes major implementation difficulties, and (c) investigate some of its operational implications.

* Henry J. Claycamp and William F. Massy are associate professors of marketing, Stanford University.

THE ECONOMIC THEORY OF SEGMENTATION

The concept of market segmentation was developed in economic theory to show how a firm selling a homogeneous product in a market characterized by heterogeneous demand could maximize profits.[1] The theory shows that optimal profits can be achieved if the firm uses consumers' marginal responses to price, i.e., price elasticities, to define mutually exclusive segments and sets price (or output) so that marginal profits achieved in each segment are equal. Extension of these results to include other marketing variables (besides price) is easily done.

Since there seems to be little doubt that marketers are interested in the segmentation concept because of its profit implications and because the economic theory model shows how the concept is related to profit maximization, it can be considered as an "ideal" or "optimal" approach.

Although optimal segmentation is a simple and appealing concept, at least four kinds of problems can be identified that make it exceedingly difficult to utilize:

1. problems of defining mutually exclusive market segments
2. problems of measuring response elasticities on a segment by segment basis
3. information constraints that affect the possibility of reaching segments selectively (i.e., the marketer ordinarily has only socioeconomic or demographic information about audiences reached by promotional media or areas covered by distribution outlets, and it is usually difficult to find relationships between these variables and marginal response differentials)

[1] See [7] for example.

Journal of Marketing Research,
Vol. V (November 1968), 388–94

4. institutional constraints that limit the ability to use existing means of reaching segments with the desired degree of price or promotional selectivity.

The economic theory of segmented markets does not deal with these problems. If the obvious advantages of optimal segmentation are to be realized, we need to extend the theory to include fundamental problems of implementation. The primary purpose here is to present an extension of the segmentation process as a mathematical model.

A MULTISTAGE THEORY OF MARKET SEGMENTATION

As indicated, our goal is to extend the classical microeconomic theory of market segmentation to take account of problems in defining segments, and the existence of institutional and informational constraints on managers' ability to design promotional strategies that will reach specified segments. The problem of using demographic, socioeconomic, or other "practical" variables to describe consumer groups will also enter the analysis. Although our model uses only price and promotional variables, it could be revised to include any of the marketing mix elements.

We assume a market with firms sufficiently decoupled such that strategies can be planned without direct reference to problems of possible competitive retaliation, at least in the short run and for some marketing variables. The analysis considers profit maximization strategies for a single product. The model will be developed in five stages, representing successively more aggregative and easier to apply approaches to market segmentation. Though each stage is best considered separately, we emphasize that we are not presenting five different models of market segmentation. Our "final" model is represented by Stage 4. Stage 5 shows how the procedures developed in Stages 1–4, if carried to the limit, lead logically to a nonsegmentation or "mass market" strategy.

Stage 1: Segmentation by Perfect Discrimination Among Customers

Suppose that a firm attempts to market its product to N customers, each with the demand function:

$$d_i = f_i(p_i, x_i), \qquad i = 1, \cdots N,$$

where p_i is the price and, x_i a vector of m nonprice promotional variable offered to the ith customer. If the unit cost of distribution (not including promotion) to the ith customer is c_i, the firm's gross revenue equation can be written as:

$$R = \sum_{i=1}^{N} (p_i - c_i)d_i = \sum_{i=1}^{N} (p_i - c_i)f_i(p_i, x_i).$$

The firm's cost equation, also easily defined, includes the costs of supplying and promoting the product:

$$C = g\left\{\sum_{i=1}^{N} d_i\right\} + \sum_{i=1}^{N} q_i,$$

where d_i is the product demand and q_i is the total cost of implementing the promotional package, denoted by x_i, both for the ith customer. The function $g\{\cdot\}$ is the typical cost function of manufacturing the product (a function of the sum of all customers' demands) and any fixed costs of operating the firm; it does not include distribution or promotion costs, which are handled elsewhere in the model. The cost equation is best rewritten as an explicit function of the controllable variables in the marketing mix, p_i and x_i. Let v_i^1 be the vector of per unit costs of promotion to the ith customer, so that $q_i = v_i^1 x_i$:

$$C = g\left\{\sum_{i=1}^{N} f_i(p_i, x_i)\right\} + \sum_{i=1}^{N} v_i^1 x_i.$$

Given revenues and costs, we can write the following profit equation for the firm:

$$\Pi = R - C = \sum_i (p_i - c_i)f_i(p_i, x_i) - g\left\{\sum_i f_i(p_i, x_i)\right\} - \sum_i v_i^1 x_i. \tag{1}$$

The firm's optimal marketing mix will be obtained when Equation 1 is maximized with respect to p_i and the elements of x_i (for $i = 1, \cdots N$).

By differentiating (1) partially with respect to each of the controllable marketing variables and setting the derivatives equal to zero,[2] we get the familiar decision rules:[3]

$$(2) \qquad \begin{aligned} &(p_i - c_i - MC)\frac{\partial f_i}{\partial p_i} = -f_i, && i = 1, \cdots N, \\ &(p_i - c_i - MC)\frac{\partial f_i}{\partial x_{ij}} = v_{ij}, && i = 1, \cdots N \\ & && j = 1, \cdots m; \end{aligned}$$

where MC is the cost function derivative with respect to total demand. This development is a direct descendent of the microeconomic model of the perfectly discriminating monopolist (see [7] for example), generalized to include both price and nonprice competitive variables. The $(m + 1)N$ equations must be solved to determine the optimal price-promotional mix for each market (here, an individual customer) supplied by the firm.

This model represents the market segmentation strategy in its extreme form. Each customer is identified as a segment in his right because each person's demand function may be at least slightly different from his neighbor's.

---

[2] Second-order conditions will be ignored here.

[3] Profit maximization subject to a constraint on the total promotional budget would lead to slightly different decision rules but would not change the conclusions drawn from this or any of the following models.

Besides the obviously severe computational problems that are likely involved in solving (2), other problems ordinarily preclude this approach in practical situations, that is,

1. Marketers are rarely able to know the form or parameters of individual customer's demand functions. When such knowledge must be obtained by statistical methods, it is almost always necessary to deal with customer groups rather than individual customers.
2. It is rarely possible to pinpoint promotional efforts to specific customers or to maintain perfect price discrimination strategies. Marketers are faced with legal constraints on pricing and price leakage for resale by customers. For promotion, they must usually use standing promotional vehicles, such as advertising media, with predetermined audience characteristics.

The model must be generalized to account for difficulties. We will deal with Point 2 first because it will provide the groundwork for analyzing Point 1.

Stage 2: Customer Segmentation with Institutional Constraints

Suppose that the firm faces a fixed set of promotional vehicles through which it must exercise its nonprice marketing efforts. The fixed set of promotional vehicles will be denoted by the vector $y_i, y_2, \cdots y_n$, which we shall call "media." Thus, there are n media to reach the firm's N customers. We would expect $n < N$—though the model does not specifically require this relation. Now, the elements of the nonprice promotional vector for the ith customer can be related to the media by the set of equations:

$$x_{ij} = \Psi_i(y) \quad j = 1, \cdots m; i = 1, \cdots N.$$

We shall assume that the Ψ-functions are all linear (a reasonable assumption, as demonstrated by the following example). Thus, we may write

$$x_{ij} = \sum_{k=1}^{n} b_{ijk} y_k ,$$

where the "media characteristic paramaters" b_{ijk} represent the contribution of the kth kind of promotional input for the ith customer. The above is more concisely written in matrix form:

$$(3) \qquad x_i = \underline{\underline{B}}_i y, \qquad i = 1, \cdots N,$$

where $\underline{\underline{B}}_i$ is the $m \times n$ matrix of media characteristic parameters. Equation 3 implies that when the firm sets the variables' values in the media vector y, it determines the level of all nonprice promotional variables for the customers. Thus the y's, instead of the x's, should be seen as the controllable marketing variables.

We can clarify the meaning of (3) by using a simplified numerical example. Suppose that the market consists of three customers: one with high, one with middle, and one with low socioeconomic status. Now, assume there are two forms of nonprice competition (magazine ads and the amount of shelf space the retailers allocated to the brand) and three media available to the manufacturer (a "class" or prestige consumer magazine, a "pulp" or low-status consumer magazine, and a trade newspaper read by retailers). Assume also that all three customers shop at the same retail outlet. (These assumptions are made only to simplify the example; they are not an essential part of the model.) The table gives the matrices $\underline{\underline{B}}_i (i = 1, 2, 3)$ that might be expected to occur in the kind of situation just described. The matrices indicate that the high-status customer reads Magazine A, the low-status customer reads B, and the middle-status customer reads neither. Also, retailers are assumed to adjust shelf space in response to advertising in both the prestige consumer Magazine A and the trade paper, though the last is ten times more effective in this regard than is the first. The level of the nonprice competition vector x, e.g., the number of advertising exposures or shelf facings, that stimulate each customer can now be determined by Equation 3, using the matrices in the table.

Before changing the profit equation developed in Stage 1 to include the existence of promotional media with specific characteristics, we will invoke two simplifying assumptions. Suppose that for legal reasons, the firm decides not to follow a strategy of price discrimination. That is, the price to each customer is to be the same so that $p_i = p$ for all i. Suppose also that

HYPOTHETICAL MEDIA CHARACTERISTIC MATRICES FOR TWO-VARIABLE, THREE-MEDIA, THREE-CUSTOMER EXAMPLE

Customer (N = 3)	*Nonprice competition variable (m = 2)*	*Media (n = 3)*		
		Consumer magazine A (prestige)	*Consumer magazine B (pulp)*	*Trade magazine*
Customer 1 (high status)	Advertising exposures	1.0	0.0	0.0
	Shelf space	0.1	0.0	0.0
Customer 2 (middle status)	Advertising exposures	0.0	0.0	0.0
	Shelf space	0.1	0.0	1.0
Customer 3 (low status)	Advertising exposures	0.0	1.0	0.0
	Shelf space	0.1	0.0	1.0

the distribution costs are the same for all customers so that $c_i = c$ in all cases. These assumptions are made only for convenience (the notation is simpler if the subscripts on p and c are omitted). The model's development does not depend on these assumptions, nor do we suggest that they are realistic.

Recall from the numerical example that the elements of the media vector y are dimensioned in physical terms, e.g., number of exposures or number of shelf facings. Let the vector $w' = w_1, w_2, \cdots w_n$ be the per unit costs of using the media. For example, if $w_1 =$ one cent, a one-unit "buy" in medium 1 (the high-status magazine in the example) will cost one cent per exposure or ten dollars per thousand. Substitution of Equation 3 for x_i, w' for v'_i, and setting $p_i = p$ and $c_i = c$ in Equation 1 yields the following Stage 2 profit equation:

$$\Pi = R - C = (p - c) \sum_i f_i(p, \mathrm{B}_i y) - w'y - g\left\{\sum_i f_i(p, \mathrm{B}_i y)\right\}. \tag{4}$$

Equation 4 must be maximized with respect to p and the n elements of y.

Differentiation with respect to price leads to a simple aggregation of the first part of Equation 2.

$$(p - c - MC \sum_i \frac{\partial f_i}{\partial p} = - \sum_i f_i .$$

(This relation is the same for all subsequent models and will not be considered further.) The derivative with respect to a given medium variable y_k is:

$$\frac{\partial \Pi}{\partial y_k} = 0 = (p - c) \sum_i \sum_j \frac{\partial f_i}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial y_k} - w_k - MC \sum_i \sum_j \frac{\partial f_i}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial y_k}.$$

Transposing and recognizing that $\partial x_{ij}/\partial y_k = b_{ijk}$, we have the following set of decision rules for media:

$$(p - c - MC) \sum_i \sum_j b_{ijk} \frac{\partial f_i}{\partial x_{ij}} = w_k, k = 1, \cdots n. \tag{5}$$

This result differs from the one for Stage 1 (the second part of Equation 2) because weighted averages of the response derivatives $\partial f_i/\partial x_{ij}$ are used in aggregated equations, instead of individual terms in individual equations.

Stage 2 accounts for what might be called *institutional constraints* that restrict the marketing manager's freedom of action. Note, however, that the information requirements for Stage 2 are even more demanding than they were in Stage 1. Besides knowing all individual response derivatives $\partial f_i/\partial x_{ij}$, the manager must get estimates of the parameters of the media transfer functions $\Psi_i(y)$.

Information at the required level of detail is rarely available even for widely researched mass media (possibly except for certain industrial buying situations). Instead of relating to individual customers, audience data are usually broken down by demographic and socioeconomic variables, or at the very most, coverage may be reported by consumption level for certain key product classes [4]. Therefore, we must extend our Stage 2 model to account for these information constraints.

Stage 3: Microsegmentation

Suppose that media circulation is known only for a total of M mutually exclusive and exhaustive consumer classes, which are defined by socioeconomic, demographic, or similar variables. (These classes will be called *media descriptor classes* or, alternatively, *microsegments*.) The media characteristic coefficient matrices now refer to the descriptor classes rather than to individual customers—we have $\underline{\underline{B}}_l$, $l = 1, \cdots M$, where, for example, a given matrix might refer to "high-income, high-educated persons over 65." In principle, these matrices can be determined from audience survey information.[4] Introducing descriptor classes leads to the following modification of the Stage 2 decision rule presented in Equation 5:

$$(p - c - MC) \sum_j \sum_l b_{ljk} \sum_{i \varepsilon l} \frac{\partial f_i}{\partial x_{ij}} = w_k ; k = 1, \cdots n, \tag{6}$$

where the notation $i\varepsilon l$ means all persons within the lth descriptor cell.

It is obvious from Equation 6 that the constraint on media audience information leads to equal weighting of all members in each media descriptor class. The term $\sum_{i\varepsilon l} \partial f_i/\partial x_{ij}$ represents the aggregate marginal response to be expected from all persons in descriptive cell l. Thus, imposing a constraint on media information automatically relaxes the information requirements with respect to the individual response derivatives.

We have named the segmentation level represented by Stage 3 *microsegmentation* because segmentation by media descriptor classes is the least aggregative degree of promotional discrimination possible given existing institutional constraints and media research methods.

Stage 4: Macrosegmentation

Now let us consider the problem of estimating the marginal response of sales to promotion. It seems likely that purely judgmental methods will be insufficient to determine the demand functions or their derivatives. Some kind of empirical approach is needed

[4] This model implicitly assumes that media cover a given descriptive class homogeneously, i.e., problems of audience accumulation and duplication are ignored.

if we are to obtain estimates of the response derivatives in (6).

Empirical analyses of sales response to price or promotional variables may be discussed in terms of either individual or aggregative demand functions. The first approach is very difficult. The only substantive effort to deal with individual demand functions of which we are aware is included in Duhammel's study [1]. Though the results were interesting, they do not give us confidence in the practical efficacy of a fully disaggregative approach. If we are to conduct more aggregative statistical demand analyses, however, we must decide how much to aggregate and on what variables the process should be based. Equation 6 gives an immediate answer—at least in part. According to the results of Stage 3 we can always aggregate to the level of the smallest microsegment for which media information is commonly available.

But aggregation to the level of (6) may not be enough. Demand function analysis involves estimation of the change in sales to be expected per unit change in promotion, but media audience research concerns average audience levels for descriptor classes. If we consider data sources for these analyses, it becomes apparent that sample sizes sufficient to measure the elements of B_l for a given descriptor class, with reasonable degree of accuracy, may be insufficient to measure $\sum_{i\varepsilon l} \partial f_i/\partial x_{ij}$. As Frank and Massy [2, 3], and others have shown, the estimation of response coefficients is not easy. Even if the analysis is based on time series data, the time series must be based on the buying behavior of a sufficient number of families to avoid gross instabilities.

Since the sample sizes necessary to estimate response sensitivities for a given microsegment must often be rather large, the researcher is faced with two alternatives: (1) using the maximum number of descriptor cells, and hence a very large overall sample size, or (2) aggregating over descriptor classes to form a smaller number of new classes with adequate numbers of respondents in each, while keeping the total sample within bounds. The total sample size for a study is often fixed (as when working with syndicated panel data), in which case the second alternative is the only feasible one. Or perhaps the cost of data collection precludes using a large overall sample size. Finally, the cost of audience research is usually absorbed by the media (or agencies), suggesting that larger sample sizes will be possible there than for product or brand specific sensitivity analyses, if only because the costs can be divided among many users.

This reasoning implies that aggregation beyond the minimal descriptor class sizes dictated by available media audience statistics will be the rule rather than the exception if statistical methods for estimating demand sensitivities are to be used practically. How should the aggregation be performed? Given the potentialities of media research for dealing with relatively disaggregative microsegments, it seems reasonable to use these media descriptors to build more aggregative demand descriptor classes. Therefore, we define a *macrosegment* as follows: macrosegment h consists of the customers in media descriptive cells $l\varepsilon h$. This definition ensures that it will be possible to make media decisions for each segment. Since media characteristic coefficients can be found for each microsegment l, the media characteristics for macrosegments h can be found by simple aggregation.

The promotion rule for Stage 3 (Equation 6) is easily modified to accommodate the higher level of aggregation.

$$(7)\quad (p - c - MC) \sum_j \sum_h \left\{\sum_{l\varepsilon h} b_{ljk}\right\} \left\{\sum_{l\varepsilon h} \sum_{i\varepsilon l} \frac{\partial f_i}{\partial x_{ij}}\right\} = w_k, k = 1, \cdots n.$$

The sensitivity term $\left\{\sum_{l\varepsilon h} \sum_{i\varepsilon l} \partial f_i/\partial x_{ij}\right\}$ might be written simply as $\partial f_h/\partial x_{hj}$ to emphasize that it refers to the aggregate demand function for the hth macrosegment.

Stage 5: The "Mass Market" Concept

It will be useful to present another generalization of our market segmentation model, this one corresponding to the case in which no segmentation strategy is practiced at all. Profit maximization without segmentation leads to the following decision rule for promotion:

$$(8)\quad (p - c - MC) \sum_j \left\{\sum_{l=1}^{M} b_{ljk}\right\} \left\{\sum_{l=1}^{M} \sum_{i\varepsilon l} \frac{\partial f_i}{\partial x_{ij}}\right\} = w_k, k = 1, \cdots, n,$$

where the first term in the brackets represents the total impact of medium k in terms of promotion type j for all numbers of the population, and the second term is the derivative of the total market demand function. As easily recognized, (8) is the same as (7) if there is only one inclusive macrosegment.

SOME IMPLICATIONS OF THE THEORY

The theoretical models presented in the preceding section have several implications for both the philosophy and practice of segmentation.

First, it seems clear that segmentation should be considered as a process of *aggregation* rather than *disaggregation*. For example, recall that the theory's five stages dealt with the full range of segmentation possibilities. That is Stages 1 and 2 treat individual consumer units as segments; Stage 3 deals with aggregation of consumer units into microsegments; Stage 4 considers aggregations of microsegments into larger groups (macrosegments), and Stage 5 deals with the mass market or complete aggregation of consumers.

It should be obvious that if consumers have different responses to the firm's marketing variables and if there are no scale diseconomies in fitting specialized programs to individual consumers, segmentation at the level of individual consumer units (the case of the perfectly discriminating monopolist) would yield maximum profits. However, this discussion showed that even if the dubious assumption of no scale economies in the marketing mix is valid, other constraints typically preclude this form of segmentation. For example, lack of information about the response characteristics of groups reached by promotional media and institutional constraints on the flexibility of their use require aggregation at least to the level of microsegments (Stage 3). Additional aggregation to macrosegments—or ultimately to a single segment—may be required because of difficulties in measuring response differentials for specific groups.

It is easy to see that addition of the successive constraints and corresponding higher levels of aggregation must reduce the level of the firm's profit. (This statement ignores research costs required to implement a given level of segmentation.) That is, the application of Stage 1 segmentation yields more profit than Stage 2, Stage 2 more than Stage 3, etc. This is a direct result of the mathematical properties of constrained versus unconstrained maxima. Thus the fundamental problem of market segmentation can be characterized as finding the point at which the marginal reduction of profits caused by the imposition of another constraint, or level of aggregation, is just balanced by the marginal reduction in research and administration costs made possible by the constraint.

We argued earlier that this balance will likely occur at the macrosegmentation stage. Thus the basic resource allocation problem in segmentation involves finding the values of the controllable marketing variables that bring the decision rule in (7) to equality. The problem of finding the solution to (7) can be handled with standard mathematical programming procedures when the necessary data have been collected and the macrosegments have been defined. Though the problems of finding a solution are not trivial, they will not be explored here.

The concept that segmentation is a process of aggregation implies building to a viable segmentation strategy rather than tearing a market apart to find one. This may be a fine point in regard to the philosophy of segmentation, but it appears to be important for the implementation of the strategy. It is impossible to form meaningful market segments without taking institutional and information constraints into account and this means building from the point at which the constraints are felt, namely from persons to microsegments to macrosegments.

Criteria for Forming Macrosegments

Let us now consider some implications of the theory for the formation of meaningful macrosegments. In particular, we consider the question of how media descriptor cells (microsegments) should be allocated to macrosegments.

It is clear that if the response derivatives for all microsegments included in each of the macrosegments are identical, Equation 7 is merely a factored form of Equation 6. For the jth type of nonprice competition we have:

$$\sum_l b_{ljk} \sum_{i \varepsilon l} \frac{\partial f_i}{\partial x_{ij}} = \sum_h \left\{ \sum_{l \varepsilon h} b_{ljk} \right\} \cdot \left\{ \sum_{l \varepsilon h} \sum_{i \varepsilon l} \frac{\partial f_i}{\partial x_{ij}} \right\}, \qquad k = 1, \cdots n,$$

if

$$\sum_{i \varepsilon l} \frac{\partial f_i}{\partial x_{ij}} = \sum_{i \varepsilon l^*} \frac{\partial f_i}{\partial x_{ij}} \qquad \text{for all } l, l^* \varepsilon h.$$

(Recall that we can do nothing about any possibility that the $\partial f_i / \partial x_{ij}$ are heterogeneous within a given microsegment without changing audience research procedures.)

This consideration suggests that for a given kind of competitive activity it will be useful to form segments such that

$$\text{variance} \quad \left\{ \sum_{i \varepsilon l} \frac{\partial f_i}{\partial x_{ij}} \right\}, \qquad \text{for } l \varepsilon h,$$

is as small as possible for each of the macrosegments. This means that microsegments (l) should be assigned to macrosegments (h) in such a way that the within-group variances of microsegment response coefficients are as small as possible relative to the between-group variance.[5] (The media characteristic coefficients b_{ljk} are not included in the variance formula because the macrosegments must be developed before a specific promotional program is chosen.)

Macrosegment formation requires grouping microsegments by similarities among their promotional responses. How can information on the response derivatives for each microsegment be obtained? It will generally be necessary to determine the kind of promotional variables that are likely to be used in the marketing mix and define specific response variables for each of them. (The change in the buying probability caused by an additional advertising exposure would be one such measure.) Often it will be necessary to use surrogates for the response variables, as when the

[5] The average within-group variance will surely decline as more macrosegments are permitted, assuming that optimal allocations are made at each stage. As noted earlier, the question of determining the optimal number of segments must be resolved by analyzing the trade-off between costs of research and marketing administration (which rise with the number of segments considered) and the gains to be expected from reductions in within-group variance. The analysis will depend on factors specific to problem and product class and is beyond the scope of this article.

response of consumer attitude or awareness measures to changes in advertising are used in place of the sales response.

Looser response measures, e.g., less adequate surrogates, will more usually be dictated when working at the microsegment level than will subsequently be used when dealing with macrosegments. This will occur because of the many microsegments for which response measures must be estimated and the difficulty of obtaining data on individual microsegments. However, this difficulty will be reduced because the response measures will be reestimated at the macrosegment level.

The last problem raised by the theory is to find a method for optimally grouping microsegments into macrosegments, assuming that information on the relevant response derivatives is available for each microsegment. That is, we want to make assignments such that the resulting macrosegments consist of microsegments with large within-group and small between-group homogeniety of response derivatives.

Making such assignments is a problem in optimal taxonomy, also called cluster analysis.[6] It is fortunate that such programs have recently been developed, because without them macrosegment formation would largely be guesswork. Although many detailed problems in using these procedures must be solved by future research, we conclude that the data sources and research methods now available are sufficient to permit their application. The value of our theory of market segmentation will be proved if it leads researchers to experiment with these techniques that allow a systematic attack on the practical problems of segmenting markets.

[6] For a description of one optimal taxonomy procedure, see [8]. For an application of the clustering method to a marketing problem, see [5]. Claycamp and Massy are doing research on using the Rubin cluster analysis program to form macrosegments as indicated above. Results will be in a future article.

REFERENCES

1. William F. Duhammel, "The Use of Variable Markov Processes as a Partial Basis for the Determination and Analysis of Market Segments," Unpublished doctoral dissertation, Graduate School of Business, Stanford University, March 1966.
2. Ronald E. Frank and William F. Massy, "Market Segmentation and the Effectiveness of a Brand's Price and Dealing Policies," *Journal of Business*, 38 (April 1965), 186–200.
3. "Estimating the Effects of Short-Term Promotional Strategy in Selecting Market Segments," in Patrick J. Robinson and Charles L. Hinkle, eds., *Sales Promotion Analysis: Some Applications of Quantitative Techniques*, Philadelphia: Marketing Science Institute (in press).
4. Harper W. Boyd, "Correlates of Grocery Product Consumption Rates," *Journal of Marketing Research*, 4 (May 1967), 184–90.
5. Paul Green, Ronald Frank, and Patrick Robinson, "Cluster Analysis in Test Market Selection," *Management Science*, 13 (April 1967), B387–400.
6. William F. Massy, "The Cost of Uncertainty in Advertising Media Selection," Graduate School of Industrial Administration, Carnegie Institute of Technology, Working Paper No. 2, December 1966.
7. Joan Robinson, *The Economics of Imperfect Competition*, London: Macmillan & Co., 1954, 179–88.
8. Jerrold Rubin, "Optimal Classification into Groups: An Approach for Solving the Taxonomy Problem," International Business Machines Corporation, New York Scientific Center, May 1965.