Title

Chapter 6 Market-share models

Author

Cooper, Lee G

Publication Date

1993

DOI

10.1016/s0927-0507(05)80029-5

J. Eliashberg and G.L. Lilien, Eds., *Handbooks in OR & MS, Vol. 5*

Chapter 6

Market-Share Models*

Lee G. Cooper

Anderson Graduate School of Management, University of California at Los Angeles, Los Angeles, CA 90024-1481, USA

1. Introduction

The topic of market-share models overlaps substantially with Roberts and Lilien's discussion of choice models (This Handbook, Chapter 2), since consumer choice underlies the process of market-share formation; with Gatignon's discussion of marketing-mix models (This Handbook, Chapter 15), since market-share models should be useful in planning the marketing mix for a brand in light of competitive forces, with the Hanssens and Parsons discussion of enconometric and time-series models (This Handbook, Chapter 9), since market-share models deal with market response over time as well as over geography and over competitors, and with the Blattberg and Neslin discussion of sales-promotion models (This Handbook, Chapter 12), since market-share models have been used extensively to analyze the effects of retail promotions. What this chapter contributes to management science in marketing beyond that offered by these other chapters is hopefully an understanding of how increasingly rich specification of market-share models can guide the systematic study of market and competitive structures.

By *market and competitive structures* we mean the complex pattern of competitive interplay within a market. Take the cereal market as an example. The presweetened cereals such as Frosted Flakes or Honey Nut Cheerios serve a predominantly different segment than the mainly unsweetened cereals such as Cheerios, Kellogg's Corn Flakes, or Wheaties. Substitution *within* these submarkets is likely to be more frequent than *between* them. But we cannot treat them as totally separate markets since dramatic price reductions for Cheerios might entice consumers to switch from the presweetened segment and sweeten at home. Even within the unsweetened cereals, we expect more substitution within varieties of corn flakes, than between corn flakes and wheat flakes, or more substitution within varieties

*This chapter relies heavily on the developments in Chapters 2 and 3 of Cooper & Nakanishi [1988]. The numerous contributions of Masso Nakanishi to this work are gratefully acknowledged. I also thank M.J. Snyder for her assistance with the numerical illustration, A.C. Nielsen for providing the data used in the numerical illustration, and the reviewers for their many helpful comments.

of granolas than between granolas and single-grain cereals. Media advertising for one of Kellogg's brands may help other Kellog brands, have little effect on other premium national brands, and hurt each store's private-label brand. In return the large price differential between a store's private-label brand and the corresponding national brand (e.g. Honey Nut Cheerios and Honey Nut Tasteeos–two to one according to a recent *Wall Street Journal* story, Gibson [1991]) can create long-term cross-competitive influences within this dyad that are not felt throughout the market.

Three basic principles motivate the specification of the market-share models discussed in this chapter. Market-share models should be *competitive*, *descriptive* as well as predictive, and *profit-oriented*, Being fundamentally *competitive* implies that we cannot know the effect or effectiveness of a marketing action without accounting for the actions of competitors. Say we try an eight-week promotion for a bar soap that combines a newspaper feature with a high-value coupon and a low feature price, and observe only a 10% gain over baseline sales. Management's reaction to such a disappointing response would be very different if all other brands were sitting quietly at shelf price, than if the other national brands were temporally exhausting demand with four-for-the-price-of-three multipacks. While single-brand sales models can be specified to include the actions of particular competitors, only market-share models include competition as a fundamental part of their composition. Market-share models are models for understanding how the marketing efforts of every brand impact the results in a competitive marketplace. Only by *describing* the influence of each marketing instrument can we gain a basis for marketing planning. *Prediction* alone is not enough. Time-series models that forecast the future from the past sales provide no insight into how those sales were generated. The emphasis on being *descriptive* also embraces the need to understand the areas in which consumer choice probabilities are synonymous with market shares as well. Part of the goal of *description* transcends what can be done by market-share models alone. Managers need to understand that their efforts have (potentially) competitive effects and (potentially) market-expansive effects. In sales-response models these effects are comingled, but by combining *descriptive* market-share models for the competitive effects with *descriptive* category-volume models for the market-expansive effects, managers obtain a much richer understanding of the market. The *profit-oriented* goal of market-share analysis urges us to ask how the firm's allocations of resources to aspects of the marketing mix produce bottom-line results. It reminds us that maximizing market share is not the same as maximizing profits. Systematically understanding how the actions available to a brand impact the market results of all competitors would be a very important gain for marketing science.

We begin by discussing the fundamental relations between marketing effort and market shares. In this context we discuss the derivation of the basic theorem that relates marketing effort to market share. We then develop, present and discuss five alternative market-share models, and evaluate them in terms of how we expect market-share elasticities to behave. We then discuss the relations between the aggregate concept of *market share* and the disaggregate concept of *choice probability*, with special emphasis on the threats to parameter stability in market-share models. The topic then shifts to the sources and representation of asymmetries in markets and competition, and the issue of the distinctiveness of marketing activities. A numerical example illustrates the diagnostic richness of these models. A concluding section discusses some of the challenges that market-share models must confront in the next decade.

2. Market share and marketing effort

We use the term *market share* to signify the share of total market sales (in quantity sold) for a product in a given period and in a given geographical area. This concept of market share is more explicitly stated in the following manner:

$$s_i = \frac{Q_i}{Q} \tag{1}$$

where

s_i = the market share of brand i,
Q_i = the sales of brand i's product,
Q = the total sales for the market,
$Q = \sum_{j=1}^{m} Q_j$,
m = the number of competing brands.

The quantity Q in the above equation is commonly called the *industry sales, primary demand*, all commodity volume (ACV), or simply *the market*. Market shares are temporally and spatially specific. They are defined and measured only for a specific period and a specific geographical area, but the time or occasion subscript is dropped in these preliminary developments for notational convenience. It does not make much sense for one to talk about a brand's share in general; one must say instead a brand's share of the Chicago trading area in 1990, that of the New York market in 1991. This is because both the numerator and denominator of Equation (1) are time- and area-specific, and will have to be matched for the same period and geographical area for this ratio to make sense.

Kotler [1984] sets up the basics for using market-share models in brand planning by asserting that a brand's market share is proportional to the *marketing effort* supporting it. We may represent this by:

$$s_i = k \cdot M_i \tag{2}$$

where

M_i = the marketing effort of brand i,
k = a constant of proportionality.

This simplest representation asserts that the greater the brand's marketing effort, the greater should be its market share.

Setting aside for the time being the question of how one might measure marketing effort, one must know the value of the proportionality constant, k, before Equation (2) is useful. But since market shares for an industry must sum to one, i.e.

$$\sum_{i=1}^{m} s_i = 1,$$

we know that

$$\sum_{i=1}^{m} k \cdot M_i = 1$$

or

$$\sum_{i=1}^{m} M_i = \frac{1}{k}.$$

Hence

$$k = \frac{1}{\sum_{i=1}^{m} M_i}.$$

By substituting this value of k in Equation (2), we have

$$s_i = \frac{M_i}{\sum_{j=1}^{m} M_j}. \tag{3}$$

This last equation says that the market share of brand i is equal to the brand's share of the total marketing effort, a statement which certainly seems plausible. Equation (3) is what Kotler calls the *fundamental theorem* of market share [Kotler, 1984, p. 231].

On important variation on this fundamental theorem indicates if brands tended to differ in terms of the effectiveness of their marketing effort; one may write

$$s_i = \frac{\alpha_i \cdot M_i}{\sum_{j=1}^{m} \alpha_j \cdot M_j} \tag{4}$$

where α_i is the effectiveness coefficient for brand i's marketing effort. This implies that, even if two brands expend the same amount of marketing effort, they may not have the same market share. If one brand's marketing effort is twice as effective as that of the other, the former will achieve a market share twice as large as the other's share.

Specification and measurement of what goes into *marketing effort* is an important topic. Kotler assumes that a brand's marketing effort is a function of its marketing mix, both past and current. Mathematically, we may write

$$M_i = f(P_i, A_i, D_i, \ldots) \tag{5}$$

where

P_i = the price of brand i's product,
A_i = the advertising expenditures of brand i,
D_i = the distribution efforts (e.g. the percent of stores carrying brand i).

There are wide choices in the specification of the functional form for Equation (5). For example, if we believe that components of the marketing mix interact we may choose a multiplicative function

$$M_i = P_i^p \cdot A_i^a \cdot D_i^d$$

where p, a, d are parameters to be estimated reflecting the importance of each component of the marketing mix. If one substitutes this expression in (3) or (4), the resultant market-share model will be the simplest version of an MCI (multiplicative, competitive-interaction) model. Or if we choose an exponential function

$$M_i = \exp(p \cdot P_i + a \cdot A_i + d \cdot D_i),$$

the market-share model is the simplest version of the multinomial logit (MNL) model.

2.1. Market-share theorem

Kotler's *market-share-as-share-of-marketing-effort* representation makes a lot of intuitive sense, but there are other ways to derive such a representation. We will review some of them in a later section, and only look here at one important theorem derived by Bell, Keeney & Little [1975].

Bell, Keeney & Little (BKL) consider a situation where, in purchasing a product, consumers must choose one brand from a set of alternative brands available in the market. They posit that the only determinant of market shares is the *attraction* which consumers feel toward each alternative brand, and make the following assumptions about attractions. Letting $\mathscr{A}_i$ be the attraction of brand i ($i = 1, 2, \ldots, m$) and s_i be its market share,

Axiom A1. $\mathscr{A}_i \geqslant 0$ for all i and $\sum_{i=1}^{m} \mathscr{A}_i > 0$ (i.e., attractions are nonnegative and their sum is positive).

Axiom A2. $\mathscr{A}_i = 0 \Rightarrow s_i = 0$. (The symbol $\Rightarrow$ should be read 'implies', i.e., zero attraction implies zero market share.)

Axiom A3. $\mathscr{A}_i = \mathscr{A}_j \Rightarrow s_i = s_j$ $(i \neq j)$ (i.e., equal attraction implies equal market share).

Axiom A4. When $\mathscr{A}_j$ changes by Δ, the corresponding change in s_i $(i \neq j)$ is independent of j (e.g. a change in attraction has a symmetrically or proportionally distributed effect on competitive market share).

From those four *axioms* they show that the following relationship between attractions and market shares may be derived:

$$s_i = \frac{\mathscr{A}_i}{\sum_{j=1}^{m} \mathscr{A}_j}. \tag{6}$$

The first three axioms are not controversial. But, as will be developed later, Axiom A4 has been the subject of much critical discussion. While Equations (3) and (6) represent two rather distinct schools of thought regarding the determinants of market shares (a brand's marketing effort for the former and consumer attraction for the latter), few would argue the fact that these equations are extremely similar. An additional assumption that the attraction of a brand is proportional to its marketing effort (which has intuitive appeal) is all that is required to reconcile these two equations.

BKL also show that a slightly different set of assumptions also yield Equation (6). Let C be the set of all alternative brands from which consumers make their choice.

Axiom B1. $\mathscr{A}_i \geqslant 0$.

Axiom B2. The attraction of a subset $S(\subseteq C)$ is equal to the sum of the attractions of elements in S.

Axiom B3. $\mathscr{A}_i$ is finite for all i and nonzero for at least one element in C.

Axiom B4. If the attractions of subsets $S^{(1)}$ and $S^{(2)}$ are equal, their market shares are equal.

The last axiom establishes the relationship between attractions and market shares. BKL observe that, if we add an assumption that

$$\sum_{i=1}^{m} \mathscr{A}_i = 1$$

in lieu of B4, $\mathscr{A}_i$ in this set of axioms satisfies the assumptions for probabilities in a finite (discrete) sample space. Because of this BKL suggest that attractions may be interpreted as *unnormalized probabilities*. However, this in turn suggests that if attractions were to follow axioms B1 through B4, by normalizing the $\mathscr{A}_i$'s through (6), market shares (s_i) may be interpreted as probabilities. This confuses an *aggregate* concept (*market shares*) with an *individual* (or *disaggregated*) concept (*probabilities*). When the market is homogeneous (i.e. not composed of systematically different consumer segments), market shares and choice probabilities may be used interchangeably. But, as is discussed later, in other circumstances we must be careful not to use these concepts interchangeably.

2.2. Alternative models of market share

The previous sections gave the rationales behind the MCI model and its close cousin, the MNL model. We now give explicit specifications to the simplest versions of those models.

MCI model:

$$\mathscr{A}_i = \exp(\alpha_i) \cdot \prod_{k=1}^{K} X_{ki}^{\beta_k} \cdot \varepsilon_i, \tag{7}$$

$$s_i = \frac{\mathscr{A}_i}{\sum_{j=1}^{m} \mathscr{A}_j}.$$

MNL model:

$$\mathscr{A}_i = \exp\left(\alpha_i + \sum_{k=1}^{K} \beta_k \cdot X_{ki} + \varepsilon_i \right), \tag{8}$$

$$s_i = \frac{\mathscr{A}_i}{\sum_{j=1}^{m} \mathscr{A}_j},$$

where

s_i = the market share of brand i,
$\mathscr{A}_i$ = the attraction of brand i,
m = the number of brands,
X_{ki} = the value of the kth explanatory variable X_k for brand i (e.g. prices, product attributes, expenditures for advertising, distribution, sales force),
K = the number of explanatory variables,
β_k = a parameter to be estimated,
α_i = a parameter for the constant influence of brand i,
ε_i = an error term.

In what follows we will use *attraction*, rather than *marketing effort*, to describe $\mathscr{A}_i$, because it is a more accepted terminology, keeping in mind that this implies the assumption that attraction is proportional to marketing effort.

The MCI and MNL models are not the only models of market shares. The *linear model* assumes simply that a brand's market share is a linear function in marketing-mix variables and other relevant variables. The *multiplicative* and *exponential models* represent market shares as a product of a number of variables (shrunk or stretched by appropriate parameters). Although there are other more complicated market-share models, for our purposes at present we need only define these three alternative models.

Linear model:

$$s_i = \alpha_i + \sum_{k=1}^{K} \beta_k \cdot X_{ki} + \varepsilon_i. \tag{9}$$

Multiplicative model:

$$s_i = \exp(\alpha_i) \cdot \prod_{k=1}^{K} X_{ki}^{\beta_k} \cdot \varepsilon_i. \tag{10}$$

Exponential model:

$$s_i = \exp\left(\alpha_i + \sum_{k=1}^{K} \beta_k \cdot X_{ki} + \varepsilon_i \right). \tag{11}$$

The five models – MCI, MNL, linear, multiplicative and exponential – are closely related to each other. For example, if we take the logarithm of both sides of either the multiplicative or exponential model, we will have a linear model (linear in the parameters of the respective models, and not in variables). In other words, while the conceptual difference may be great, the mathematical difference between the linear model and the multiplicative and exponential models is merely in the choice of transformations for variables, that is, whether or not the logarithmic transformation is applied to the explanatory variables. (The specification for the error term may be different in those three models, but this is a technical issue that will not be addressed here.)

The most interesting relationship is, however, the one between the MCI and multiplicative models (and the corresponding relationship between the MNL and exponential models). The multiplicative model, of course, assumes that market shares are a multiplicative function in explanatory variables, while in the MCI model attractions are multiplicative in variables and market shares are computed by normalizing attraction (making the sum of market shares to be equal to one). Obviously, the key difference between the two is normalization. In this connection, Naert & Bultez [1973] proposed the following important conditions for a market-share model:

(1) Estimated market shares from the model are nonnegative.
(2) The estimated market shares sum to one over all the competitors.

These conditions, commonly known as the *logical-consistency requirements*, are clearly not met by either the multiplicative or the exponential model, but are met by their respective normalized forms (i.e., MCI and MNL) – a clear advantage for MCI and MNL models. Note that the linear model does not satisfy the logical-consistency requirements.

Why, then, are the MCI and MNL models not used more extensively? The answer is that for a time both of those models were considered to be intrinsically nonlinear models, requiring estimation schemes that were expensive in analysts' time and computer resources. This, however, turned out to be a hasty judgment because Nakanishi [1972; Nakanishi & Cooper, 1974] showed that those models may be changed into a linear model (in the model parameters) by a simple transformation. Using the MCI model, for example, we can first take the logarithm of both sides.

$$\log s_i = \alpha_i + \sum_{k=1}^{K} \beta_k \log X_{ki} + \log \varepsilon_i - \log\left\{\sum_{j=1}^{m}\left(\alpha_j \prod_{k=1}^{K} X_{kj}^{\beta_k}\varepsilon_j\right)\right\}.$$

If we sum the above equation over i $(i = 1, 2, \ldots, m)$ and divide by m, we have

$$\log \tilde{s} = \bar{\alpha} + \sum_{k=1}^{K} \beta_k \log \tilde{X}_k + \log \tilde{\varepsilon} - \log\left\{\sum_{j=1}^{m}\left(\alpha_j \prod_{k=1}^{K} X_{kj}^{\beta_k}\varepsilon_j\right)\right\}$$

where $\tilde{s}$, $\tilde{X}_k$ and $\tilde{\varepsilon}$ are the geometric means of s_i, X_{ki} and ε_i, respectively. Subtracting the above from the previous equation, we obtain

$$\log\left(\frac{s_i}{\tilde{s}}\right) = \alpha_i^* + \sum_{k=1}^{K} \beta_k \log\left(\frac{X_{ki}}{\tilde{X}_k}\right) + \varepsilon_i^* \qquad (12)$$

where

$$\alpha_i^* = (\alpha_i - \bar{\alpha}), \qquad \varepsilon_i^* = \log(\varepsilon_i/\tilde{\varepsilon}).$$

The last equation is linear in model parameters α_i^* $(i = 1, 2, \ldots, m)$ and β_k $(k = 1, 2, \ldots, K)$. (In addition, there is another parameter σ_ε^2, the variance of ε_i, to be estimated, but this parameter does not concern us here.) This transformation is called the *log-centering* transformation. The importance of this transformaton is that it leads to ways to estimate the parameters of the original nonlinear model using linear-regression techniques. Note also that if we apply the inverse of this transformation to the estimates that result from a linear regression we must obtain market-share estimates that satisfy the logical-consistency conditions. That is,

$$\hat{s}_i = \frac{\exp \hat{y}_i}{\sum_{j=1}^{m} \exp \hat{y}_j}$$

where $\hat{s}_i$ is the estimate market share for brand i and $\hat{y}_i$ is the estimated dependent variable from the *reduced form* of the regression model. This is known as the *inverse log-centering transformation.*

If we apply the log-centering transformation to the MNL model, we obtain the following linear form:

$$\log\left(\frac{s_i}{\tilde{s}}\right) = (\alpha_i - \bar{\alpha}) + \sum_{k=1}^{K} \beta_k (X_{ki} - \bar{X}_k) + (\varepsilon_i - \bar{\varepsilon})$$

where $\bar{\alpha}$, $\bar{X}_k$ and $\bar{\varepsilon}$ are the arithmetic means of α_i, X_{ki} and ε_i, respectively. If we let $\alpha_i^* = (\alpha_i - \bar{\alpha})$ and $\varepsilon_i^* = (\varepsilon_i - \bar{\varepsilon})$,

$$\log\left(\frac{s_i}{\tilde{s}}\right) = \alpha_i^* + \sum_{k=1}^{K} \beta_k (X_{ki} - \bar{X}_k) + \varepsilon_i^*. \tag{13}$$

Both Equations (12) and (13) are linear functions of the model parameters, and hence can be called *log-linear* models. The multiplicative and exponential models are also log-linear models. In other words, both the MCI and MNL models are really special cases of log-linear models. Compare the *reduced forms* of these models.

Linear model:

$$s_i = \alpha_i + \sum_{k=1}^{K} \beta_k X_{ki} + \varepsilon_i.$$

Multiplicative model:

$$\log s_i = \alpha_i + \sum_{k=1}^{K} \beta_k \log X_{ki} + \log \varepsilon_i.$$

Exponential model:

$$\log s_i = \alpha_i + \sum_{k=1}^{K} \beta_k X_{ki} + \varepsilon_i.$$

MCI model:

$$\log\left(\frac{s_i}{\tilde{s}}\right) = \alpha_i^* + \sum_{k=1}^{K} \beta_k \log\left(\frac{X_{ki}}{\tilde{X}_k}\right) + \varepsilon_i^*.$$

MNL model:

$$\log\left(\frac{s_i}{\tilde{s}}\right) = \alpha_i^* + \sum_{k=1}^{K} \beta_k (X_{ki} - \bar{X}_k) + \varepsilon_i^*.$$

In all five equations the right-hand side is linear in both α_i or α_i^* $(i = 1, 2, \ldots, m)$ and β_k $(k = 1, 2, \ldots, K)$. The left-hand side is a market share, the logarithm of a market share, or a log-centered form of a market share. Ignoring the differences in the specification of the error term (ε_i, $\log \varepsilon_i$ or ε_i^*), note that the number of parameters in the five formulations are the same, one would expect that those models would be just as accurate in predicting the *dependent variable*, namely, the left-hand side of each equation. Which one, then, makes the most accurate prediction of market shares? Many studies on predictive accuracy of market-share models [Brodie & De Kluyer, 1984; Ghosh, Neslin & Shoemaker, 1984; Leeflang & Reuyl, 1984; Naert & Weverbergh, 1981, 1985] found the logical-consistency property of the MCI and MNL models to produce only marginally better predictions than the linear and multiplicative models. Whey then all this fuss about the MCI and MNL models? First, these tests did not include the more sophisticated versions of the models specified later in this chapter. And second, as was stated in the Introduction, we do not believe that predictive accuracy is the only important criterion for judging the value of a model. We would rather find the answer in the *construct validity* (i.e. intrinsic meaningfulness) of those models. Is the model *descriptive* in a way that facilitates brand planning? Is the model *profit-oriented* so that the revenue implications can be investigated? Do the elasticities make managerial sense? This latter issue is addressed in the next section.

2.3. Market-share elasticities

Simply stated, *market-share elasticity is the ratio of the relative change in a market share corresponding to a relative change in a marketing-mix variable.* Expressed mathematically,

$$e_{s_i} = \frac{\Delta s_i / s_i}{\Delta X_{ki} / X_{ki}} = \frac{\Delta s_i}{\Delta X_{ki}} \cdot \frac{X_{ki}}{s_i} \tag{14}$$

where s_i is the market share and X_{ki} is the value of the kth marketing-mix variable, for brand i. The symbol Δ indicates a change in respective variables. There is nothing conceptually difficult in market-share elasticity. For example, if a brand's share increased 10% (say from 30 share points to 33) corresponding to a price reduction of 5%, the above equation would give a (price) elasticity of -2; or if advertising expenditures were increased by 3% and as a result the share increased by 1%, the (advertising) elasticity would be 0.33; and so forth. We should really use $e_{s_i}^{(k)}$ to denote the elasticity of share with respect to a particular instrument k, but drop the superscript for notational convenience.

There is no way to estimate elasticities directly from empirical data without adopting a model. This may not be intuitively clear because the formula for computing elasticities (14) appears to contain only those terms which may be empirically measurable. But note that the Δs_i term in Equation (14) must correspond to the change in a specific marketing-mix variable, ΔX_{ki}. Suppose that one observed that a brand's share increased 3% in a period. How does one know how much of that increased share is due to price reduction? Or due to increased advertising? To assess those so-called *partial effects* one needs a market-share model.

The reader may be cautioned at this point that the estimated values of elasticities vary from one model to another, and hence one must choose the model that conceptually fits the situation best. To illustrate, we will derive the share elasticity with respect to X_{ki} for the simplest version of each model. For that purpose, however, one needs another concept of share elasticity which is slightly different from the one defined by (14). Technically, (14) is called the *arc elasticity*. This is because both Δs_i and ΔX_{ki} span a range over the market-response curve which gives the relationship between market shares and marketing-mix variables. The other elasticitiy formula is called the *point elasticity* and takes the following form:

$$e_{s_i} = \frac{\partial s_i}{\partial X_{ki}} \cdot \frac{X_{ki}}{s_i}. \tag{15}$$

Note that the only difference between the two formulas is that $(\Delta s_i/\Delta X_{ki})$ in Equation (14) is replaced by $(\partial s_i/\partial X_{ki})$ in (15). Formula (15) utilizes the slope of the market-response curve at a specific value of X_{ki}. The reason for using the point-elasticity formula rather than the arc formula is that the former gives much simpler expressions of share elasticity. We may add that (15) is a close approximation of (14) for a small value of ΔX_{ki}, that is, when the change in X_{ki} is very small. The point elasticity for each model is given below.

Linear model:

$$e_{s_i} = \beta_k X_{ki}/s_i.$$

Multiplicative model:

$$e_{s_i} = \beta_k.$$

Exponential model:

$$e_{s_i} = \beta_k X_{ki}.$$

MCI model:

$$e_{s_i} = \beta_k(1 - s_i).$$

MNL model:

$$e_{s_i} = \beta_k(1 - s_i)X_{ki}.$$

Though the five market-share models are similar in the sense that they are either linear or log-linear models, the share elasticities implied from the models are quite different. One may wish to disqualify some models on the basis of those expressions on some a priori grounds.

Market-share elasticities should have the following properties:

(1) Since $Q_i = Q \cdot s_i$, we can show that $e_{Q_i.x} = e_{Q.x} + e_{s_i.x}$,
(2) If s_i is an increasing function in $X, e_{s_i.x} \to 0$ as $s_i \to 1$. This is because, as $s_i \to 1$, $Q_i \to Q$ and therefore $e_{Q_i.x} \to e_{Q.x}$ in property (1) above.
(3) If s_i is a strictly increasing function in $X, e_{s_i.x} \to 0$ as $X \to \infty$. This is derived from property (2) above, since as $X \to \infty$, $s_i \to 1$. If s_i is an increasing function in X but approaches a constant (<1) as $X \to \infty$, then $e_{s_i.x} \to 0$ (since $\mathrm{d}s_i/\mathrm{d}X \to 0$).

The first property is a strict mathematical consequence of the relation of brand sales to market share and industry sales. While none of the currently considered models fails to satisfy this condition, it should be kept in mind for evaluating other alternatives. The second property reflects that a brand's share elasticity should approach zero as the share for that brand approaches one. The multiplicative model implies that share elasticity is constant regardless of the share level, and therefore seems rather inappropriate as a market-share model.

The third property reflects the generally accepted view that it becomes harder to gain market shares as a brand increases its marketing effort. In other words, one would expect market-share elasticity to approach zero as X_{ki} goes to infinity (or minus infinity, depending on the variable in question). But the exponential model implies an opposite: share elasticity may be increased indefinitely as the value of X_{ki} increases. This is an uncomfortable situation, especially if variable X_{ki} is a promotional variable (such as advertising expenditures, number of salesmen, etc., since it implies ever-increasing market-share returns for increasing promotional expenditures. In addition, the exponential model has the same problem as the multiplicative model: for a fixed value of X_{ki}, e_{s_i} is constant for all levels of s_i.

Note that the elasticity expression for the linear model reflects that share elasticity declines as the share increases, but, when the share approaches one, the elasticity does not approach zero. In fact, share elasticity approaches 1 as X_{ki} increases to infinity (or minus infinity, as the case may be). Thus the linear model produces a highly unreasonable share-elasticity expression.

Considering what we expect from share elasticities, one may conclude that the linear, multiplicative and exponential models are not proper market-share models for use in marketing decision-making. This leaves us the MCI and MNL models as feasible alternatives. Figure 6.1 shows the change in share elasticity over the positive range of X_{ki} values.

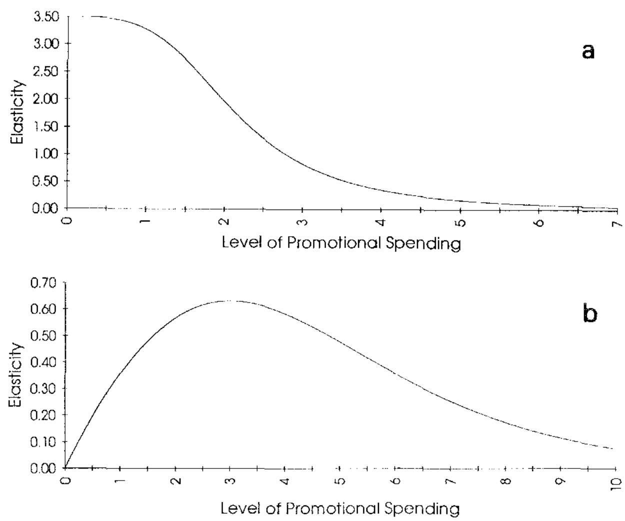

Fig. 6.1. Share elasticities for (a) MCI and (b) MNL models.

The share is assumed to increase as X_{ki} increases. Accounting for this share increase, the share elasticity for the MCI model monotonically declines as X_{ki} increases (Figure 6.1a), while that for the MNL model increases to a point and then declines. Which expression is a better one for share elasticity depends on the nature of the explanatory variable X_{ki}. The relevant issue is how share elasticity should behave for low values of the variable. If X_{ki} is product price, for example, it is more likely that share elasticity is fairly large even when price is near zero. Hence, one would be inclined to use the MCI model for price. On the other hand, if the variable is advertising expenditure, it is not unreasonable to assume that, at an extremely low level of expenditure, advertising is not very effective. This assumption, of course, leads to the adoption of the MNL model for advertising expenditure. Gruca & Sudharshan [1991], however, point out some issues when using MNL forms of variable as a basis for resource allocation. For firms with less than half of the market, the MNL formulation of an advertising variable implies that it is optimal to increase spending as much as possible. While these authors advise using the MCI version for advertising, another interpretation is possible. It might in fact be optimal for a firm to allocate as much to advertising as possible until it gains 50% of the market. In any case, careful consideration of the nature of each explanatory variable could rightfully lead to a mixture of MCI and MNL variables within a single market-share model. The general attraction framework encompasses both MCI and MNL models.

General attraction model:

$$\mathscr{A}_i = \exp(\alpha_i + \varepsilon_i) \prod_{k=1}^{K} f_k(X_{ki})^{\beta k}, \tag{16}$$

$$s_i = \frac{\mathscr{A}_i}{\sum_{j=1}^{m} \mathscr{A}_j},$$

where f_k is a monotone transformation of X_{ki}. If one chooses an identity transformation for k (that is, $f_k(X_{ki}) = X_{ki}$), (16) becomes the MCI model; if f_k is an exponential function (that is, $f_k(X_{ki}) = \exp(X_{ki})$), then (16) becomes the MNL model. But there is no reason for one to have to choose either the identity or exponential transformation for all f_k's in (16). Depending on the nature of variable X_k, one should be free to choose either the identity or exponential transformation (or any other appropriate monotone transformation, for that matter). This is why in (16) f_k has subscript k.

3. Market shares and choice probabilities

So far we have chosen to treat market shares as an aggregate quantity, namely, the ratio of a firm's (brand) sales to the relevant industry sales. But, since aggregate sales are composites of many purchases made by individual buyers (consumers and industrial buyers), market-share figures must be related to individual buyers' choices of various brands. In analyzing the relationships between market shares and individual choice probabilities, we will have to consider the variability of two factors – *choice probabilities* and *purchase frequency* for individual buyers – over the population of buyers in the market. Let us first define those two concepts.

Suppose that each buyer purchases a number of units of the product per period. We will assume that the *purchase frequency* (i.e. the number of purchases per period) by an individual buyer is a random variable that has a statistical distribution. We shall call this distribution an individual *purchase-frequency distribution*, since it is defined for each buyer in the market. The specific form of this distribution does not concern us here except that it has its own mean (*mean purchase frequency*).

Let us assume, not unreasonably, that a buyer does not always purchase the same brand from a set of alternative brands. In additon, it is assumed that a buyer's choice of a brand at one *purchase occasion* is made independently from his/her previous purchase and the buyer's selection is governed by probabilities specific to each brand (i.e. each buyer's brand selection in a period follows a zero-order Bernoulli process). While this *zero-order assumption* has been debated for the last thirty years [Frank, 1962; Kuehn, 1962], and it is a more reasonable assumption for the *household* than for the individual within a household [Kahn, Morrison and Wright, 1986], it is a palatable starting point for these developments [Bass, Givon, Kalwani, Reibstein & Wright, 1984; Givon 1984]. The probabilities for alternative brands in the industry are called *individual choice probabilities*.

Whether or not the buyer's behavior is truly probabilistic or deterministic is not an issue here. A buyer's choice behavior may be largely deterministic, but the environmental conditions surrounding purchase occasions may be such that they involve probabilistic elements that, from the viewpoint of an outside observer, make the buyers' choices appear probabilistic. We also posit that the attractions of alternative brands affect choice probabilities. This is to be consistent with our position that brand attractions are the determinants of market shares.

We distinguish four cases regarding the homogeneity (or heterogeneity) of the buyer population with respect to *individual choice probabilities* and *purchase frequencies.*

Case 1: Homogeneous purchase frequencies and choice probabilities
If mean individual-purchase frequencies are equal for all buyers and the brand selection of every buyer in the market is governed by the same set of choice probabilities, it is rather obvious that the market share of a brand will be approximately equal to the choice probability for the brand. (Actually, it is the expected value of the market share of a brand which will be equal to the choice probability for the brand.) In this case market shares may be interpreted as individual choice probabilities. For example, if the market share for a brand is 30 share points, one may say that each buyer chooses this brand with a 0.3 probability.

Case 2: Homogeneous purchase frequencies and heterogeneous choice probabilities
The interpretation of market shares will have to be changed a little if each buyer has a different set of choice-probability values for alternative brands. We still assume that mean purchase frequencies are equal for all buyers in the market. Under those assumptions it is easy to show that the expected value of a brand's market share is equal to the (population) average of choice probabilities for that brand. In other words, a market share of 30 share points may be interpreted as meaning that the average of the choice probabilities across the buyer population is 0.3.

Case 3: Heterogeneous purchase frequencies and homogeneous choice probabilities
This is the case where, while a common set of choice probabilities is shared by all buyers, mean purchase frequencies vary over buyers and have a statistical distribution over the buyer population. In this case the expected value of a brand's market share is still equal to its choice probability.

Case 4: Heterogeneous purchase frequencies and choice probabilities
In this case both choice probabilities and purchase frequencies are assumed to be variable over the buyer population. We need to distinguish further two cases within this.

(a) Uncorrelated case: choice probabilities and purchase frequencies are uncorrelated (i.e. independently distributed) over the buyer population.
(b) Correlated case: choice probabilities and purchase frequencies are correlated over the buyer population.

If purchase frequencies and choice probabilities are uncorrelated, the expected value of market shares is, as is shown below, still equal to population averages of choice probabilities (as in the case of homogeneous purchase frequencies). Turning to the correlated case, one finds that market shares are no longer directly related to choice probabilities. Case 4(b) is perhaps more realistic for many products, since one often hears that so-called *heavy users* and *light users* exhibit remarkably different purchase behavior. Heavy users are said to be more discriminating in taste, to be more price-conscious, and to tend to purchase *family-size* or economy packages. It is not surprising, then, to find heavy users, preferring some brands or brand/size combinations to those preferred by light users. If there were differences in the value of choice probability for a brand between heavy and light users, individual purchase frequencies and choice probabilities would be correlated and the market share for the brand will be biased toward the choice probability values for heavy users simply because they purchase more units of the brand. Thus market shares and choice probabilities generally do not coincide in this case. The results above may be stated more formally. The expected value of unit sales for brand i is obtained by averaging (over the buyer population) individual purchase frequencies multiplied by the individual's choice probability for the brand. Hence the expected value of market share for brand i is given by:

$$\text{market share}_i = \frac{\text{average number of units purchased for brand } i}{\text{average purchase frequency (for all brands)}}$$

or

$$\mathrm{E}(s_i) = \frac{1}{\bar{\mu}} \int_0^{\infty} \int_0^1 \mu \pi_i g(\mu, \pi_i)\, \mathrm{d}\pi_i \mathrm{d}\mu \tag{17}$$

where

$\mathrm{E}(s_i)$ = the expected value of market share for brand i,
μ = the mean purchase frequency per period (per individual),
$\bar{\mu}$ = the population mean of μ,
π_i = the individual choice probability for brand i,
$g(\mu, \pi_i)$ = the joint density function for μ and π_i.

Equation (17) shows that the expected value of market share for brand i is a weighted average of choice probabilities (weights are individual mean purchase frequencies) divided by average (over individuals) mean purchase frequency $\bar{\mu}$. From (17) we directly obtain the following result.

$$\mathrm{E}(s_i) = \bar{\pi}_i + \mathrm{cov}(\mu, \pi_i)/\bar{\mu}$$

where

$\bar{\pi}_i$ = the population mean of π_i,
$\mathrm{cov}(\mu, \pi_i)$ = the covariance of μ and π_i.

This is because, by definition, $\mathrm{cov}(\mu, \pi_i) = \bar{\mu}\mathrm{E}(s_i) - \bar{\mu}\bar{\pi}_i$. This equation shows that in general $\mathrm{E}(s_i)$ is not equal to $\bar{\pi}_i$. Since $\mathrm{cov}(\mu, \pi_i)$ may be positive or negative, one cannot say if market shares are greater or smaller than population mean of the

Table 6.1
Relations between market shares and choice probabilities

Choice probabilities	Purchase frequencies	
	Homogeneous	Heterogeneous
Homogeneous	Case 1: $E(s_i) = \pi_i$	Case 3: $E(s_i) = \pi_i$
Heterogeneous	Case 2: $E(s_i) = \bar{\pi}_i$	Case 4(a) Uncorrelated: $E(s_i) = \bar{\pi}_i$ Case 4(b) Correlated: $E(s_i) = \bar{\pi}_i + \mathrm{cov}(\mu, \pi_i)/\bar{\mu}$

choice probabilities. But if μ and π_i are positively correlated, $E(s_i)$ is greater than $\bar{\pi}_i$. If the correlation is negative, $E(s_i)$ is less than $\bar{\pi}_i$. Note also that, if $\mathrm{cov}(\mu, \pi_i) = 0$ (that is, if there is no correlation between the market share and choice probability), then the expected market share and the average choice probability are equal. In other words, in the uncorrelated case the expected value of a brand's market share is equal to its average choice probability ($\bar{\pi}_i$). The foregoing results are summarized in Table 6.1.

It is apparent from the table that the only case where we are uncertain of the correspondence between market shares and choice probabilities is Case 4(b). This fact might tempt one to look at this case as an exception or anomaly, but it is probably the most prevalent condition in the market. A practical implication of the preponderance of Case 4(b) is that, for the purpose of market-share forecasts, it is not sufficient for one to be able to predict the choice behavior of individuals accurately; rather it becomes necessary for one to be able to predict choice probabilities for each different level of purchase frequencies.

Of course, the situation cannot be changed by merely assuming that μ and π_i are uncorrelated over the buyer population (Case 4(a)). Since μ and π_i are arithmetically related, that is, $\pi_i = \mu_i/\mu$ where μ_i is the expected number of units of brand i purchased by an individual, and $\sum_{i=1}^{m} \mu_i = \mu$ where m is the number of alternative brands in the industry, the assumption that $\mathrm{cov}(\mu, \pi_i) = 0$ (for all i) implies a very restrictive form of joint distribution for μ and π_i. Indeed, it may be shown that μ is distributed as a gamma function and the π_i's are jointly distributed as a Dirichlet distribution. No other distributional assumption will give $\mathrm{cov}(\mu, \pi_i) = 0$. (See Cooper & Nakanishi [1988, pp. 52–54] for the proof of this result.)

What does all this argument about the relationship between purchase frequencies and choice probabilities suggest to analysts and marketing managers? If the correlation between purchase frequencies and choice probabilities is suspect, there may be an aggregation problem. One should segment the market in terms of purchase frequencies and analyze each segment separately. One may discover that marketing instruments have different effects on different segments and may be able to allocate marketing resources more efficiently. Also forecasting of brand sales and market shares will become more accurate if market shares are forecast for each segment and weighted by the mean purchase frequencies for the segments to obtain the estimate of overall market shares. Segmentation analysis of this type, however, requires more refined data than the usual aggregate market-share data, such as consumer-diary or scanner-panel data. Such data are currently available in the *single-source* data sets, and an illustration integrating the aggregate store data and the information in the associated consumer panels is presented later.

3.1. *Individual choice probabilities*

The focus of this section is to relate individual choice probabilities to attractions of alternative brands. We can of course assume that the choice probability for a brand is proportional to its attraction, and obtain a result similar to *Kotler's fundamental theorem* discussed earlier. But there are other more axiomatic approaches to deriving choice probabilities, and here we will be dealing with two basic models which are closely related to each other. It may be added that the terms *attraction* and *utility* will be used interchangeably in this section.

3.1.1. *Constant-utility models*

The simplest model for choice probabilities is the constant-utility model which is also called the Luce model or Bradley–Terry–Luce model. Its basic assumption (or *axiom*) may be stated as follows.

Axiom 1. Let an object, x, be an element of the choice set (i.e. set of choice alternatives), C, and also of a subset of C, S (i.e. $S \subseteq C$). The probability that x is chosen from C is equal to the product of the probability that x is chosen from S and the probability that (an element of) S is chosen from C.

Luce [1959] calls this assumption the *individual choice axiom*, which may be expressed mathematically as:

$$\Pr(x|C) = \Pr(x|S)\Pr(S|C)$$

where $\Pr(x|C)$ is read as 'the probability that x is chosen from C'.

This *axiom* for choice probabilities leads to results similar to that of the market-share theorem for market shares. If we let

$$u_x = \frac{\Pr(x|C)}{\Pr(z|C)}$$

for an arbitrary object in C, then for two objects x and y in C

$$\frac{u_x}{u_y} = \frac{\Pr(x|C)}{\Pr(y|C)}$$

and this ratio does not change with the choice of z. Also, since

$$\sum_{y \in C} \Pr(y|C) = \Pr(x|C) \sum_{y \in C} \frac{u_y}{u_x} = 1$$

we have

$$\Pr(x|C) = u_x \Big/ \sum_{y \in C} u_y.$$

The quantity u_x is called the *constant utility* of object x, and presumably determined for each individual as a function of marketing activities for x.

This model formed a basis of various models of individual choice behavior [Huff, 1962; Haines, Simon & Alexis 1972; Nakanishi & Cooper, 1974] and was also implicitly adopted for many market-share models. But this model exhibits the so-called *independence from irrelevant alternatives* (IIA) property which produces some quite counterintuitive results. From the axiom we have

$$\frac{\Pr(x|C)}{\Pr(y|C)} = \frac{\Pr(x|S)}{\Pr(y|S)}$$

for any subset of S of C which contains both x and y. Since this relationship must hold for set $\{x, y\}$,

$$\frac{u_x}{u_y} = \frac{\Pr(x|\{x,y\})}{\Pr(y|\{x,y\})}.$$

This ratio is independent of the choice of z. Since z is supposedly irrelevant to the odds of choosing x over y, this has been called the *independence of irrelevant alternatives* (IIA) property. The classic counterexamples are from Debreu [1960]. Although Debreu proposed a record-buying situation, the issues are often more clearly illustrated using a transportation-choice example. Suppose a person is indifferent between riding on a red bus (RB) or a blue bus (BB) if offered just these two alternatives, but prefers riding a taxi (T) four-to-one over a red bus, if offered this pair, or four-to-one over a blue bus, if offered that pair of alternatives. The choice axiom would summarize this case by noting that $\Pr(\mathrm{RB}|\{\mathrm{RB},\mathrm{BB}\}) = 0.5$, $\Pr(\mathrm{T}|\{\mathrm{T},\mathrm{RB}\}) = 0.8$, and $\Pr(\mathrm{T}|\{\mathrm{T},\mathrm{BB}\}) = 0.8$. While it seems clear that the probability of choosing a taxi shouldn't decrease if it is offered in a choice set along with both a red bus and a blue bus (i.e. $\Pr(\mathrm{T}|\{\mathrm{RB},\mathrm{BB},\mathrm{T}\})$ should still be 0.8), the choice axiom insists that $\Pr(\mathrm{T}|\{\mathrm{RB},\mathrm{BB},\mathrm{T}\}) = 0.67$ and $\Pr(\mathrm{RB}|\{\mathrm{RB},\mathrm{BB},\mathrm{T}\}) = 0.16$. The choice axiom forces this so that the *ratio* of the utility of RB to T is *constant* regardless of the choice set in which they are offered.

The concept of *constant utility* for a brand and the IIA property are really two sides of the same coin. If we think of *utility* is an inherent property of a brand that doesn't change regardless of the context in which a choice is made, we will be trapped by the IIA property into counterintuitive positions. There are two ways out of this problem. First, we can explicitly consider how the context in which choices are made affects the attractions of the alternatives. This is the path we follow in discussing the *temporal distinctiveness* of marketing efforts. Second, we can consider *utility* to be a random variable, rather than a constant. This is the topic of the next section.

3.1.2. Random-utility models

A choice model can be based on the assumption that the attractions an individual feels toward various objects (in our application, brands in an industry) on each purchase occasion are random variables, and the individual selects the brand that happens to have the largest utility value among the alternatives on that occasion. This *random-utility model* is defined as follows. Let $U_1, U_2, \ldots, U_m$ be the utilities for alternative brands where m is the number of brands in the choice set, C, of all competing brands in the industry, and let $g(U_1, U_2, \ldots, U_m)$ be the joint density function for them. The probability that brand i is chosen on a purchase occasion is given by

$$\Pr(i|C) = \Pr(U_i \geqslant U_j \text{ for all } j \in C).$$

In order to evaluate this probability, however, one must evaluate an integral function. For three brands, the probability that brand 1 is chosen is given by the following integral:

$$\Pr(1|C) = \int_{-\infty}^{\infty} \int_{-\infty}^{u_1} \int_{-\infty}^{u_1} g(u_1, u_2, u_3)\, du_3\, du_2\, du_1. \tag{18}$$

Similarly, $\Pr(2|C)$ and $\Pr(3|C)$ are given by suitably changing the upper limits of integration. Integral (18) may be defined for any number of choice objects (e.g. brands).

A large number of variants of random utility models may be created from this definition by selecting different specifications for g. However, the usefulness of random-utility models is limited because, unless the density function g is so very special as to give an analytical solution, the evaluation of this integral will in general require numerical integration. For example, if g is a joint-normal density (known as a probit or multivariate-probit model), there is no analytical (or closed-form) solution to this integeral. The probit model is a reasonable model for many applications, but its use has been hampered by the fact that the evaluation of (18) for a large number of objects involves tedious numerical integration.

There is one noted exception, however, to the need for cumbersome numerical integration of (18). McFadden [1974] showed that, if the joint distribution for random utilities $\{U_1, U_2, \ldots, U_m\}$ is a so-called multivariate *extreme-value distribution* of type I, then Integral (18) has a closed-form solution. A multivariate extreme-value distribution takes the following form:

$$G(u_1, u_2, \ldots, u_m) = \prod_{i=1}^{m} \exp[-\exp(\alpha_i - u_i)]$$

where $\alpha_i (i = 1, 2, \ldots, m)$ are parameters. This distribution is convenient because the maximum value among a sample of random utilities $\{u_1, u_2, \ldots, u_m\}$ from this distribution is also distributed as an extreme-value distribution of the following form:

$$F(u_{\max}) = \exp\left[-\exp(-u_{\max}) \sum_{i=1}^{m} \exp(\alpha_i)\right]$$

where $u_{\max}$ is a realization of a new random variable,

$$U_{\max} = \max(U_1, U_2, \ldots, U_m).$$

Using this property, the distribution function for random variable

$$U_{\max^*_i} = \max\{U_j\text{: for all } j \neq i\}$$

is given by

$$F(u_{\max^* i}) = \exp\left[-\exp(-u_{\max^* i}) \sum_{j \neq i}^{m} \exp(\alpha_j)\right].$$

Then the probability that brand i is chosen at a purchase occasion is given by

$$\begin{aligned}
\Pr(i \mid C) &= \Pr(U_i > U_{\max^* i}) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{u_i} \mathrm{d}G(u_i)\mathrm{d}F(u_{\max^* i}) \\
&= \int_{-\infty}^{\infty} \exp[-\exp(\alpha_i - u_i)] \exp(\alpha_i - u_i) \\
&\quad \cdot \exp\left[-\exp(u_i) \sum_{j \neq i}^{m} \exp(\alpha_j)\right] \mathrm{d}u_i = \exp(\alpha_i) \Big/ \sum_{j=1}^{m} \exp(\alpha_j).
\end{aligned} \tag{19}$$

If we let the attraction of brand i, $\mathscr{A}_i$, be equal to $\exp(\alpha_i)$, this expression is similar to an MNL model. Indeed the foregoing argument has been used to derive MNL models for individual-choice behavior. However, one may derive an expression similar to more straightforward attraction models in Equation (6), if, instead of an extreme-value distribution of type I, one chooses an extreme-value distribution of type II, namely,

$$G(u_1, u_2, \ldots, u_m) = \sum_{i=1}^{m} \exp[-\mathscr{A}_i u_i^{-b}]$$

where $\mathscr{A}_i$ $(i = 1,2,\ldots,m)$ are parameters. To show this, first note that the distribution function for random variable

$$U_{\max^* i} = \max\{U_j\text{: for all } j \neq i\}$$

is given by

$$F(u_{\max^* i}) = \exp\left[-u_{\max^* i}^{-b} \sum_{j \neq i}^{m} \mathscr{A}_j\right].$$

Using this,

$$\begin{aligned}\Pr(i|C) &= \Pr(U_i > U_{\max^* i}) \qquad (20)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{u_i} \mathrm{d}G(u_i)\mathrm{d}F(u_{\max^* i})\\ &= \int_{-\infty}^{\infty} \exp[-\mathscr{A}_i u_i^{-b}](\mathscr{A}_i b u_i^{-b-1})\exp\left[-u_i^{-b}\sum_{j\neq i}^{m}\mathscr{A}_j\right]\mathrm{d}u_i\\ &= \mathscr{A}_i \Big/ \sum_{j=1}^{m} \mathscr{A}_j.\end{aligned}$$

Thus individual-level versions of MCI models as well as MNL models are derivable if extreme-value distributions are assumed for the joint distribution of random utilities. So if we substitute an extreme-value distribution of type II, then what McFadden showed for MNL models can be generalized to MCI models.

Although both Equations (19) and (20) are derived for individual choice probabilities, one may derive an attraction model for aggregate market shares, if the definition of distribution functions is slightly changed. Suppose that random utilities for alternative brands, $U_1, U_2, \ldots, U_m$, are jointly distributed over the population of individual buyers, rather than within an individual. Each individual has a set of realized values for utilities, $u_1, u_2, \ldots, u_m$, and will select that brand which has the maximum utility value among m brands. Cast in this manner, the problem is to find the proportion of buyers who will purchase brand i, but Equations (19) and (20) give precisely this proportion (that is, market share) for two extreme-value functions.

Although random-utility models in general do not have the IIA property, it should be noted that some random-utility models do. Yellott [1977] proved that a random-utility model is equivalent to a constant-utility model (and hence possesses the IIA property) if and only if the joint distribution for random utilities follows a multivariate extreme-value distribution. The basic forms of MNL and MCI models belong to this special case. But we wish to emphasize that it is possible to construct attraction models of probabilistic choice which do not have the IIA property. We will discuss two such models – the *fully extended* and *distinctiveness* models – later in this chapter.

4. Asymmetries in markets and competition

Up to this point we have dealt with the simplest attraction models and the foundations that underlie them. These models have been symmetric in that a change in the market share of one leads to symmetrically distributed changes in the market shares of the other brands (i.e. each brand gains or loses from the actions of others in proportion to its market share). An even cursory observation of competitive interactions in the marketplace reveals that some firms (brands) are capable of exerting inordinately strong influence over the shaping of demand and competition, while other firms (brands) are not. On-package ads such as General Mills' Total Corn Flakes claim "more nutritious and better tasting than Kellogg's Corn Flakes" are offered in the managerial belief that going after specific competitors can have a disproportionate impact [Gibson, 1991]. Split-cable ad tests have shown some companies how to choose ads that will give more market-share boost for the same media buck. Recent work by Blattberg & Wisniewski [1989] emphasizes how national brands can draw market shares from regional brands much more readily than vice versa, and how regional brands can draw from economy brands more easily than economy brands can fight back. Further, all brands do not receive the same return from advertising or promotional expenditures. Such cases illustrate both differential effectiveness of brands and asymmetries in market and competitive structures. Differential effectiveness among brands reflects that firms (brands) have different degrees of effectiveness in carrying out their marketing activities. That such differences exist in real markets is obvious, but differential effectiveness alone does not create or reflect asymmetries. Shares will still be gained or lost according to the rules of the symmetric market. There are two fundamental sources of asymmetries. First, asymmetries are reflected in stable differential cross-effects between brands. Brands are differentially effective not only with respect to their own shares and sales, but also with respect to their ability to influence the shares and sales of other brands (that is, *clout*). Furthermore, brands seem to differ in the degree to which they are influenced by other brands' actions (that is, *vulnerability*). The second source of asymmetries derives from the *temporal distinctiveness* that a brand achieves in the marketplace by differentiating itself from the competition. We will first deal with differential effect and differential cross-effects, and look at market and competitive structure in terms of the patterns of clout and vulnerability of the competitors. Then we will address the issues concerning the temporal distinctiveness of marketing activity.

4.1. Differential effectiveness

In the initial discussion of differential effectiveness a solution at that time was to include some parameters in market-share models to take account of the overall *marketing effectiveness* of each brand. In the general specification of an attraction model,

$$\mathscr{A}_i = \exp(\alpha_i + \varepsilon_i) \sum_{k=1}^{K} f_k(X_{ki})^{\beta_k},$$

$$s_i = \mathscr{A}_i \Big/ \sum_{j=1}^{m} \mathscr{A}_j,$$

the parameters α_i $(i = 1, 2, \ldots, m)$ represented the marketing effectiveness of each brand. But the inclusion of the α's in attraction models, does not account fully for differential effectiveness among brands. The differential effectiveness may be specific to each marketing instrument, such as a brand that has a particularly effective pricing policy or an effective advertising campaign. The α_i's do not appear directly in the elasticity formulas for a particular marketing instrument, X_k (namely, $e_{s_i} = \beta_k(1 - s_i)$ for MCI models and $e_{s_i} = \beta_k X_{ki}(1 - s_i)$ for MNL models). The *marketing-effectiveness parameters* may reflect differences in *brand equity*, the *brand franchise*, or *brand loyalty* – literally, they are the constant component of each brand's attraction, but have nothing directly to do with elasticities. As a result, elasticity formulas for simple-effects attraction models do not reflect differential effectiveness.

If we wish to modify market-share elasticities to account for differential effectiveness, this may be achieved in only one way, that is, by specifying parameters β_k's in such a manner that each brand has a special parameter, β_{ki}, for variable X_k. The attraction model becomes:

$$\mathscr{A}_i = \exp(\alpha_i + \varepsilon_i) \prod_{k=1}^{K} f_k(X_{ki})^{\beta_{ki}}, \tag{21}$$

$$s_i = \mathscr{A}_i \Big/ \sum_{j=1}^{m} \mathscr{A}_j.$$

This modification does not change the basic structure of direct and cross-elasticities for attraction models. The self-elasticities, for example, are

MCI model:

$$e_{si} = \beta_{ki}(1 - s_i).$$

MNL model:

$$e_{s_i} = \beta_{ki} X_{ki}(1 - s_i).$$

As variable X_{ki} increases, the elasticity decreases for MCI models, but it increases and then decreases for MNL models. (The cross elasticities are given in the next section.) By expanding the parameterization of the model we are now able to capture brand-by-brand differences in market responsiveness to each element of the marketing mix. If all brands are equally effective then $\beta_{ki} = \beta_{kj} = \beta_k \; \forall \, i,j$, and the elasticity expressions reduce to those for simple-effects attraction models.

4.2. *Differential cross-elasticities*

If we are interested in knowing what effects other brands' actions will have on a brand's share, or what effects a brand's marketing actions will have on other brands' shares, then we need to consider *cross elasticities*. Suppose that brand j changed variable X_{kj} by a small amount ΔX_{kj}. The cross elasticity of brand i's $(i \neq j)$ share with respect to variable X_{kj} may be verbally expressed as 'the ratio of the proportion of change in market share for brand i corresponding to the proportion of change in variable X_k for brand j,' and is defined as follows:

$$e_{s_i \cdot j} = \frac{\Delta s_i / s_i}{\Delta X_{kj} / X_{kj}} = \frac{\Delta s_i}{\Delta X_{kj}} \cdot \frac{X_{kj}}{s_i}. \tag{22}$$

Note that $e_{si \cdot j}$ has two subscripts: the first indicates the brand that is influenced and the second, the brand that exerts influences. This is an *arc* cross-elasticity formula and the corresponding *point* cross-elasticity is defined as:

$$e_{s_i \cdot j} = \frac{\partial s_i / s_i}{\partial X_{kj} / X_{kj}} = \frac{\partial s_i}{\partial X_{kj}} \cdot \frac{X_{kj}}{s_i}. \tag{23}$$

Point cross-elasticities for differential-effects attraction models are as follows:

MCI model:

$$e_{s_i \cdot j} = -\beta_{kj} s_j.$$

MNL model:

$$e_{s_i \cdot j} = -\beta_{kj} X_{kj} s_j.$$

Let us consider what the above formulas imply. For the raw-score versions of both MCI and MNL models cross elasticities with respect to variable X_{kj} are constant for any brand $i (i \neq j)$. This means that the relative changes of other brands' shares (i.e. $\partial s_i / s_i$) caused by brand j's actions are the same for any brand, though actual changes in shares (i.e. ∂s_i) are different from one brand to another, depending on the current share level for each brand (i.e. s_i).

Table 6.2 shows the pattern of direct and cross elasticities we might get from a differential-effects attraction model. The direct elasticities are on the diagonal of

Table 6.2.
Elasticities for a differential-effects model

Brand	1	2	3	4	5	6
1	−1.2	0.6	0.5	0.1	0.3	0.7
2	0.3	−1.7	0.5	0.1	0.3	0.7
3	0.3	0.6	−0.9	0.1	0.3	0.7
4	0.3	0.6	0.5	−0.8	0.3	0.7
5	0.3	0.6	0.5	0.1	−1.6	0.7
6	0.3	0.6	0.5	0.1	0.3	−1.7

this matrix and the cross elasticities are off the diagonal – showing how the brand indicated by the column exerts influence on the brand indicated by the row. The diagonal entries can differ between columns due to different market shares between columns as well as due to differing parameters β_{kj}. The off-diagonal entries can also differ between columns due to different market shares as well as due to differing parameters β_{kj}. But within each column the off-diagonal entries are identical. This reflects the symmetry of Axiom 4 of the market-share theorem discussed earlier. It is a pattern that tells us that the differential-effects attraction model developed so far is an IIA model.

The equality of cross elasticities implied by such attraction models does not fit what we observe in the marketplace. There are brands that seem to be nearly immune from other brands' price changes; some brands seem to be able to ignore promotional activities of other brands with little loss of their shares, while others seem to be particularly vulnerable. It is therefore desirable to specify market-share models that reflect inequality of cross elasticities that we believe exists in the marketplace. As indicated before, there are two ways to attack this problem. On one hand, we could reflect the asymmetries which might arise from the *temporal distinctiveness* of marketing efforts. This is pursued in the section on the distinctiveness of marketing activities. The other way involves extending the parameters of the attraction model to reflect asymmetries due to *systematic and stable cross-competitive effects*. Fortunately, this can be accomplished with relative ease within the framework of attraction models as shown below.

The fully extended attraction model

$$\mathscr{A}_i = \exp(\alpha_i + \varepsilon_i) \prod_{k=1}^{K} \prod_{j=1}^{m} f_k(X_{kj})^{\beta_{kij}}, \tag{24}$$

$$s_i = \mathscr{A}_i \bigg/ \sum_{j=1}^{m} \mathscr{A}_j,$$

where β_{kij} is the parameter for the cross-competitive effect of variable X_{kj} on brand i.

Equation (24) is called an attraction model with differential cross-competitive effects or a *fully extended* attraction model to distinguish it from the differential-effects attraction model shown in Equation (21). The most important feature of the fully extended model is that the attraction for brand i is now a function not only of the brand's own actions (variables X_{ki}'s, $k = 1, 2, \ldots, m$) but also of all other brands' actions (variables X_{kj}'s, $k = 1, 2, \ldots, K$; $j = 1, 2, \ldots, m$). The β_{kij}'s for which i is different from j are the *cross-competitive effects* parameters, which partly determine cross elasticities. The β_{kij}'s for which j equals i (i.e., β_{kii}) are *direct-effects* parameters and are equivalent to the β_{ki}'s in the differential-effects model (21). This notation is cumbersome, but it is necessary to keep track of who is influencing whom. Note that the fully extended model has many more parameters (with $m^2 \times K$ β_{kij}'s and m α_i's) than the original attraction model (with $K + m$ parameters) and the differential-effects model (with $mK + m$ parameters). But even for extremely large markets these parameters are possible to estimate by the procedures discussed in Cooper & Nakanishi [1988, Chapter 5]. Since we get $(m - 1)$ independent observations from each store each week, scanner data provides enough degrees of freedom to overdetermine most models. For example, there are around 100 brands of cereal. If we track price, feature, display, coupon and brand-specific effects we have to estimate around 40 100 parameters. While this is a daunting numerical task that no one has yet undertaken, data from only ten stores over a year's time provide enough independent observations (51 480) to overdetermine a solution. This, of course, assumes homogeneity of parameters across stores. If differences in store format (e.g. every-day low-price versus advertised specials) indicate this assumption is questionable, then data from only ten stores per format would still overdetermine a solution. When other, less obvious, forms of heterogeneity across stores exist, methods for discovering latent classes in other marketing arenas may be adaptable to the current context [cf. Kamakura & Russell, 1989; Ramaswamy & DeSarbo, 1990].

4.3. Properties of fully extended attraction models

The fully extended attraction model is not an IIA because the choice between two brands can be affected by the cross-competitive influences of the other brands available in the market. To see what market and competitive structures can be revealed by the fully extended model (24), let us look at the direct and cross elasticities for this model.

MCI model:

$$e_{s_i \cdot j} = \beta_{kij} - \sum_{h=1}^{m} s_h \beta_{khj}.$$

MNL model:

$$e_{s_i \cdot j} = \left(\beta_{kij} - \sum_{h=1}^{m} s_h \beta_{khj} \right) X_{kj}.$$

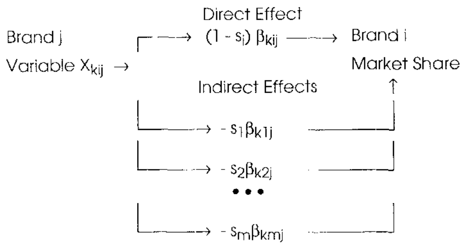

Fig. 6.2. Cross elasticities in the fully extended model.

These formulas apply to both direct and cross elasticities. If i is equal to j, then we have the direct elasticities for brand i, otherwise we have the cross elasticity, $e_{s_i \cdot j}$ of market share for brand i with respect to changes in marketing variable X_{kj} for brand j, and it is given by β_{kij} minus the weighted average of β_{khj}'s over h, where the weights are the market shares of respective brands (s_h) [cf. McGuire, Weiss & Houston, 1977]. Figure 6.2 describes the effects that influence market shares according to the fully extended model.

Assume that variable X_{kj} is the price for brand j. Then parameter β_{kij} for which i is not equal to j is likely to have a positive value. When brand j reduces its price the share of brand i tends to decrease. This effect of brand j's price change on brand i's share is shown in Figure 6.2 as a *direct competitive effect*. Note that a direct competitive effect is weighted by the complement of brand i's share (as we show below). When brand i's share is large, brand i is less affected directly by the moves by brand j. The influence of brand j's price change is not limited to the direct competitive effect on brand i, however. When brand j reduces its price, its own share should increase. Furthermore, the market shares of brands 1 through m (other than brands i and j) will also receive a negative effect, which in turn should have a positive effect on brand i's share. Influences of these kinds are *indirect competitive effects* shown in Figure 6.2.

In order to examine formally the points raised above, rewrite the cross-elasticity formula for MCI models as follows:

$$e_{s_i \cdot j} = (1 - s_i)\beta_{kij} - s_j\beta_{kjj} - \sum_{h \neq i,j}^{m} s_h\beta_{khj}.$$

The first term, of course, represents the direct competitive effects. The second term shows the indirect competitive effects through brand j. The last term consists of indirect competitive effects through all other brands. If X_{kj} is brand j's price, one expects that $\beta_{kjj} < 0$ and $\beta_{kij} > 0$ (for $i \neq j$). Since the first and last terms are expected to be positive and the second term negative, we do not know the sign of $e_{s_i \cdot j}$. It is dependent on the relative size of $(1 - s_i)\beta_{kij} - s_j\beta_{kjj}$ and $\sum_{h \neq i,j}^{m} s_h\beta_{khj}$.

Consider the following special cases.

Case 1: All cross-elasticity parameters (β_{kij}, $i \neq j$) are zero. In this case, $e_{s_i \cdot j} = -s_j\beta_{kjj}$. This is the same as the cross-elasticity formula for the differential-effects MCI models.

Case 2: All cross-elasticity parameters ($\beta_{kij}, i \neq j$) are approximately equal. In this case,

$$\sum_{h \neq i,j}^{m} s_h \beta_{khj} \approx (1 - s_i - s_j)\beta_{kij}.$$

Then

$$e_{s_i \cdot j} \approx s_j(\beta_{kij} - \beta_{kjj}).$$

This suggests that $e_{s_i \cdot j}$ has the same sign as β_{kij}.

Case 3: β_{kij} is nearly zero, but

$$\sum_{h \neq i,j}^{m} s_h \beta_{khj} > s_j \beta_{kjj}.$$

In this case $e_{s_i \cdot j}$ may have a sign different from β_{kij}.

Case 3 is an interesting situation because, in this case, it is possible that brand i even gains share when brand j reduces its price. For Case 3 to occur brand j's share should be relatively small, but the impact of its actions on brands other than i must be large. (This brings to our mind an image of an aggressive small brand j which is frequently engaged in guerilla price-wars.) In addition, brand i must be reasonably isolated from the rest of the market, implying that it is a *niche*-er. This case illustrates the richness of market and competitive structures that can be revealed by the fully extended attraction models.

It may be added that if $i = j$, we may write

$$e_{s_i \cdot j} = (1 - s_i)\beta_{kii} - \sum_{h \neq i}^{m} s_h \beta_{khi}.$$

The first term represents the direct effect of X_{ki} on brand i's share. The second term gives the sum of all indirect competitive effects on brand i's share. This formula suggests a possibility that, even if the direct effect is negligible (e.g. β_{kii} is small), direct elasticity, $e_{s.i}$, may be sizeable due to the combination of indirect competitive effects. In other words, a brand may be able to increase its share merely by reducing other brands' shares. Simple-effects or differential-effects attraction models do not allow such a possibility. This is another indication of the descriptive richness of the fully extended attraction models.

While issues in estimation are beyond the mandate of this chapter, readers may have noted that not all of the parameters of the fully extended model are identified. McGuire, Weiss & Houston [1977] showed that only the deviations $\beta^*_{kij} = \beta_{kij} - \bar{\beta}_{k \cdot j}$ and the brand-specific parameters α_i are estimable. But Cooper & Nakanishi [1988, p. 145] showed that only β^*_{kij} is needed to estimate all the elasticities.

4.4. *Revealing competitive structures*

Once cross-elasticity parameters are introduced in market-share models, it becomes possible to specify market and competitive structures on the basis of cross elasticities among brands. An example will serve to illustrate this concept. Suppose that the marketing variable in question is price. One may estimate share elasticities with respect to price using a differential cross-elasticities market-share model. Table 6.3 shows the matrix of direct and cross elasticities among six brands in a hypothetical market.

Some readers familiar with matrices may manage to see the market and competitive structure with a little work. However, we can always graph the pattern of an elasticity matrix using the simplest form of a *competitive map* [Cooper, 1988]. If we think of brands as vectors emanating from the origin of a space, the stronger the cross elasticity between two brands, the more correlated those brands' vectors would be. The closer the competition the smaller the angle between the vectors for two brands. The more complementary two brands are, the more opposite (closer to 180°) these brands should be in the map. If two brands to not compete at all, the vectors should be at right angles to each other. If some subset of brands does not compete with any of the brands in the market, the noncompeting brands should reside at the origin of the space. This space will be a joint space – showing both how brands exert influence (i.e. *clout*) and how they are influenced by other brands' actions (*vulnerability*). The cross elasticities are represented in a competitive map as if they were the *scalar products* of the coordinates of the brands.

$$ {}_iE_j = {}_iV_t{}'{}_tC_j + {}_i\Delta_j \tag{25} $$

where ${}_iE_j$ is the matrix of cross elasticities, ${}_iV_t$ is a matrix showing the coordinates of the brands on dimensions that describe the vulnerability of the brands, ${}_tC_j$ is a matrix showing the coordinates of the brands on dimensions that describe the clout of the brands, and ${}_i\Delta_j$ is a matrix of errors. Singular-value decomposition can be used to obtain the smallest (least-squares) error for any given choice of the dimensionality of the competitive map.

The two-dimensional competitive map for Table 6.3 is given in Figure 6.3. We expect the marketing actions of a brand to help itself. In a map this would be

Table 6.3.
Elasticities for a fully extended attraction model

Brand	1	2	3	4	5	6
1	−1.2	1.0	0.2	0.2	0.3	1.2
2	0.9	−1.6	0.3	0.4	0.1	1.4
3	0.1	0.2	−0.4	0.6	0.2	0.2
4	0.2	0.3	0.0	−1.7	1.1	0.3
5	0.3	0.5	0.4	1.0	−1.8	0.3
6	0.9	1.0	0.3	0.1	0.2	−0.7

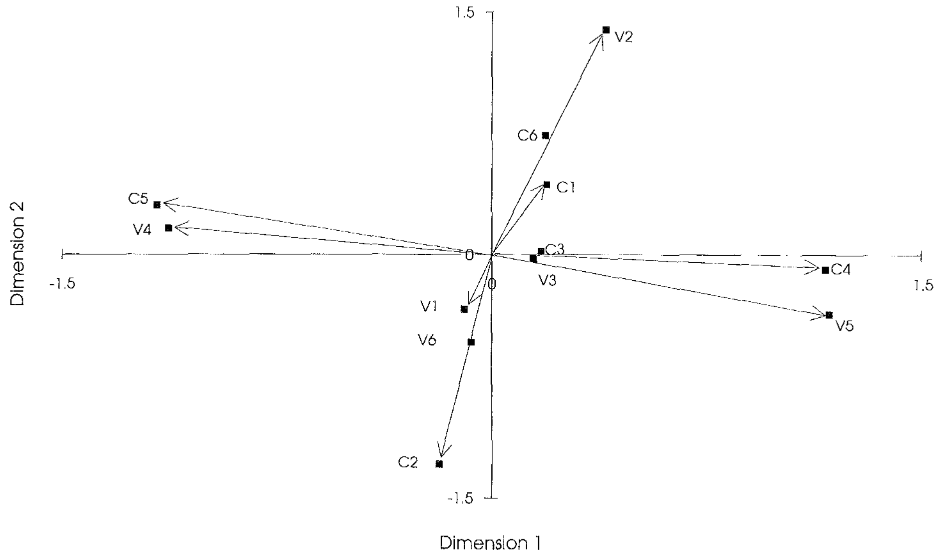

Fig. 6.3. Scaling of elasticities from Table 6.3, based on singular value decomposition.

reflected by having the vector representing the clout for one brand be about 180° from the vector representing vulnerability for that same brand. This occurs throughout Figure 6.3. Note that brand 2 is aligned to exert a good deal of competitive pressure on both brands 1 and 6; but these brands are positioned to return the pressure. Brands 4 and 5 are aligned to compete with each other, but there is not much competition between these two submarkets. Additionally, brand 3 is isolated at the origin – competing with neither group of brands.

This simple version of competitive mapping provides a snapshot of the competitive structure implied by a single matrix of elasticities. Mapping the elasticities for this marketing instrument over time and/or regions can give brand managers a view of what events produce changes in competitive patterns. And the full story on the competitive structure of this market is not known until the elasticities for each marketing instrument are mapped. Some analysts worry that looking at maps for different instruments will reveal brands as substitutes with respect to one instrument and complements with respect to another. This is certainly possible, but not necessarily undesirable. National brands seem to compete fiercely with each other on price maps, but ad maps might show that ads for one premium national brand help other premium national brands (at the expense of regional and private-label brands). Brand managers need to know if such patterns exist, for surely brand planning is affected by such knowledge.

4.5. Hierarchies of market segments

Ambiguities in interpreting the nature of competition from the tables of elasticities may be caused by the aggregation, that is, by not explicitly recognizing segments of buyers in the market. As was already pointed out, the overt pattern of brand grouping does not necessarily give hints about the underlying patterns of buyer demand. There are two possible interpretations on the nature of brand groups in Figure 6.3, for example. One interpretation is that two market segments exist and the cross elasticities reflect the difference in product perception between segments. The buyers who belong to the first segment might have only brands 1, 2 and 6 in their consideration set; those who belong to the second segment might consider brands 4 and 5 as the only relevant alternatives. In this interpretation brand groups correspond one-to-one with market segments of buyers.

The second interpretation of brand groups is that brands tend to be grouped in accordance with different types of buyer needs they serve. Suppose that the consumer uses plain and fruited yogurts for different occasions (e.g. plain yogurt for salad dressings and sauces, but fruited yogurt for desserts and snacks). This will cause the yogurt market to be divided into the plain and fruited brand groups, and minor price differences between the two groups will not affect demand for either. This type of segmentation on the basis of needs or a *benefit* segmentation, does not produce distinct buyer segments in the market, since the same household purchases both plain and fruited yogurts. Of course, if two brand groups serve two entirely isolated buyer needs, they should be treated as two distinct industries rather than one. But if the price of plain yogurt is drastically reduced, buyers might start cutting up their own fruit and the demand for fruited yogurt may be affected. A plain-yogurt brand with an aggressive price policy may have some cross elasticities with fruited brands, or vice versa. Moderate cross elasticites between groups would force one to treat them as a single market.

As in the above example, if the elasticities are measured only for the entire market, if will be impossible to establish the propriety of the above two interpretations solely on the basis of tables such as Table 6.3. In order to evaluate the correctness of these two interpretations, one will need data sets such as consumer panels (either diary or scanner panels). Moreover, it is desirable to have accompanying data on the buyer perception of alternative brands. Lacking such detailed data sets, however, one should at least understand well the aggregate implications of variabilities in elasticities among buyer segments. We will first look at the nature of elasticities in a multisegmented market.

Suppose that there are two segments in the market, containing N_1 and N_2 buyers, respectively. We sill use the notation $q_{i(l)}$ and $s_{i(l)}$ to indicate, respectively, the sales volume and market share of brand i in the lth segment. Since

$$s_i = [q_{(1)}s_{i(1)} + q_{(2)}s_{i(2)}]/q$$

where

$q_{(l)}$ = sales volume in segment l ($l = 1,2$),
q = total sales volume ($q_{(1)} + q_{(2)}$).

The point share elasticity of brand i with respect to X_{kj} is given by

$$e_{s_i \cdot j} = [(q_{(1)}/q)(\partial s_{i(1)}/\partial X_{kj}) + (q_{(2)}/q)(\partial s_{i(2)}/\partial X_{kj})](X_{kj}/s_i).$$

This shows that an overall elasticity is the weighted average of corresponding segment elasticities, weights being the relative sales volumes for respective segments. If we write the segment elasticity as $e_{s_i \cdot j(l)}$, then the general expression for $e_{s_i \cdot j}$ is given by

$$e_{s_i \cdot j} = \sum_{l=1}^{L} (q_{i(l)}/q_i) e_{s_i \cdot j(l)} \tag{26}$$

where L is the number of segments and q_i is the sales volume for brand i for the entire market. This expression gives one the means to compute the overall elasticity matrix from matrices for segments. It also hints at ways to redress the problems created by the heterogeneity in purchase frequency discussed earlier. If we segment the market into heavy users and light users, and parameterize an attraction model in each segment, we still can compute the overall elasticities from the segment elasticities.

We could achieve this segmentation by using panel data in each trading area to identify the heavy users and the light users in a category. The problem then becomes how to obtain an estimate of the market-level aggregate sales of a brand for each segment, when we have only the corresponding sales in the panel. Simple aggregation of the contemporaneous panel sales is not a good idea. Early research indicated that data aggregated this way did not allow for good estimation of cross-competitive effects. Panel data in a particular week is too sparse, even with large panels, to allow for reliable estimation of cross-competitive effects.

A simple remedy for this is to think that each panel segment has a long-term attraction to a brand as well as a contemporaneous attraction that incorporates the impact of the immediate promotional environment. Using long-term average sales in the panel segment and contemporaneous sales in each panel segment, we wish to develop weights that split the total sales for a brand in a period (S_{jt}) into parts that correspond to each segment. In the two-segment case we want to estimate weights $W_{\mathrm{L}jt}$ and $W_{\mathrm{H}jt}$ for that light- and heavy-user segments respectively, such that

$$W_{\mathrm{L}jt} + W_{\mathrm{H}jt} = 1,$$
$$W_{\mathrm{L}jt} \times S_{jt} = S_{\mathrm{L}jt},$$
$$W_{\mathrm{H}jt} \times S_{jt} = S_{\mathrm{H}jt}.$$

This way the segment sales we estimate add up to the known aggregate sales in the market. The theoretical relations could be specified in a simple linear model as below:

$$S_{\mathrm{L}jt} = \gamma_j \bar{P}_{\mathrm{L}j.} + \lambda_{\mathrm{L}j} P_{\mathrm{L}jt},$$
$$S_{\mathrm{H}jt} = \gamma_j \bar{P}_{\mathrm{H}j.} + \lambda_{\mathrm{H}j} P_{\mathrm{H}jt},$$

where

S_{Ljt} = the aggregate sales for brand j in period t in the light-user segment,
S_{Hjt} = the aggregate sales for brand j in period t in the heavy-user segment,
P_{Ljt} = the panel sales for brand j in period t in the light-user segment,
P_{Hjt} = the panel sales for brand j in period t in the heavy-user segment,
γ_j = a parameter expressing the influence of long-term panels sales for brand j in each segment,
λ_{Lj} = a parameter expressing the influence of the contemporaneous marketing environment on the relation between panel sales for brand j in period t and the corresponding aggregate sales for the light-user segment,
λ_{Hj} = a parameter expressing the influence of the contemporaneous marketing environment on the relation between panel sales for brand j in period t and the corresponding aggregate sales for the heavy-user segment.

The weights we desire could be based on a simple linear model such as

$$S_{jt} = \hat{\gamma}_j(\bar{P}_{Lj.} + \bar{P}_{Hj.}) + \hat{\lambda}_{Lj}P_{Ljt} + \hat{\lambda}_{Hj}P_{Hjt} + e_{jt}. \tag{27}$$

We then obtain estimates of the required weights from

$$\hat{W}_{Ljt} = \hat{S}_{Ljt}/\hat{S}_{jt}, \tag{28}$$

$$\hat{W}_{Hjt} = \hat{S}_{Hjt}/\hat{S}_{jt}. \tag{29}$$

A similar segmentation scheme could be developed for any number of mutually exclusive and exhaustive segmentation. With minor modification a scheme based on fuzzy segments could be developed. We end up with an estimated brand-sales vector for each segment. For all practical purposes we may calibrate segment models just as we do for the aggregate market sales. In this way information (such as manufacturer's coupon redemption) that has previously been the exclusive domain of panel data, can be incorporated into aggregate (store-level) attraction models. A numerical illustration of this is given in Section 5.

4.6. *Distinctiveness of marketing activities*

Fully extended attraction models have advanced our ability to reflect the complexity of market and competitive structures, but there are other aspects of competition which have not been properly dealt with even in the fully extended models. We will turn to some of the more critical issues in this section and the next one. Here we will take up the issue of *distinctiveness* of marketing activities by competing brands.

The main thesis of this section is that a brand's marketing actions must be distinct to be effective. Even casual observations bear out this proposition. Price reduction by a brand would have more effect on market shares when other brands' prices are kept high than it would when all competitors match the price reduction.

The market-share impact of one brand's promotion would be significantly greater when the brand is alone in promotion than it would when all brands engage in promotional activities.

If it is the differences between brands, rather than the absolute levels of marketing activities that materially affect buyers' preference, then we will have to bring the distinctiveness of marketing activities among brands into the market-share models. Attraction models handle the distinctiveness issue quite naturally.

Cooper & Nakanishi [1988, pp. 70–72] showed that one may express the variables in MCI and MNL models in a *deviation* form without changing the properties of the models. In other words, we may express a variable either as

$$X^*_{ki} = X_{ki}/\tilde{X}_k \quad \text{or} \quad X^*_{ki} = X_{ki} - \bar{X}_k$$

and substitute X^*_{ki} for X_{ki} in MCI or MNL models, respectively. This property of attraction models does not change if we move from the simple-effects form to differential-effects models and fully extended models, or if monotone transformations (f_k) other than identity or exponential are used. Substituting

$$f^*_k(X_{ki}) = f_k(X_{ki})/\tilde{f}_k(X_k),$$

where $\tilde{f}_k(X_k)$ is the geometric mean of $f_k(X_{ki})$ over i, for $f_k(X_{ki})$ in an attraction model, will not change the nature of the model. Thus, the variables in attraction models may be replaced by some equivalent form of deviations from the industry mean, and those models operate, in essence, on the principle of distinctiveness. If X_k is price in an MCI model, each brand's price may be expressed as deviations from the average price for the industry. If all brands charge the same price, X^*_{ki} will be equal to one, and price will not affect the shares of brands. Only when the prices for some brands deviate from the industry mean do they influence market shares of themselves and others.

The handling of distinctiveness by attraction models becomes technically difficult when the variable in question is qualitative. Product attributes are example of qualitative variables – a make of refrigerator may or may not have an ice-maker; or an automobile model may or may not have an automatic transmission. Such variables take only two values, typically, one if the product (or brand) has an attribute and zero if it does not. Of course, one may compute the industry average for a binary (two-valued) variable (which is the same as the proportion of products or brands that have the attribute) and subtract it from the value for each product/brand. But by this operation the transformed variable may take either positive or negative values, and hence it may be used only with an MNL model. In order to incorporate binary variables in an MCI model a simple, but effective, transformation – the *index of distinctiveness* – was developed [Nakanishi, Cooper & Kassarjian, 1974].

Suppose that X_k is a variable associated with the possession or nonpossession of an attribute. Let the proportion of products (or brands) in this industry which have the attribute be r. If there are ten brands and two brands have the attribute, r will be 0.2. The value of the index of distinctiveness for each brand is determined by the following simple operation.

If brand i has the attribute, $X_{ki} = 1/r$.

If brand i does not have the attribute, $X_{ki} = 1 - r$.

Thus if r equals 0.2, those brands with the attribute are given the value of 5 and those without the attribute will be given the value of 0.8. Note that the smaller r, the greater the value of X_k for those brands that have the attribute. This represents, in essence, the effect of the distinctiveness of a brand. If a brand is the only one which has the attribute the index value $(1/r)$ becomes maximal.

This index has a rather convenient property that it is ratio-wise symmetrical to the reversal of coding a particular attribute. If we reversed the coding of possession and nonpossession of an attribute in the previous numerical example, r would be 0.8, and the value of X_k for those brands with the attribute would be 1.25 $(= 1/0.8)$ and that for the brands without the attribute would be 0.2 $(= 1/5)$. In other words, those brands without the attribute become distinctive in the reverse direction.

The index of distinctiveness shown above transforms a binary variable such that it is usable in an MCI model. Cooper & Nakanishi [1983] found that this index is a special case of a more general transformation applicable not only for qualitative variables but also for any quantitative variable. First, convert any variable X_{ki} to a standardized score by the usual formula:

$$z_{ki} = (X_{ki} - \bar{X}_k)/\sigma_k$$

where

$\bar{X}_k$ = the arithmetic mean of X_{ki} over i,
σ_k = the standard deviation of X_{ki} over i.

Since standardized z-scores (z_{ki}'s) may take both positive and negative values, they may be used in an MNL model in the form of $\exp(z_{ki})$, but cannot be used in an MCI model. To create a variable usable in the latter model transform z-scores in turn in the following manner:

$$\zeta_{ki} = (1 + z_{ki}^2)^{1/2}, \quad \text{if } z_{ki} \geqslant 0,$$
$$\zeta_{ki} = (1 + z_{ki}^2)^{-1/2}, \quad \text{if } z_{ki} \leqslant 0. \tag{30}$$

This new transform, ζ_k, (to be called the *zeta-score* for X_k) takes only positive values and has a property that it is ratio-wise symmetrical when the positive and the negative directions of variable X_k are reversed. For example, let the value of ζ_{ki} be 2.5. If X_{ki} is multiplied by -1, ζ_{ki} will take a value of 0.4 $(= 1/2.5)$. It may be easily shown that the zeta-score includes the index of distinctiveness as a special case for binary variables. For a binary variable X_k, $\bar{X}_k = r$ and $\sigma_k = r(1-r)$. Hence

$$z_{ki}^2 = (1-r)/r \quad \text{if } X_{ki} = 1,$$
$$z_{ki}^2 = r/(1-r) \quad \text{if } X_{ki} = 0.$$

Substitution of the equivalent of z_{ki}^2's in the zeta-score formula yields squared roots of distinctiveness indices.

We can think of brands as point masses in space. The zeta-score is based on the ratio of the noncentral moment of inertia about brand i to the central moment of inertia – thus reflecting how an object stands out from a group. This ratio is not affected by a general linear transformation of X_k, making it an appropriate transformation of interval-scale ratings – thus allowing interval-scale rating to be used in MCI as well as MNL models. The ratio has a minimum value of one for brands at the center (i.e., the mean of X_k), and increases as a particular brand gets farther away from the center. To translate this ratio into a usable index we invert it at the mean of the underlying variable. This allows us to tell if a brand is distinctively high or distinctively low in an attribute compared to the other brands in the competitive offering. (Instead of thinking of brands as point masses in space, we could think of brands as points in space with masses equal to their market shares. This leads to a new, and as yet untested, version of distinctiveness that has great appeal to this one-time student of physics.) Figure 6.4 gives the comparison of the zeta-score with the $\exp(z_{ki})$ transform.

Although the shapes of the two transforms are quite similar, the choice between the two may be made by the form of the elasticities. The direct and cross elasticities for the $\exp(z_{ki})$ transforms are given by

$$E = (I - JD_s)BSD_X$$

and those for the zeta-transforms are given by

$$E = (I - JD_s)BSD_z$$

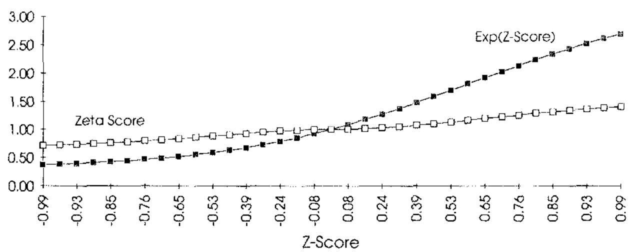

Fig. 6.4. Comparison of zeta and exp z.

where

S = the $m \times m$ matrix with elements $\{\partial z_{kj}/\partial X_{ki}\}$, i.e.,

$$S = \frac{1}{\sigma_k}\left[I - \frac{1}{m}J - \frac{1}{m}ZZ'\right],$$

D_s = an $m \times m$ diagonal matrix with the ith diagonal element s_i,
D_X = an $m \times m$ diagonal matrix with the ith diagonal element X_{ki},
σ_k = the standard deviation of X_k over i,
J = an $m \times m$ matrix of 1's,
Z = an $m \times 1$ vector of standardized scores (i.e. $z_{ki} = (X_{ki} - \bar{X}_k)/\sigma_k$),
D_z = an $m \times m$ diagonal matrix with the ith diagonal element $|z_{ki}|/(1 + z_{ki}^2)$.

Figure 6.5 compares the elasticities of the zeta-score with the $\exp(z_{ki})$ transform. The dip in the middle of the elasticity plot for zeta-scores corresponds to the flat portion of the zeta-score function depicted in figure 6.5. With zeta-scores, change is always depicted as slower in the undifferentiated middle portion of the distribution. Consider what this might imply for a frequently purchased branded good (FPBG). If it establishes an initial *sale price* about one-half a standard deviation below the average price in the category, the price is distinctively low and market-share change is relatively rapid. If the price drops further from this point, market share increases, but at a slower and slower rate. Bargain-hunting brand-switchers have already been attracted to the brand, and little more is to be gained from further price cuts. If the price increases from this initial sale price, market share drops rapidly at first, as the value of being distinctively low-priced is dissipated. At the undifferentiated position at the middle of the price distribution, market share is changing least rapidly as minor changes on either side of the average price go largely unnoticed. This indistinct region is similar to what DeSarbo, Rao, Steckel, Wind & Columbo [1987] represent in their friction-pricing model and similar to what Gurumurthy & Little [1986] discuss in their pricing model based on Helson's adaptation-level theory. On the high-priced side an analogous series of events happen. Small price increases around the average price are not noticed, but once the brand price is high enough to be distinguished from the mass, the loss of market share becomes more rapid. At some point, however, the change in market share must decline, as the brand loses all but its most loyal following.

In many categories of FPBG's the brands pulse between a relatively high self-price and a relatively low sale-price. In such cases the middle of the elasticity curve is vacant and the values of the elasticities for zeta-scores and exp(z-scores) might be quite similar. The exp(z-score) elasticities might be most descriptive of the path of market-share change from aggregate advertising expenditures, with increasing market-share growth as the expenditures move from zero up to the industry average, and diminishing growth rate for additional expenditures.

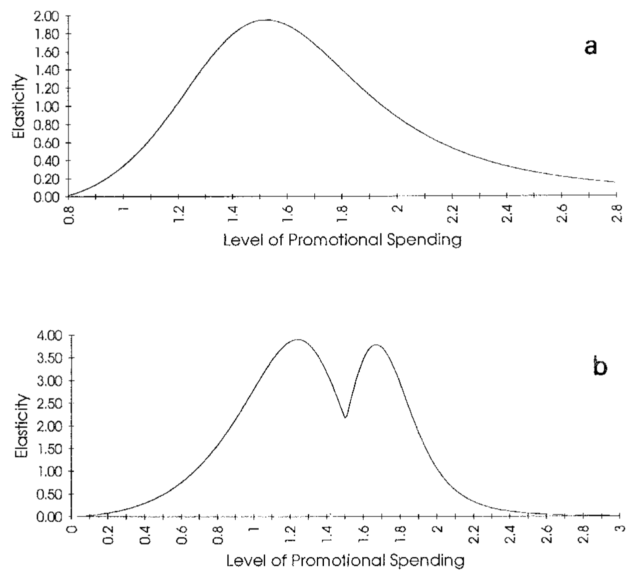

Fig. 6.5. Comparison of (a) zeta-score and (b) exp (z-score) elasticities.

Zeta-scores or exp(z-scores) explicitly model the *temporal distinctiveness* of brands by separating the underlying importance of a feature from the particular pattern of shared features in any given choice context (i.e. its *salience*). If two brands are both *promoted* in a store they do not each get the same boost in market share as if they were promoted alone. By specifically modeling such contextual effects we overcome the limitations imposed by the IIA (context-free) assumption of Luce choice models discussed earlier. Since the IIA assumption does not recognize that the value of a major promotion is somehow shared by all the brands on sale in that time period, the parameters of a raw-score (Luce-type) model will always reflect both the underlying value of the feature and the particular pattern of shared features in the contexts used for calibration. By using a distinctiveness index explicitly to model the changing pattern of feature-sharing from one period to the next, the parameters of the market-share model are free to reflect the underlying value of a feature. In forecasting, one again uses either zeta-scores or exp(z-scores) to help translate the underlying value of a feature to the particular pattern of shared features in new periods. So while the models in Nakanishi & Cooper [1974] were IIA models, as indicated above, the first real application of MCI models [Nakanishi, Cooper & Kassarjian, 1974] found a way out of the counterintuitive results imposed by the IIA assumption. All we need to do is provide a representation of how the context in which choices are made affects those choices.

The dramatic swings in market shares from week to week that analysts observe in scanner records of store sales encourages us to apply distinctiveness transformations in each time period to produce a generalized attraction model:

$$\mathscr{A}_{it} = \exp(\alpha_i + \varepsilon_i) \prod_{k=1}^{K} \prod_{j=1}^{m} f_{kt}(X_{kjt})^{\beta kij},$$

$$s_{it} = \mathscr{A}_{it} \Big/ \sum_{j=1}^{M} \mathscr{A}_{jt}.$$

Here f_{kt} has two subscripts to indicate not only a choice of functional form for the representation of an explanatory variable (i.e. MCI versus MNL representations), but also the kind of distinctiveness transformation to be applied within each period (i.e. zeta-scores or exp(z-scores)). Transformations such as exp(z-scores) and zeta-scores in each time period or choice situation not only highlight the distinctiveness of brands but serve to standardize variables. This reduces the collinearity otherwise inherent in the differential-effects forms of market-response models [cf. Cooper & Nakanishi, 1988, Chapter 5].

5. A numerical illustration from single-source data

The last decade has witnessed a tremendous explosion of research on how to use scanner data to learn about consumer and market response. The intense efforts at developing methods to deal with scanner data have fallen into two separate camps. On one hand, there has been extensive development of discrete choice (mainly logit-based) models and methods for understanding individual-level choice processes [Bawa & Ghosh, 1991; Bawa & Shoemaker, 1987; Bucklin & Lattin, 1991; Carpenter & Lehmann, 1985; Currim, Meyer & Le, 1988; Currim & Schneider, 1991; Elrod, 1988; Fader & McAlister, 1990; Grover & Srinivasan, 1989; Guadagni & Little 1983; Gupta, 1988, 1991; Jain & Vilcassim, 1991; Kalwani, Yim, Rinne & Sugita, 1990; Kamakura & Russell, 1989; Kannan & Wright, 1991; Krishnamurthi & Raj, 1991; Lattin & Bucklin, 1989; Meyer & Cooper, 1987; Moore & Winer, 1987; Neslin, 1990; Neslin, Henderson & Quelch, 1985; Pedrick & Zufryden, 1991; Ramaswamy & DeSarbo, 1990; Tellis, 1988; Vilcassim & Jain, 1991; Wheat & Morrison, 1990a; Winer, 1986, 1989; and Zufryden, 1984, 1987]. On the other hand, there have been substantial efforts to develop market-response models and methods that address both the vast size and detail of store-tracking data [Abraham & Lodish, 1987; Allenby, 1989, 1990; Bemmaor & Mouchoux, 1991; Blattberg & Wisniewski, 1989; Bolton, 1989; Bultez & Naert, 1988; Cooper, 1988; Cooper & Nakanishi, 1988; Doyle & Saunders, 1990; Kumar & Leone, 1988; Russell & Bolton, 1988; Shugan, 1987; Vilcassim, 1989; and Walters & MacKenzie, 1988]. The few efforts that have tried to use both panel and store data have been like Wheat & Morrison [1990b] which uses the percentage of times deals occur (computed from store-level data) in a panel-data investigation of purchase-timing models, or like Pedrick & Zufryden [1991] where store-level causal variables are used in panel-level modeling efforts. In spite of the common modeling frameworks (i.e. discrete MNL models and aggregate MNL models) used in some of these undertakings referenced above, no work has focused on bridging between these two research streams.

There are very pragmatic reasons for attempting to merge the two traditions. Consumer panels contain information that is missing from the store-tracking data. Panel data tell for example about brand-usage frequency, about redemption of manufacturers' coupons, and about television-advertising exposures – information that is simply not available in store-tracking databases. If such data are not incorporated into market-response models, the diagnostic value of the models is necessarily impaired, and the forecasting ability of the models may also be affected.

This section proceeds by:

(1) developing a segmentation scheme that divides the consumer panel into mutually exclusive and exhaustive groups based on usage frequency;
(2) aggregating panel variables to the segment level;
(3) estimating the store-level sales for each segment (for each brand);
(4) calibrating an asymmetric market-share model for the store-level sales of each segment;
(5) combining segment-level market-share models into a single forecast;
(6) calibrating and forecasting from an aggregate market-share model; and
(7) comparing forecasts and diagnostics.

5.1. Developing a segmentation scheme

While there are many useful bases for segmenting consumer panels [cf. Wind, 1978], one of the most fruitful areas for investigation of market-share models concerns differences in usage frequency. As pointed out earlier, parameters from aggregate market-share models will not reflect the individual choice processes if the heavy users in a category systematically purchase different brands than the light users (Case 4(b)).

Without a consumer panel to go along with store-tracking data a model-developer would not know if there was a correlation between brand choice and usage frequency. But with single-source databases come both the knowledge of the problem (if it exists) and the opportunity to address it through segmentation.

The problematic correlation between brand choice and purchase frequency can be detected with a simple χ^2 test on expected purchase frequencies. The theoretical market shares can be estimated from the long-term sales data in the store-tracking data or the panel (using the overall panel would eliminate any rejections due to the panel failing to track the store). The market shares times the total number of purchases in each panel segment provides the expected purchase frequency for each brand, while the observed purchase frequency is readily obtainable for each panel segment. The great size of the panels may make it very likely that trivial differences will be confirmed as statistically significant. Only experience over diverse categories will tell if extremely conservative testing procedure (i.e. extreme type-one error rates) should be invoked.

But before we can test for the correlation between brand choice and purchase frequency, we have to decide who are the heavy users and who are the light users. Cooper [1989a, 1989b] split Nielsen's single-source panels into heavy- and light-user segments in 19 different ways (depending on different definitions of how heavy the heavy users had to be). All but the most extreme definitions led to groups that were significantly discriminable on the basis of actionable variables such as income, number of members in the family, number of years at the same residence, number of TV's in the household, hours-per-week worked by the male head-of-household (HOH), highest degree obtained by male HOH, and number of hours-per-week worked by female HOH. Given this latitude, we decided to define heavy users as those panel households that purchased the top half of the total volume in the category. We know that in general it is tougher to predict more extreme events. This specification of segment membership means we will have ultimately to forecast less extreme volumes than with any other split.

In the catsup category 50% of volume is purchased by the top 20% of the households. In the yogurt category the top 10% of the households purchased 50% of the volume. We expect the concentration to vary from one category to another. Further, in any historical period the record of which households have purchased the most in a category is an imperfect indicator of who are the *real* heavy users. But since Schmittlein, Cooper & Morrison [1993] have shown that the longer the term on which manifest concentration is estimated the closer the manifest concentration is to the *true concentration*, our approach is to use all available panel data to classify households into heavy- and light-user segments.

Using the approximately 50–50 split of volume to define segments in the catsup market, the market and segment shares for Heinz, Hunts, Del Monte, and a combination of the private-label and generic catsups appear in Table 6.4. That the store shares for Heinz and Del Monte are slightly higher than either panel-segment share or that the store shares for the private-label and generic brands are slightly higher, indicates that to a minor extent the panels do not track the stores. While such minor discrepancies should not harm the overall analysis, in such a case the expected frequencies for the χ^2 should (and did) come from overall panel sales. The χ^2 for the independence of segment and share is 372 with 3 degrees of freedom. The significance indicates that we are in a Case 4(b) situation and segmentation should be undertaken. This extreme significance, however, is due

Table 6.4.
Store and segment market shares

Brand	Store share (%)	Heavy-segment share (%)	Light-segment share (%)
Heinz	67.40	68.79	68.25
Hunts	15.80	15.98	15.31
Del Monte	7.60	8.15	8.52
PvtGeneric	9.20	7.08	7.91

mainly to the large size of the panel, since the Φ coefficient is only 0.019 for this same cross tabulation.

5.2. *Aggregate panel variables to the segment level*

The kinds of variables that would be useful to incorporate from the panel file into a market-response analysis are relatively obvious. We aren't interested in the number of toasters or dogs in the household, not only because we have no hypotheses as to why certain brands would differentially benefit from the various levels, but more fundamentally because these kinds of variables do not vary across weeks to any appreciable degree. They would act like segment-specific constants or intercepts.

The obvious candidates for inclusion are those variables we wished were in the store-tracking database to begin with. Highest on this list would be information on manufacturer's coupon redemptions. While store-coupon redemptions have been reported, and used in market-share analysis [cf. Cooper & Nakanishi, 1988], manufacturer-coupon redemptions have not. It may be that inclusion of this information will alter our understanding of price or store-coupon sensitivity.

A less obvious, but no less important, variable is commercial exposures. GRP's or TRP's could be recorded in a store-tracking databse. But it is not obvious how commercial exposures would be routinely included in anything but a panel database. In the catsup category both manufacturer coupons and commercial exposures were initially aggregated for each segment. But in the catsup category too few brands engaged in TV advertising to obtain meaningful results. Thus only manufacturer coupons are used in this first application.

5.3. *Estimating the store-level sales for each segment (for each brand)*

There are three basic assumptions that lead to the estimation Equations (27)–(29).

(1) The long-term average panel sales for a brand has an influence on contemporaneous store sales. This implies that even if there are no sales in the panel segment for a particular brand in a given week there still might be sales in the corresponding store segment. The higher the historic average sales the more likely it is that there will be sales in the store segment *ceteris paribus*, even in the absence of sales in the panel segment.

(2) The contemporaneous panel sales for a brand has an influence on contemporaneous store sales. Of course, we expect the contemporaneous brand sales in a panel segment to be positively related to brand sales for the corresponding store segment. Each panel segment can have its unique sensitivity, i.e. the regression weight that relates contemporaneous panel sales for the heavy panel segment to contemporaneous store sales for the heavy store segment may be different from the regression weight that relates the corresponding sales figure for the light segments in the panel and in the store.

(3) The sensitivity of contemporaneous store sales to long-term panel sales is

the same within brand across segments. There is no reason that would force us to speculate that long-term historic sales for a brand affects one segment differently than another. We do, of course, expect different levels of average sales across segments, but this restriction is much more specific than that. It merely says that the sensitivity of contemporaneous sales to historic sales for a brand is the same for each panel segment. This amounts to a necessary restriction on the intercept for the regression model.

The regression model in (27) was estimated using the Nielsen single-source database tracking catsup sales in Sioux Falls, SD, for 138 weeks. Table 6.5 presents the regression-model results for Heinz, Hunts, Del Monte, control (private-label) brands, generic catsup, and aggregation of all other branded catsups, and an aggregation of all catsup called ALLBRAND in the analysis. The top 20% of users in this market purchased approximately 53% of the catsup volume. These households constituted the heavy-users segment in the panel.

The models have very high R-square values (except for the aggregate of all other branded catsup), but one must be careful to recognize that since the normal intercept is suppressed, R-square is more like a congruence coefficient than the coefficient of multiple determination we expect. All of these models are highly significant when compared to the simulation results for congruence coefficients reported by Korth & Tucker [1975]. All of the parameters are also statistically significant. These tables also report the split of the baseline volume into the average historical weekly volume (in ounces) for each panel segment.

Table 6.5.
Summary of regression analyses – Sioux Falls panel

Brand	R-square	F(3, 138)	Alpha baseline	Beta light	Beta heavy	Average volume	
						Light	Heavy
Heinz	0.976	1827	8.596	15.137	10.562	2403	2784
(t-values)			10.811	5.248	3.429		
Hunts	0.965	1233	8.273	11.786	15.617	539	647
(t-values)			11.98	5.837	8.761		
Del Monte	0.923	539	5.024	14.287	13.298	300	330
(t-values)			6.095	6.746	7.314		
Control-brands	0.967	1315	13.381	8.623	14.585	233	243
(t-values)			17.993	5.173	10.519		
Generic	0.931	604	26.599	11.813	6.014	45	44
(t-values)			17.95	7.504	3.11		
All others	0.578	62	10.496	4.737	6.843	3	3
(t-values)			6.401	4.301	6.283		
All brands	0.983	2645	9.719	8.545	14.798	3524	4050
(t-values)			11.01	3.268	5.024		

5.4. Calibrating an asymmetric market-share model for the store-level sales of each segment

Remember we have created one sales stream that we believe corresponds to the part of total brand sales that was purchased by the entire light-users segment, and a corresponding sales stream for the entire heavy-users segment. Except for the aggregated information from the panel file (which in this illustration concerned only the redemption of manufacturer's coupons), the causal environment encountered by these segments is the same. So we calibrated an asymmetric MCI model using the procedures outlined in Cooper & Nakanishi [1988, p. 168] and Carpenter et al. [1988]. Basically these procedures estimate all differential effects and brand-specific intercepts, search among the residuals for potentially significant cross effects, and finally re-estimate all effects using weighted least squares. Given the tiny shares of the aggregate representing all other branded catsups, this aggregate was dropped from the analysis. The private-label and generic catsups were combined into an aggregate simply called Private Labels. Each brand was represented by a brand-specific intercept (Int) and a differential effect for line ads (Ad-L), major ads (Ad-M), end-of-aisle display (D-EA), front-aisle display (D-FA),

Table 6.6.
Summary of market-share model – light-users segment

Brand	Int	Ad-L	Ad-M	D-EA	D-FA	D-IN	D-AO	Cp-S	Cp-M	Price
Heinz	2.00	0.14	+NS	+NS	0.29	+NS	+NS	0.57	−NS	−0.49
Cross effects	Heinz is less influenced by Del Monte's D-FA, Heinz is more influenced by Hunt's Cp-M, Heinz's D-AO has more influence on private labels, than predicted by the symmetric model.									
Hunts	−NS	0.26	0.36	0.22	−NS	0.24	0.28	0.72	0.16	−NS
Cross effects	Del Monte's Price has more influence on Hunts, Heinz is more influenced by Hunt's Cp-M, Del Monte's D-EA has more influence on Hunts, Private labels' D-EA has less influence on Hunts, Hunt's Price has more influence on private labels, than predicted by the symmetric model.									
Del Monte	0.00	−NS	0.29	0.75	1.09	0.15	0.38	−NS	0.51	−0.61
Cross effects	Heinz is less influenced by Del Monte's D-FA, Del Monte's Price has more influence on Hunts, Del Monte's D-EA has more influence on Hunts, Private labels' Price, D-EA and Cp-M have more influence on Del Monte, than predicted by the symmetric model.									
Private labels	−0.62	−NS	0.17	0.35	+NS	0.27	+NS	0.68	+NS	−0.73
Cross effects	Heinz's D-AO has more influence on private labels, Private labels' D-EA has less influence on Hunts, Hunt's Price has more influence on private labels, Private labels' Price, D-EA and Cp-M have more influence on Del Monte, than predicted by the symmetric model.									

in-aisle display (D-In), an aggregate of all other displays (D-AO), percent of brand volume sold on store coupon (Cp-S), and price (Price). Manufacturer's coupon (Cp-M) was represented as the percent of panel-segment sales on which a manufacturer's coupon was redeemed. All variables were represented as exp(z-scores).

Tables 6.6 and 6.7 report the parameters estimated for the market-share models for the light and heavy segments, respectively. Both models fit very well in calibration. The light-segment R-square is 0.824 ($F^{52}_{3210} = 288$). The heavy segment R-square is 0.828 ($F^{48}_{3214} = 323$). One would hope that this would be the case since the cross-competitive effects are selected so that they are very likely to be significant in the calibration data set (approximately two-thirds of the store weeks are used for calibration). The models also fit quite well in cross calibration in which the specification developed in the calibration dataset is applied to the remain one-third of the store weeks. Here the R-squares were 0.803 and 0.814 respectively. And finally the models also cross-validated quite well. The cross validation used the parameter values form the calibration dataset to forecast market shares for the untouched data in the final one-third of the store weeks. Here we form a single composite for each segment and perform a simple regression on the appropriately transformed (log-centered) dependent measure. The R-squares are 0.756 and 0.766 respectively for models with one predictor and 1607 degrees of freedom. The comparison to the aggregate model is discussed in Section 5.7. The cross effects are described verbally, by comparison with what would be expected from a symmetric market-share model [cf. Bell, Keeney & Little, 1975]. In symmetric market-share models, when one brand loses share the other brands are represented as if

Table 6.7.
Summary of market-share-model – heavy-users segment

Brand	Int	Ad-L	Ad-M	D-EA	D-FA	D-IN	D-AO	Cp-S	Cp-M	Price
Heinz	2.30	0.15	+NS	0.07	0.22	0.03	+NS	0.61	0.11	−NS
Cross effects	Heinz is less influenced by Del Monte's D-FA and private labels' Cp-M than predicted by the symmetric model.									
Hunts	−NS	0.21	0.34	0.21	+NS	0.24	0.25	0.71	0.16	−0.50
Cross effects	Hunts is less influenced by private labels' Price than predicted by the symmetric model.									
Del Monte	0.00	−NS	0.44	0.78	1.07	+NS	0.32	−NS	0.26	−1.13
Cross effects	Del Monte's D-FA has less influence on Heinz, Private labels' D-EA has more influence on Del Monte, and Del Monte's Price has less influence on private labels than predicted by the symmetric model.									
Private labels	−0.87	+NS	0.22	0.38	−NS	0.35	0.16	0.54	−NS	−1.30
Cross effects	Heinz is less influenced by private labels' Cp-M, Hunts is less influenced by private labels' Price, Private labels' D-EA has more influence on Del Monte, and Del Monte's Price has less influence on private labels than predicted by the symmetric model.									

they gain share strictly in proportion to their prior market shares. Further interpretation is deferred to the section below where diagnostics are also compared.

5.5. *Combining segment-level market-share models into a single forecast*

Combining the segment-level forecasts mentioned in the previous section into a single market-share forecast is relatively straightforward since we already have the weights from Equations (28) and (29). Combined forecast of market share for each brand merely applies these weights to the estimated brand shares from the segment models.

The variance accounted for (VAF) in the forecast period of the combined forecast is over 71%. This is a strong indication of very good forecasting ability. The root-mean-squared error (RMSE) of 0.128 might seem high, but we must remember that there are tremendous swings in market share when looked at on a week-by-week and store-by-store basis.

5.6. *Calibrating and forecasting from an aggregate market-share model*

Excluding the manufacturer's coupons, the procedures for specification of the aggregate market-share model were the same as those for the segment-level models. The R-square in calibration was 0.845 ($F^{45}_{3217} = 390$), in cross calibration

Table 6.8.
Summary of market-share model – aggregate market

Brand	Int	Ad-L	Ad-M	D-EA	D-FA	D-IN	D-AO	Cp-S	Cp-M	Price
Heinz	2.06	0.17	0.07	0.08	0.20	+NS	0.13	0.60		−0.70
Cross effects	Heinz is less influenced by Del Monte's D-FA, Heinz is less influenced by Hunt's Price, than predicted by the symmetric model.									
Hunts	−NS	0.26	0.35	0.19	+NS	0.29	0.22	0.84		−0.40
Cross effects	Heinz is less influenced by Hunt's Price, Del Monte's Price has more influence on Hunts, Private labels' D-IN has less influence on Hunts, than predicted by the symmetric model.									
Del Monte	0.00	−NS	0.38	0.75	1.08	+NS	0.33	−NS		−0.57
Cross effects	Heinz is less influenced by Del Monte's D-FA, Del Monte's Price has more influnce on Hunts, Private labels' Price and D-EA have more influence on Del Monte, than predicted by the symmetric model.									
Private labels	−0.79	−NS	0.19	0.35	−NS	0.39	0.17	0.61		−0.84
Cross effects	Private labels' D-EA has more influence on Hunts, Hunt's Price has more influence on private labels, Private labels' Price, D-EA and Cp-M have more influence on Del Monte, than predicted by the symmetric model.									

was 0.816, and in cross validation was 0.772 – just slightly higher in all cases than the corresponding values for the segment-level models.

The parameter values are reported in Table 6.8. Here too the cross effects are described verbally, in relation to what we expect from a symmetric market-share model. Further interpretation of the parameters is postponed to the next section.

5.7. *Comparing forecasts and diagnostics*

Table 6.9 compares the summary statistics on model calibration, cross calibration, cross validation, and forecasts. The similarities are remarkable. Although a slight edge might seem to go to the aggregate model (there are no statistical tests to compare these R-square values), when it comes to forecasting, the models are identical to three decimal places.

The *condition index* is also reported for the models in each of the calibration datasets. This number is the ratio of the largest singular value (square root of the eigenvalue) to the smallest singular value in the sum-of-squares-and-crossproducts (SSCP) matrix for the (reduced-form) regression model. This index is discussed by Belsley, Kuh & Welsch [1980] as an indicator of the degree of collinearity in the regression system. These authors indicate that condition indices over 100 can cause 'substantial variance inflation and great potential harm to regression estimate' [p. 153]. Indices from 21 to 35 are moderate at worst and give further evidence that asymmetric market-share models that use exp(z-scores) [cf. Cooper & Nakanishi, 1988, pp. 141–143] are exempt from the warnings concerning collinearity first given by Bultez & Naert [1975].

We may begin the comparison of the diagnostic value of this approach by looking at the parameter values listed in Table 6.10. This table shows only the differential-effect parameters for the three models calibrated earlier. There are no statistical tests for the equality of parameters across equations of this sort. What

Table 6.9.
Comparison of market-share models

	Light-users segment R-square	Heavy-users segment R-square	Aggregate market R-square
Calibration	0.824	0.829	0.845
D.F.	52–3210	48–3214	45–3217
Condition index	32.5	34.8	21.6
Cross calibration	0.803	0.814	0.816
D.F.	51–1558	47–1562	44–1565
Cross validation	0.756	0.766	0.772
D.F.	1–1607	1–1607	1–1607
Forecast accuracy		*Combined*	*Aggregate*
	RMSE	0.128	0.128
	VAF	0.711	0.711

Table 6.10.
Brand summary – across models

Brand	Int	Ad-L	Ad-M	D-EA	D-FA	D-IN	D-AO	Cp-S	Cp-M	Price
Heinz										
Aggregate	2.06	0.17	0.07	0.08	0.20	+NS	0.13	0.60		−0.70
Heavy users	2.30	0.15	+NS	0.07	0.22	0.03	+NS	0.61	0.11	−NS
Light users	2.00	0.14	+NS	+NS	0.29	+NS	+NS	0.57	−NS	−0.49
Hunts										
Aggregate	−NS	0.26	0.35	0.19	+NS	0.29	0.22	0.84		−0.40
Heavy users	−NS	0.21	0.34	0.21	+NS	0.24	0.25	0.71	0.16	−0.50
Light users	−NS	0.26	0.36	0.22	−NS	0.24	0.28	0.72	0.16	−NS
Del Monte										
Aggregate	0.00	−NS	0.38	0.75	1.08	+NS	0.33	−NS		−0.57
Heavy users	0.00	−NS	0.44	0.78	1.07	+NS	0.32	−NS	0.26	−1.13
Light users	0.00	−NS	0.29	0.75	1.09	0.15	0.38	−NS	0.51	−0.61
Private labels										
Aggregate	−0.79	−NS	0.19	0.35	−NS	0.39	0.17	0.61		−0.84
Heavy users	−0.87	+NS	0.22	0.38	−NS	0.35	0.16	0.54	−NS	−1.30
Light users	−0.62	−NS	0.17	0.35	+NS	0.27	+NS	0.68	+NS	−0.73

we are undertaking is not a statistical comparison, but rather a comparison of how differently these models would be interpreted by managers having to develop brand plans in a competitive environment.

For convenience the parameters have been grouped into four classes. Those with no border are relatively stable across analyses. Note that 28 of the 40 groups (70%) fall into this stable class, which is reassuring in many ways. The brand-specific intercepts, line-ad parameters, front-aisle display parameters, and store-coupon effects all seem stable within brands across analyses.

The solid, dark lines highlight parameter groups with major differences in levels (although no changes in the pattern of statistical significance across analyses). For Del Monte the aggregate price parameter is −0.57, but the introduction of the manufacturer's coupon and segmentation leads to a heavy-user segment that is very price-sensitive ($\beta_{\mathrm{H}j\cdot\mathrm{Price}} = -1.13$) and somewhat coupon-sensitive ($\beta_{\mathrm{H}j\cdot\mathrm{Coupon}} = 0.26$), while the light-user segment is much less price-sensitive ($\beta_{\mathrm{L}j\cdot\mathrm{Price}} = -0.61$) and more opportunistic users of manufacturer's coupons ($\beta_{\mathrm{L}j\cdot\mathrm{Coupon}} = 0.51$) for *private-label* brands the manufacturer's coupon effect is not significant in either segment (as should be the case), but the segmentation still leads to major differences in the price parameter – with the heavy-user segment being much more price-sensitive than the light-user segment.

The dotted lines highlight groups in which the analyses reveal differences in the patterns of significance across analyses. In the aggregate analysis of Heinz there is a significant price effect. Segment-level analysis indicates that the heavy-user segment is not price-sensitive, but is sensitive to manufacturer's coupons, while the light-user segment is somewhat price-sensitive and not affected by manufacturer's coupons. For Hunts we see both segments are somewhat sensitive to the presence of manufacturer's coupons, but all of the price sensitivity that appears in the aggregate is due to the sensitivity of the heavy-users segment.

Particularly for Heinz (the dominant brand in the category) we see diagnostically different patterns comparing the segment analyses and the aggregate analysis. Major ads and other displays appear effective in the aggregate, but have no significant effect in either segment. The end-of-aisle displays that seem effective in the aggregate analyses influence the heavy users but not the light users. All of these differences lead us to believe that segment-based analyses provide a diagnostically richer picture of this brand and of the market – without a sacrifice in forecast accuracy.

5.8. *Conclusion*

The goal of this section was to illustrate a method for bridging the too-long-separate traditions of individual-choice modeling and market-response modeling. We have shown that some kinds of information in panel databases can be used to segment otherwise aggregate market-response models, and other kinds of information can be integrated into these segment models to provide a diagnostically richer representation of market and competitive influences. And we have shown that these diagnostically rich, asymmetric market-share models can be estimated without the fear of collinearity.

This was a modest effort involving segmentation by usage frequency and the integration of a single variable from the panel database. But the underlying methods are so simple and robust that more venturesome applications seem readily doable.

6. Issues facing attraction models

This section deals with three issues that attraction models need to confront in the coming decade. The first feals with the tension between tactical and strategic uses of these models. The second concerns incorporating effects with greatly varying temporal impacts. And the third concerns new areas of application.

6.1. *Strategy versus tactics*

Resource-allocation decisions are typically made at the brand (or strategic-business-unit) level. Optimal allocation rules are elasticity-based. So we need models calibrated at this level to support strategic decisions. Promotion-planning involves decisions such as what combinations of brand versions and sizes should be co-promoted. Should fruited 4 oz. yogurts be promoted with plain yogurt of all sizes? Which topping combination of frozen pizza is best to feature in an ad? Such questions require very detailed specification of market-share models. Cross effects may relate to aggregates of versions and sizes within a brand line. Current illustrations of market-share models are between these two extreme levels of specification.

What is needed is a way of specifying and estimating attraction models in great detail (to satisfy tactical requirements), and a means of aggregating elasticities over versions and sizes of a brand to satisfy strategic requirements. As computations become amazingly cheaper, estimating detailed models becomes more practical. But aggregating elasticities over brand versions or sizes (without estimating a more aggregate model) involves mathematical developments that have yet to be undertaken. Specification of a more aggregated model for strategy than for tactics seems an obvious alternative, but there is no guarantee that the complex of competitive forces that appears in detailed models will appear in models with more aggregated competitors. Detailed study of the competitive structures revealed by models of the same underlying data, at different levels of brand aggregation is needed.

6.2. *Differences in temporal impact*

Scanner data have helped us grow accustomed to calibrating models at the weekly level. The promotional environment is largely constant throughout a week within a trading area. Models that account for the huge swings in sales (and shares) that accompany promotions are useful for planning coupons, displays and newspaper features. But what about advertising? TV ads are much slower in showing any effect they might have. If they are effective, the impact is spread out over a much longer time period. In the numerical illustration of the previous section, advertising effects are probably lumped into the brand-specific effects α_i. So effects that are *slow* with respect to the periodicity of a calibrated model might show up as the *constant* component of a brand's attraction. This makes evaluating advertising effectiveness very difficult.

We are greatly aided in the search for a remedy, by the proof in Cooper & Nakanishi [1988, pp. 78–85] that distributed-lag models can be specified for both explanatory variables and lagged market-share effects without losing the logical-consistency property of attraction models, if the specification is done in the *log-centered* form of the variables. But even with this capability we might face a situation in which each week there are very weak, lagged advertising effects from the prior twelve (or more) weeks of advertising. With the possibility of differential lagged effects and cross-competitive lagged effects, the number of possible parameters could become unmanageable for dynamic forces that we expect to be weak in each time period.

Another alternative might be to estimate the brand-specific parameters as *sequentially varying parameters* with their values varying *systematically* as functions of advertising expenditures. Chapter 9 by Hanssens and Parsons discusses such models.

6.3. *New application areas*

The rapid evolution of single-source, scanner databases, with their rich description of the competitive influences affecting all brands in a market, has given new momentum to the study and development of attraction models. The trend is bound to continue and strengthen as manufacturers and retailers come to grips with the complexities of what it means to manage in an information-rich environment. But the new areas of application are likely to be far removed from the brand-management context that has dominated the discussion in this chapter so far.

Attraction models have their roots in the interface between marketing and geography [cf. Huff, 1962, 1963; Haines, Simon & Alexis, 1972]. How shoppers in neighborhood j were attracted to retail center i, was the type of problem broached in these early studies. But the Huff model was an IIA model incapable of reflecting asymmetries in attraction. We believe that some of the most challenging problems to which attraction models can be applied return to these roots, but in a much more mature form.

For example, the nations of the world can be thought of as centers of attraction. The asymmetric flows of trade between nations could be represented by extended attraction models. Even at an aggregate level implied by investigating the balance of trade in gross dollars (or other monetary units), asymmetric attraction models provide a framework for valuing the various forces that are theorized to drive trade imbalances (e.g. differences in labor rates, tariff barriers, barriers in the distribution channels). The relative values of currencies themselves could be represented with asymmetric attraction models. If these representations are *descriptive* as well as predictive, we would enrich the empirical basis for understanding currency markets.

There are fundamental inadequacies, however, in representing the asymmetric flows of resources between countries in terms of capital. Some capital has its roots in largely renewable resources, while other capital is generated by the depletion of finite and fixed resources. The implications of the differences in the resource base are grave, but glossed over by purely aggregate representations of capital flows. For example, the trade surpluses that most *first-world* countries have with many *third-world* countries drains capital from the third world. It makes a considerable difference if this need for capital is sated by the export of rapidly renewable resources (e.g. coffee, sugar beet, or other cultivated agricultural commodities), versus the diminishing supply of exotic hardwoods from the world's rainforests. It even makes a difference if the agricultural land is claimed by the destruction of rainforests as opposed to balanced reuse of existing agricultural lands.

The weighty and complex problems we face in establishing a sustainable balance of resource flows between nations requires modeling techniques rich enough to capture the diversity of the causal influences driving these problems. In the hands of creative researchers, asymmetric attraction models can help us understand the forces driving imbalance, and perhaps help provide insight into what can be done.

References

Abraham, M.M., and L.M. Lodish (1987). PROMOTER: An automated promotion evaluation system. *Marketing Sci.* 6(2), 101–123.

Allenby, G.M. (1989). A unified approach to identifying, estimating and testing demand structures with aggregate scanner data. *Marketing Sci.* 8(3), 265–280.

Allenby, G.M. (1990). Hypothesis testing with scanner data: The advantage of Bayesian methods. *J. Marketing Res.* 27 (November), 379–389.

Bass, F.M., M.M. Givon, M.U. Kalwani, D. Reibstein and G.P. Wright (1984). An investigation into the order of the brand choice process. *Marketing Sci.* 3 (Fall), 267–287.

Bawa, K., and A. Ghosh (1991). The covariates of regularity in purchase timing. *Marketing Lett.* 2(2), 147–157.

Bawa, K., and R.W. Shoemaker (1987). The effects of direct mail coupon on brand choice behavior. *J. Marketing Res.* 24 (November), 370–376.

Bell, D.E., R.L. Keeney and J.D.C. Little (1975). A market share theorem. *J. Marketing Res.* 12 (May), 136–141.

Belsley, D.A., E. Kuh and R.E. Welsch (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* Wiley, New York.

Bemmaor, A.C., and D. Mouchoux (1991). Measuring the short-term effect of in-store promotion and retail advertising on brand sales: A factorial experiment. *J. Marketing Res.* 28 (May), 202–214.

Blattberg, R.C., and K.J. Wisniewski (1989). Price induced patterns of competition, *Marketing Sci.* 8(4), 291–309.

Bolton, R.N. (1989). The relationship between market characteristics and promotional price elasticities. *Marketing Sci.* 8(2), 153–169.

Brodie, R., and C.A. De Kluyver (1984). Attraction versus linear and multiplicative market share models: An empirical evaluation. *J. Marketing Res.* 21 (May), 194–201.

Bucklin, R.E., and J.M. Lattin (1991). A two-state model of purchase incidence and brand choice. *Marketing Sci.* 10(1), 24–39.

Bultez, A.V., and P.A. Naert (1975). Consistent sum-constrained models. *J. Amer. Statist. Assoc.* 70 (September), 529–535.

Bultez, A.V., and P.A. Naert (1988). SH.A.R.P.: SHelf Allocation for Retailer's Profit, *Marketing Sci.* 7(3) 211–231.

Carpenter, G.S., L.G. Cooper, D.M. Hanssens and D.F. Midgley (1988). Modeling asymmetric competition, *Marketing Sci.* 7(4), 393–412.

Carpenter, G.S., and D.R. Lehmann (1985). A model of marketing mix, brand switching, and competition. *J. Marketing Res.* 22 (August), 318–329.

Cooper, L.G. (1988). Competitive maps: The structure underlying asymmetric cross elasticities. *Management Sci.* 34(6), 707–723.

Cooper, L.G. (1989a). Integrating scanner-panel and store-level data. Seventh Marketing Science Conference, Duke, March.

Cooper, L.G. (1989b). Toward integrating scanner-panel and store-level data. Joint UCLA–USC Marketing Colloquium, May, UCS, Los Angeles.

Cooper, L.G., and M. Nakanishi (1983). Standardizing variables in multiplicative choice models. *J. Consumer Res.* 10 (June), 96–108.

Cooper, L.G., and M. Nakanishi (1988). *Market-Share Analysis: Evaluating Competitive Marketing Effectiveness.* Kluwer Academic Publishers, Boston.

Currim, I.S., R.J. Meyer and N.T. Le (1988). Disaggregate tree-structured modeling of consumer choice data. *J. Marketing Res.* 25 (August), 253–265.

Currim, I.S., and L.G. Schneider (1991). A taxonomy of consumer purchase strategies in a promotion intense environment. *Marketing Sci.* 10(2) 91–110.

Debreu, G. (1960). Review of R.D. Luce's *Individual Choice Behavior: A Theoretical Analysis, Amer. Econom. Rev.* 50(1), 186–188.

DeSarbo, W.S., V. Rao, J.H. Steckel, Y. Wind and R. Columbo (1987). A friction model for describing and forecasting price changes. *Marketing Sci.* 6(4), 299–319.

Doyle, P., and J. Saunders (1990). Multiproduct advertising budgeting. *Marketing Sci.* 9(2), 97–113.

Elrod, T. (1988). Choice Map: Inferring a product-market map from panel data, *Marketing Sci.* 7(1), 21–40.

Fader, P.S., and L. McAlister (1990). An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data. *J. Marketing Res.* 7 (August), 322–332.

Frank, R.E. (1962). Brand choice as a probability process. *J. Business* 35, 43–56.

Ghosh, A., S. Neslir, and R. Shoemaker (1984). A comparison of market share models and estimation procedures. *J. Marketing Res.* 21 (May), 202–210.

Gibson, R. (1991). Cereal giants battle over market share. *Wall Street J.* December 16, B1–2.

Givon, Moshe M. (1984). Variety seeking through brand switching. *Marketing Sci.* 3 (Winter), 1–22.

Grover, R., and V. Srinivasan (1989). An approach for tracking within-segment shifts in market shares. *J. Marketing Res.* 26 (May), 230–236.

Gruca, T.S., and D. Sudharshan (1991). Equilibrium characteristics of multinomial logit market share models. *J. Marketing Res.* 28 (November), 480–482.

Guadagni, P.M., and J.D.C. Little (1983). A logit model of brand choice calibrated on scanner data. *Marketing Sci.* 2(3), 203–288.

Gupta, S. (1988). Impact of sales promotion on when, what, and how much to buy. *J. Marketing Res.*, 25 (November), 342–356.

Gupta, S. (1991). Stochastic models of interpurchase time with time dependent covariates. *J. Marketing Res.* 28 (February), 1–15.

Gurumurthy, K., and J.D.C. Little (1986). A pricing model based on perception theories and its testing on scanner panel data. Working Paper Draft, Massachusetts Institute of Technology, May.

Haines, G.H., Jr. L.S. Simon and M. Alexis (1972). Maximum likelihood estimation of central-city food trading areas. *J. Marketing Res.* 9 (May), 154–159.

Huff, D.L. (1962). *Determination of Intraurban Retail Trade Areas*, Los Angeles: Real Estate Research Program. University of California, Los Angeles.

Huff, D.L. (1963). A probabilistic analysis of consumer spatial behavior, in W.S. Decker (ed.), *Emerging Concepts in Marketing*. American Marketing Association, Chicago, pp. 443–461.

Jain, D.C., and N.J. Vilcassim (1991). Investigating household purchase timing decisions: A conditional hazard function approach. *Marketing Sci.* 10(1), 1–23.

Kahn, B.E., D.G. Morrison and G.P. Wright (1986). Aggregating individual purchases to the household level. *Marketing Sci.* 5(3), 260–268.

Kalwani, M.U., C.K. Yim, H.J. Rinne and Y. Sugita (1990). A price expectations model of customer brand choice. *J. Marketing Res.* 27 (August), 251–262.

Kamakura, W.A., and G.J. Russell (1989). A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* 26 (November), 379–390.

Kannan, P.K., and G.P. Wright (1991). Modeling and testing structured markets: A nested logit approach. *Marketing Sci.* 10(1), 58–82.

Korth, B., and L.R. Tucker (1975). The distribution of chance coefficients from simulated data. *Psychometrika*, 40(3), 361–372.

Kotler, P. (1984). *Marketing Management: Analysis, Planning, and Control, 5th Edition.* Prentice-Hall, Englewood Cliffs, NJ.

Krishnamurthi, L., and S.P. Raj (1991). An empirical analysis of the relationship between brand loyalty and consumer price elasticity. *Marketing Sci.* 10(2), 172–183.

Kuehn, A.A. (1962). Consumer brand choice – a learning process. *J. Advertising Res.* 2 (December), 10–17.

Kumar, V., and R.P. Leone (1988). Measuring the effect of retail store promotions on brand and store substitution. *J. Marketing Res.* 25 (May), 178–185.

Lattin, J.M., and R.E. Bucklin (1989). Reference effects of price and promotion on brand choice behavior. *J. Marketing Res.* 26 (August), 299–310.

Leeflang, P.S.H., and J.C. Reuyl (1984). On the predictive power of market share attraction models. *J. Marketing Res.* 21 (May), 211–215.

Luce, R.D. (1959). *Individual Choice Behavior: A Theoretical Analysis.* Wiley, New York.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. in P. Zarembka. *Frontiers in Econometrics.* Academic Press, New York, pp. 105–142.

McGuire, T.W., D.L. Weiss and F.S. Houston (1977). Consistent multiplicative market share models. in B.A. Greenberg and D.N. Bellinger (eds), *Contemporary Marketing Thought.* American Marketing Association, Chicago, pp. 129–134.

Meyer, R.J., and L.G. Cooper (1987). A longitudinal choice analysis of consumer response to a product innovation, in R. Golledge and H. Timmermans (eds), *Decision Making and Choice Behaviour: Behavioural Modeling.* Croom-Helm, London, pp. 424–451.

Moore, W.L., and R.S. Winer (1987). A panel-based method for merging joint space and market response function estimation. *Marketing Sci.* 6(1), 25–42.

Naert, P.A., and A. Bultez (1973). Logically consistent market share models. *J. Marketing Res.* 10 (August), 334–340.

Naert, P.A., and M. Weverbergh (1981), On the prediction power of market share attraction models. *J. Marketing Res.* 18 (May), 146–153.

Naert, P.A., and M. Weverbergh (1985). Market share specification, estimation and validation: Toward reconciling seemingly divergent views. *J. Marketing Res.* 22 (November), 453–461.

Nakanishi, M. (1972). Measurement of sales *promotion effect at the retail level*–a new approach. *Proceedings, Spring and Full Conference*, American Marketing Association. Chicago, pp. 338–343.

Nakanishi, M., and L.G. Cooper (1974). Parameter estimation for a multiplicative competitive interaction model – least squares approach. *J. Marketing Res.* 11 (August), 303–311.

Nakanishi, M., L.G. Cooper and H.H. Kassarjian (1974). Voting for a political candidate under conditions of minimal information. *J. Consumer Res.* 1 (September), 36–43.

Neslin, S.A. (1990). A market response model for coupon promotions, *Marketing Sci.* 9(2), 125–145.

Neslin, S.A., C. Henderson, and J. Quelch (1985). Consumer promotions and the acceleration of product purchase. *Marketing Sci.* 4 (Spring), 147–165.

Pedrick, J.H., and F.S. Zufryden (1991). Evaluating the impact of advertising media plans: A model of consumer purchase dynamics using single-source data. *Marketing Sci.* 10(2), 111–130.

Ramaswamy, V., and W.S. DeSarbo (1990). SCULPTURE: A new methodology for deriving and analyzing hierarchical product-market structure from panel data. *J. Marketing Res.* 27 (November), 418–427.

Russell, G.J., R.N. Bolton (1988). Implications of market structure for elasticity structure. *J. Marketing Res.* 25 (August), 229–241.

Schmittlein, D.C., L.G. Cooper and D.G. Morrison (1993). Truth in concentration in the land of 80/20 laws. *Marketing Sci.* 12(1), forthcoming.

Shugan, S.M. (1987). Estimating brand positioning maps from supermarket scanning data. *J. Marketing Res.* 24 (February), 1–18.

Tellis, G.J. (1988). Advertising exposure, loyalty, and brand purchase: A two-stage model of choice. *J. Marketing Res.* 25 (May), 134–144.

Vilcassim, N.J. (1989). Extending the Rotterdam model to test hierarchical market structures. *Marketing Sci.* 8(2), 181–190.

Vilcassim, N.J., and D.C. Jain (1991). Modeling purchase timing and brand switching behavior incorporating explanatory variables and unobserved heterogeneity. *J. Marketing Res.* 28 (February), 29–41.

Walters, R.G., and S.B. MacKenzie (1988). A structural equation analysis of the impact of price promotions on store performance. *J. Marketing Res.* 25 (February), 51–63.

Wheat, R.D., and D.G. Morrison (1990a). Estimating purchase regularity with two interpurchase times. *J. Marketing Res.* 27 (February), 87–93.

Wheat, R.D., and D.G. Morrison (1990b). Assessing purchase timing models: Whether or not is preferable to when. *Marketing Sci.* 9(2), 162–170.

Wind, Y. (1978). Issues and advances in segmentation research. *J. Marketing Res.* 15, 317–337.

Winer, R.S. (1986). A reference price model of brand choice for frequently purchased products. *J. Consumer Res.* 13(2), 250–271.

Winer, R.S. (1989). A multi-stage model of choice incorporating reference prices. *Marketing Lett.* 1(1), 27–36.

Yellott, J.I. (1977). The relationship between Luce's choice axiom, Thurstone's theory of comparative judgments, and the double exponential distribution. *J. Math. Psychology*, 15, 109–144.

Zufryden, F.S. (1984). Modeling purchase patterns on the basis of incomplete and biased consumer purchase diary and UPC panel data. *Intern. Res. Marketing*, 1, 199–213.

Zufryden, F.S. (1987). A model for relating media exposure to purchase incidence behavior patterns. *Management Sci.* 33(10), 1253–1266.