Logically Consistent Market Share Models

PHILIPPE A. NAERT and ALAIN BULTEZ*

INTRODUCTION

A logically consistent market share model should predict market shares that are between zero and one, and sum to one. Few authors have worried about this type of problem in empirical studies, mainly because of the usual interest in a particular brand. It is then implicitly assumed that if predicted market share for that brand is MS_i, the other firms combined will get (1 − MS_i). This point may not be as obvious as it seems. If we were to estimate the market share response functions for the other brands as well, we would often find that the sum of the market shares is not one. The problem becomes apparent particularly when the response functions for all brands or for a group of brands are estimated simultaneously. An article by Neil E. Beckwith in a recent issue of JMR[2] reports an application of Zellner's joint generalized least squares method (joint GLS)[13] to the estimation of linear market share response functions of various competing brands. Beckwith's article illustrates an interesting way of obtaining more efficient estimators, but at the same time, raises the issue of logical consistency more clearly than in other applications. This article will address the problems which arise when market response models are sum-constrained. Reference will be made to Beckwith's study because it is one of the few examples where response functions of various brands have been estimated simultaneously.

We will first derive restrictions on the explanatory variables and on the parameters which are implied by a sum constraint on the dependent variable. Beckwith's study will serve as an illustration. The constraint in his work relates to the sum of the individual firms' market shares. Market share models are such that few distributions can describe the behavior of the disturbances. This is discussed in the second section. In the third section we reach the conclusion that for market share functions to be logically consistent their functional form should almost invariably be intrinsically nonlinear.1 It would seem intuitively reasonable to require logical consistency as a criterion for judging a model's appropriateness. However, logical consistency leads to more complicated market share functions, and necessitates more sophisticated estimation techniques. So there will be a trade-off between requiring logical consistency and model simplicity. This point will be examined in the final section.

IMPLICATIONS OF THE ADDITIVITY CONSTRAINT

Zellner's method will usually lead to considerable gains in efficiency of the estimators provided that disturbances of the different equations are contemporaneously highly correlated[13, pp. 353–4]. This condition is satisfied in the case of market share response functions for various brands competing in an oligopolistic market since, for example, a positive disturbance for one brand in a particular time period implies that at least one other brand will have a negative disturbance in that same period. This will be true especially in cases where the total market consists of just a few brands. Thus it would appear that Zellner's method may be profitably applied for joint estimation of a set of market share response functions.

However, Zellner's method does not guarantee that the sum of the market shares predicted from the estimated functions will add up to one, or to the sum of the market shares of the brands considered. Yet such a constraint is necessary if a logically consistent model is desired. For the time being we will assume that logical consistency is indeed desirable.

Dependent variables may be sum-constrained in the following sense: Suppose the dependent variable is market share. If all brands are considered, predicted market shares should add up to one, or to the sum of the market shares of the brands considered. Yet such a constraint is necessary if a logically consistent model is desired. For the time being we will assume that logical consistency is indeed desirable.

Dependent variables may be sum-constrained in the following sense: Suppose the dependent variable is market share. If all brands are considered, predicted market shares should add up to one. If one considers a limited number of brands which make up a well-defined submarket, then predicted market shares should add up to whatever share they actually sum to. For example, if

1 By intrinsically nonlinear forms we mean equations which cannot be linearized by transformations (e.g., logarithmic transformation), except by such approximating procedures as a first order Taylor expansion. In the sequel “linearization” will thus refer to transformation without approximation.
in period \( t \) these brands have \( r_t = .951 \) (95.1\% market share, and in \( t + 1 \), \( r_{t+1} = .953 \), then predicted shares should sum to 95.1 and 95.3 in period \( t \) and \( t + 1 \) respectively. Another example of a sum constraint which is time-dependent occurs when the dependent variable is brand sales. If \( q_{it} \) is sales of brand \( i \) in period \( t \), then the sum of brand sales over all brands should be equal to \( Q_t \), industry sales in period \( t \).

First we will examine the implications of imposing a sum constraint on the dependent variable in a set of linear functions, and then apply the results to the specific model proposed by Beckwith. Consider the following linear function (e.g., market share function for brand \( i \)):

\[
y_{it} = \beta_{i1} x_{1it} + \beta_{i2} x_{2it} + \cdots + \beta_{ip} x_{pit}, \quad \text{for } i = 1, \ldots, n.
\]

Let \( x_{j1} = (x_{1j1}, x_{2j2}, \ldots, x_{ijt}, \ldots, x_{nj}) \),

\[
y_j' = (y_{1j1}, y_{2j2}, \ldots, y_{nj}),
\]

\[
g_j' = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{ip}), \quad \text{for } j = 1, \ldots, p.
\]

Let \( j = 1 \) correspond to the constant term in the regression, that is, \( x_{1j} = 1 \) for \( i = 1, \ldots, n \) and \( t = 1, \ldots, T \). Finally let \( u \) be a \( nx1 \) column vector of ones, \( u = (1, 1, \ldots, 1) \).

**Theorem:** The necessary and sufficient conditions for a linear model, \( y_{it} = \sum_{j=1}^{p} \beta_{ij} x_{jit} \), to predict sum-constrained dependent variables, i.e., \( u'y_t = r_t \), are the following:

(a) the explanatory variables should be sum-constrained, i.e.,

\[
u'x_{jt} = c_{jt}, \quad \text{for all } j;
\]

(b) the coefficients, excluding the constant term, should be equal across equations, i.e.,

\[
\beta_{ij} = \beta_i, \quad \text{for all } i \text{ and all } j \neq 1;
\]

(c) and the following relationship between the \( \beta_i \)'s, \( c_{jt} \)'s, and \( r_t \) must hold:

\[
u'\beta_i + \sum_{j=1}^{p} \beta_j c_{jt} = r_t.
\]

The proof may be found in the Appendix.

Let us now apply the theorem to the market studied by Beckwith. He considers 5 brands which represent about 98\% of the market. We will call this total \( r_t \) in period \( t \). Two additional firms share the remaining 2\% [2, p. 169]. Each of these 2 have very specialized market segments. It is assumed that brands other than the 5 under study have no advertising expenditures [1, p. 56]. Beckwith formulates the following market share response function: market share, \( MS_{i,t} \), is a function of lagged market share \( MS_{i,t-1} \) and advertising share \( A_{Si,t} \):

\[
MS_{i,t} = \lambda_i MS_{i,t-1} + \gamma_i A_{Si,t} + \epsilon_{i,t}.
\]

The constraint on the dependent variable is,

\[
\sum_{i=1}^{n} MS_{i,t} = r_t.
\]

The explanatory variables are also sum-constrained,

\[
\sum_{i=1}^{n} MS_{i,t-1} = r_{t-1}, \quad \text{and } \sum_{i=1}^{n} A_{Si,t} = 1.
\]

Following the theorem, however, we should also have,

(a) from (3)

\[
\lambda_i = \lambda \quad \text{for all } i
\]

\[
\gamma_i = \gamma \quad \text{for all } i
\]

(b) from (4),

\[
r_t = \lambda r_{t-1} + \gamma.
\]

The conditions on \( \lambda \) and \( \gamma \) are not satisfied in Beckwith's article. For example, considering the IZEF estimates in his Table 2, the range of values for \( \lambda_i \) is 0.9814 to 1.0068, and for \( \gamma_i \), -0.0030 to 0.0133. Clearly, the various \( \lambda_i \) estimates are very similar, and with \( r_t \) approximately equal to .98, \( \lambda r_{t-1} + \gamma_i \) is close to \( r_t \) for each \( i \). The \( \gamma_i \) estimates show more variation, and therefore, given our assumption that logical consistency is a prerequisite for model acceptance, the proposed linear model has to be rejected.

McGuire, Farley, Lucas, and Ring proposed a generalized least squares procedure for estimating linear models in which the sum of the dependent variables is constrained [7]. They show that constraining the dependent variable implies restrictions on the explanatory variables as derived in the theorem above. However, they start from the assumption that for any \( j \), \( \beta_{ij} = \beta_j \) for all \( i \). Here we have shown that \( \beta_{ij} = \beta_j \) for all \( i \), and for \( j = 2, \ldots, p \) follows from the constraint on the dependent variables, and that no such restriction is needed for the constant term. Since Beckwith's formulation does not include a constant term, the procedure proposed by McGuire et al. is directly applicable. The procedure suggested by McGuire et al. as applied to Beckwith's case is briefly outlined below:

(a) Pool the data on the various brands into one big regression equation,

\[
MS_{i,t} = \lambda MS_{i,t-1} + (r_t - \lambda r_{t-1}) A_{Si,t} + \epsilon_{i,t},
\]

in which the constraints \( \lambda_i = \lambda, \gamma_i = \gamma \) for all \( i \), and \( \gamma = r_t - \lambda r_{t-1} \) are embodied. That is, we have to estimate the constrained form,


\[ MS_{i,t} - r_{i}A_{S_{i,t}} \]

(7)

Alternatively, we could have substituted

\[ \lambda = (r_{i} - \gamma)/r_{i-1}, \]

in (5). The equation to be estimated would then be,

\[ MS_{i,t} - (r_{i}/r_{i-1})MS_{i,t-1} \]

(8)

So if \( r_{i} \) is not constant, either \( \lambda \) or \( \gamma \) will be time dependent. This results from the peculiar assumption that 100% advertising leads to \( r_{i} < 1.0 \) share in a model where advertising is the only decision variable. It would be better in this case to regard the market captured by the 5 brands as the total market, i.e., express the market shares of the individual brands as a percentage of \( r_{i} \). Industry sales should then also be defined as \( r_{i} \) times true industry sales. This would seem quite appropriate since, as Beckwith states, the 2 other brands appeal to very specialized market segments.

In fact, we have no more information than that total share for the 5 brands is approximately 98% in each period; \( r_{i} \) is then, \( r_{i} = r_{i-1} = r = .98 \), for all \( t \), and (7) becomes,

\[ MS_{i,t} - .98 A_{S_{i,t}} = (MS_{i,t-1} - .98 A_{S_{i,t}}) + \epsilon_{i,t}, \]

while (8) reduces to,

\[ MS_{i,t} - MS_{i,t-1} = (\lambda (MS_{i,t-1} - .98 A_{S_{i,t}}) + \epsilon_{i,t} \]

(b) The second step is the ordinary least squares (OLS) estimation of the equation obtained in the first step. This will provide us with an estimate of the matrix of contemporaneous covariances of the various brands’ disturbances: \( \Omega \);

(c) then, we can estimate our constrained equation by GLS using \( \Phi = \Omega \otimes I \). Note, however, that \( \Omega \) is singular since if the market shares sum to \( r_{i} \),

\[ \sum_{i=1}^{n} \epsilon_{i,t} = 0 \]

(see [7, p. 1203]) as a result one of the brands observations ought to be deleted from the set, so that the method should be applied to \( n - 1 \) brands.

RESTRICTIONS ON THE DISTURBANCE TERMS

Market share is a quantity obviously confined to the interval from zero to one. This natural restriction limits the set of distributions capable of describing the behavior of the disturbances (see [7, p. 1205]) since \( 0 \leq MS_{i,t} \leq 1 \) implies:

\[ -1 \leq -(\lambda \cdot MS_{i,t-1} + \gamma \cdot A_{S_{i,t}}) \leq \epsilon_{i,t} \leq \]

\[ 1 - (\lambda \cdot MS_{i,t-1} + \gamma \cdot A_{S_{i,t}}) \leq 1 \]

which indicates that the normality assumption is irrelevant.

As pointed out by Theil [11, pp. 629 ff.], this aspect is awkward in regression-type situations and therefore a monotonic transformation is usually applied to this particular kind of dependent variable so that the new dependent variable constructed is then defined over the \([-\infty, +\infty]\) range. In fact, many transformations have this property. The logit is one of them, and in our case the first step amounts to defining a new variable:

\[ m_{i,t} = MS_{i,t}/(1 - MS_{i,t}) \]

a quantity with range \([0, \infty]\) since when \( MS_{i,t} = 0 \), \( m_{i,t} = 0 \) and when \( MS_{i,t} = 1 \), \( m_{i,t} = 1/0 = \infty \).

The log of \( m_{i,t} \) is then defined over the \([-\infty, +\infty]\) interval since \((\log 0) = -\infty \) and \((\log + \infty) = +\infty \). Adjusting the original model accordingly we obtain the equation to estimate. In Beckwith’s case we come up with:

\[ \log \left[ \frac{MS_{i,t}}{1 - MS_{i,t}} \right] = \log \left[ \frac{\lambda \cdot MS_{i,t-1} + \gamma \cdot A_{S_{i,t}}}{1 - (\lambda \cdot MS_{i,t-1} + \gamma \cdot A_{S_{i,t}})} \right] \]

or

\[ \log \left[ \frac{MS_{i,t}}{1 - MS_{i,t}} \right] = \log (\lambda \cdot MS_{i,t} + \gamma \cdot A_{S_{i,t}}) \]

\[ - \log (1 - \lambda \cdot MS_{i,t-1} - \gamma \cdot A_{S_{i,t}}) \]

which has now an intrinsically nonlinear form.

INTRINSICALLY NONLINEAR FORMS: A NECESSITY?

Now one might argue that the requirements \( \lambda_{i} = \lambda \) and \( \gamma_{i} = \gamma \) for all \( i \) are not very appealing. Indeed they do not allow for differentiation between brands, in the way market shares respond to advertising decisions. We then come to the conclusion that, if we want a logically consistent model which allows for differences in the response parameters between the various brands, the model structure should simply be nonlinear.

We may wonder at this point whether such a structure may easily be defined. Usually when market share functions are nonlinear, they are of such a form as to become linear upon transformation. For example, multiplicative functions are widely used in empirical work. The multiplicative equivalent of the expected value of Beckwith’s market share function is:

\[ MS_{i,t} = MS_{i,t-1} A_{S_{i,t}}. \]

For a multiplicative response function such as (9), it is not possible to determine meaningful restrictions on the explanatory variables and on the parameters, which will guarantee that market shares add up to one. More complex functions are generally needed.

For example, with advertising and lagged market

\[ 4 \]
share as the only explanatory variables, let:

\[ MS_{i,t} = \lambda MS_{i,t-1} + (1 - \lambda) \frac{A_{i,t}^\gamma}{\sum_j A_{j,t}^\gamma}, \]

where \( A_{i,t} \) is brand \( i \)'s advertising expenditures. It is readily seen that the sum of the market shares as defined in (10) is one, and this without specific constraints on the \( \gamma_i \) parameters. Unfortunately, such formulations have some problems of their own. Determining the values of the parameters is less straightforward than in the linear case, although many nonlinear programming procedures are now available which make nonlinear estimation much less of a problem. Also, the statistical properties of nonlinear estimators are weaker. This explains why empirical work has usually avoided intrinsically nonlinear functional forms. Only a few applications of nonlinear estimation techniques were reported in the marketing literature [4, 12].

Some readers may now ask themselves whether it is possible to design any logically consistent market share model which may be linearized. The answer is yes, but rather “heroic” assumptions have to be made and/or nontrivial transformations have to be devised. To illustrate, suppose the consumption pattern on a two-brand market may be described by a Markov-type matrix:

\[
\begin{bmatrix}
\lambda_i & \sigma_i \\
\sigma_c & \lambda_c
\end{bmatrix}
\]

where \( \lambda_i \) is the proportion of consumers loyal to brand \( i \); \( \sigma_c \) is the proportion of consumers switching from brand \( c \) to brand \( i \); \( \lambda_c \) and \( \sigma_i \) are similarly defined but refer to brand \( c \).

We can define the impact of the competitive forces on the consumption habits by making \( \lambda_i, \sigma_i, \lambda_c, \) and \( \sigma_c \) explicit functions of the 2 brands' marketing mix. Considering only the effect of advertising (A), we could postulate:

\[
\begin{align*}
\lambda_{i,t} &= 1 - \exp(-\alpha_i A_{i,t}^\gamma), \\
\sigma_c &= \exp(-\alpha_c A_{c,t}^\gamma), \\
\sigma_{i,t} &= \exp(-\alpha_i A_{i,t}^\gamma), \\
\lambda_{c,t} &= 1 - \exp(-\alpha_c A_{c,t}^\gamma)
\end{align*}
\]

where:

\( A_{i,t}^\gamma = A_{i,t}/A_{c,t} \quad \text{and} \quad A_{c,t}^\gamma = A_{c,t}/A_{i,t} \)

from which we deduce the nonlinear market share equation:

\[
E[MS_{i,t} | A_{i,t}, A_{c,t}] = [1 - \exp(-\alpha_i A_{i,t}^\gamma)] \cdot MS_{i,t-1} + \exp(-\alpha_c A_{c,t}^\gamma),
\]

and at equilibrium, assuming no change in the brands' advertising budgets until equilibrium is reached, then:

\[
E[MS_{i,t} | A_{i}, A_{c}] = \frac{\sigma_i}{\lambda_i + \sigma_i} = \frac{\exp(-\alpha_i A_{i}^\gamma)}{\exp(-\alpha_i A_{i}^\gamma) + \exp(-\alpha_c A_{c}^\gamma)}.
\]

If we are willing to assume that consumption patterns are sufficiently stable, which is very often the case when the primary demand reaches the maturity stage and when product innovations do not occur, and provided that the unit time period is sufficiently long, then the equilibrium may be achieved within period \( t \) and we may retain (13).\^4 Applying the logit transformation to \( t \) we obtain:

\[
\log \left( \frac{MS_{i,t}}{1 - MS_{i,t}} \right) = \log \left( \frac{\exp(-\alpha_i A_{i}^\gamma)}{\exp(-\alpha_c A_{c}^\gamma)} \right)
\]

\[
= \log [\exp(-\alpha_i A_{i}^\gamma)] - \log [\exp(-\alpha_c A_{c}^\gamma)],
\]

which means that we are now able to apply the OLS technique to:

\[
\log \left( \frac{MS_{i,t}}{1 - MS_{i,t}} \right) = \alpha_i A_{i}^\gamma - \alpha_c A_{c}^\gamma + \epsilon_{i,t}.
\]

The error term \( \epsilon_{i,t} \) is thus defined over the range \([ - \infty, + \infty ]\) and if we consider \( MS_{i,t} \) as a binomial frequency, then \( \epsilon_{i,t} \) is asymptotically normally distributed [11, p. 636].

It should be clear that the linearization of (13) is critically conditional upon the exponential specification chosen for \( \lambda_i, \sigma_i, \lambda_c, \) and \( \sigma_c \). It should also be observed that some quite restrictive assumptions have to be made in order to select (13) instead of the intrinsically nonlinear form (12).

Sometimes, it takes rather ingenious transformations to linearize models that look intrinsically nonlinear. Take, for example, the following market share function:

\[
MS_{i,t} = \frac{a_i \prod_{j=1}^{m} (X_{ij})^{\beta_j}}{\sum_{i=1}^{n} \left( a_i \prod_{j=1}^{m} (X_{ij})^{\beta_j} \right)},
\]

where, \( MS_{i,t}, X_{ij}, \) and \( \beta_j \) are defined as before, \( a_i \) is a coefficient which distinguishes the various brands.

Equation (15) is similar to that used by Urban [12].\^7

\^4 Provided of course that no change in the advertising strategy of both firms occurs during period \( t \). Similar assumptions were implicitly made by Hartung and Fisher [3]. We are presently investigating similar forms as well as nonlinear structures applied to the appraisal of distribution network performance [8].

\^6 An acceptable assumption since we consider the \( MS_{i,t} \) as conditional equilibrium market shares.

\^7 In Urban's formulation the \( a_i \) coefficients are not present, but brands differ in the other response coefficients. That is, the \( j \)th marketing instrument has a coefficient \( \beta_{ij} \) for brand \( i \), rather than \( \beta_j \).
He applied a direct search technique to estimate the parameters. Nakanishi [9] has shown that (15) can be made linear after applying the following transformation:

\[ \bar{a} = \left( \sum_{i=1}^{n} \log MS_{i,t} \right) / n \]

\[ \bar{X}_{it} = \left( \sum_{i=1}^{n} \log MS_{i,t} \right) / n. \]

The linearized form of (15) is then:

\[ (\log MS_{i,t} - \bar{MS}_{i}) = (\log a_{i} - \bar{a}) \]

\[ + \sum_{i=p}^{i=p} \beta_{i}(\log X_{ijt} - \bar{X}_{ij}). \]

This example illustrates that sometimes it is possible to linearize models by means of nontrivial transformations. In many cases, however, one will find that imposing logical consistency leads to response functions that are intrinsically nonlinear. In such cases, nonlinear procedures are to be applied to estimate the parameters.

**LOGICAL CONSISTENCY AS A CRITERION FOR JUDGING MODEL STRUCTURE**

Without additivity constraint, predicted market share summed over all brands might be more or less than one. Logical consistency, therefore, would seem intuitively appealing as one criterion for judging a particular market share model structure. In terms of Little's decision calculus [6] one would say that without additivity constraint a market share model lacks robustness. Yet, the advertising budgeting model that illustrates the concept of a decision calculus is itself not robust with regard to that constraint. Market share in Little's model is:

\[ MS_{i,t} = Min + (Max - Min) \]

\[ \cdot adv_{i,t} / (\delta + adv_{i,t}), \]

where:

- \( MS_{i,t} \) is market share of company \( i \) at time \( t \),
- \( adv_{i,t} \) is \( i \)'s advertising expenditure at \( t \),
- \( Min, Max, \gamma, \delta \) are parameters with \( 0 \leq Min \leq Max \leq 1 \), and \( \delta, \gamma \) positive.

If we were to define a similar market share function for competitors, no meaningful constraints on the parameters could be found to guarantee that the sum of predicted market shares be 1. The easy way out then is to assume (implicitly) that competitors' market share is \( 1 - MS_{i,t} \). However, that implies:

- Competitors' Share = \( 1 - Min \)

\[ -(Max - Min)adv_{i,t} / (\delta + adv_{i,t}), \]

that is, competitors' market share is a decreasing function of company \( i \)'s advertising expenditures, but does not depend on their own advertising. On the other hand, returning to (17), we see that for any value of \( i \)'s advertising, its market share will be between 0 and 1, and will be nondecreasing with advertising. So in many respects this model is robust. Suppose now for a moment that competitive advertising data are not available. Without such data, (17) perhaps represents the most robust possible functional form. Competitive activity is then implicitly reflected in the response parameters, and the value of additional information on competitive activity may not even be worth the cost. So, robustness is not something absolute, but has to be looked at from a cost-benefit point of view. And thus, robustness should also be related to what the model will be used for. For example, if the problem is to find out how competitors will react to changes in \( i \)'s advertising, (17) would be of little help. On the other hand if we want to know how much share to expect for a given advertising budget, (17) will probably perform quite well, assuming that the brand in question is well-established.

Returning now to the comparison between linear and nonlinear models, we do not argue then that most (intrinsically) linear models should no longer be used. We should recognize the fact that linear (or linearizable) models are easier and less expensive to estimate, the statistical results are believed to be more straightforward to interpret, and when used as aids in decision making, optimal allocation rules are simpler to derive and to apply. In a cost-benefit sense, predicted market share values from an intrinsically linear model may often provide sufficiently close approximations.\(^8\)

Furthermore, estimated parameters are often used as prior estimates, which are subsequently adjusted by managerial judgment. Defining an approximate model may then perhaps be even more acceptable. For a discussion of some of these issues, see Lambin [5, pp. 120-1].

**CONCLUSION**

In this article we derived restrictions on explanatory variables, parameters, and disturbances implied by an additivity constraint on the dependent variable, e.g., market share. These restrictions are such that linear market share structures do not allow for differences in the response parameters for various brands. We argued that to be logically consistent, market share models will often be intrinsically nonlinear. We examined logical consistency as a criterion for judging model structure and argued that this should be looked at from a cost-benefit point of view.

**APPENDIX**

Theorem: The necessary and sufficient conditions for a linear model,\(^8\)

\(^8\) The authors are currently investigating the cost-benefit aspects in more depth based on empirical studies as well as by means of simulation.
to predict sum-constrained dependent variables, i.e.,
\[ u'y_t = r_t, \]
are:
(a) \( u'X_{jt} = c_{jt}, \) for all \( j \) and \( t, \)
(b) \( \beta_{ij} = \beta_i^*, \) for all \( i \) and all \( j \neq 1, \)
(c) \( u'\beta_i + \sum_{j=2}^{p} \beta_j c_{jt} = r_t, \) for all \( t. \)

**Sufficient Condition:** If \( u'X_{jt} = c_{jt} \) for all \( j \) and \( t, \)
and \( \beta_{ij} = \beta_i^* \) for all \( i \) and for \( j = 2, \cdots, p, \) and if \( u'\beta_i + \sum_{j=2}^{p} \beta_j c_{jt} = r_t, \) it follows that \( u'y_t = r_t. \)

*Proof:* First we take the sum of the dependent variables
\[ (19) \quad u'y_t = \sum_{j=1}^{p} X_{jt} \beta_j. \]

\( X_{jt} \) is a vector of ones, and thus \( X_{jt} \beta_i = u'\beta_i. \) For
all \( j \) not equal to one \( \beta_j \) is a vector of constants \( \beta^i, \) i.e.,
\( \beta_j = (\beta_1, \beta_1, \cdots, \beta_i), \) and therefore we can write:
\[ X_{jt} \beta_j = \beta_i u'X_{jt}. \]

With \( u'X_{jt} = c_{jt}, \) we obtain after substitution in (19):
\[ (20) \quad u'y_t = u'\beta_i + \sum_{j=2}^{p} \beta_j c_{jt}. \]

With the right-hand side of (20) equal to \( r_t, \) the proof of sufficiency is complete.

**Necessary Condition:** Suppose now that the dependent variables are sum-constrained, i.e., \( u'y_t = r_t. \) To be shown is that \( u'y_t = r_t \) implies \( u'X_{jt} = c_{jt} \) for all \( j, \) \( \beta_{ij} = \beta_i^* \) for all \( i \) and for \( j = 2, \cdots, p, \) and \( u'\beta_i + \sum_{j=2}^{p} \beta_j c_{jt} = r_t. \)

*Proof:* (1) \( u'X_{jt} = c_{jt} \) for all \( j. \)

Consider two sets of admissible values for the independent variables \( X_{jt} \) and \( X^*_{jt}, \) that is,
\[ (21) \quad r_t = \sum_{j=1}^{p} X_{jt} \beta_j, \quad \text{and} \]
\[ (22) \quad r_t = \sum_{j=1}^{p} X^*_{jt} \beta_j. \]

Subtracting (22) from (21), we obtain:
\[ (23) \quad 0 = \sum_{j=1}^{p} (X_{jt} - X^*_{jt}) \beta_j. \]

Without loss of generality, we can consider the following:
\( X^*_{jt} \neq X^*_{jt} \) for \( j = k \neq 1, \)
\( X^*_{jt} = X^*_{jt} \) for \( j \neq k. \)

Equation (23) then simplifies to:
\[ (24) \quad 0 = \sum_{j=1}^{n} \beta_{ik} (X^*_{ikt} - X^*_{ikt}). \]

Without loss of generality we can assume \( \beta_{ik} \neq 0 \) only for \( i = x. \)

Suppose now that \( X_{jt} \) is not sum-constrained. The following could then be admissible \( X_{kt} \) vectors:
\[ X^*_{kt} = X^*_{ikt} \text{ for } i \neq s, \]
\[ X^*_{kt} \neq X^*_{ikt} \text{ for } i = s. \]

Equation (24) now becomes:
\[ 0 = \beta_{ik} (X^*_{ikt} - X^*_{ikt}). \]

Since \( \beta_{ik} \neq 0, \) and \( X^*_{ikt} \neq X^*_{ikt}, \) this is a contradiction and hence:
\[ u'X_{jt} = c_{jt} \quad \text{for } j = 2, \cdots, p. \]

Note that for \( j = 1, \) \( u'X_{jt} \) is by definition constant and equal to \( n. \)

*Proof:* (2) \( \beta_{ij} = \beta_i^* \) for \( j = 2, \cdots, p. \)

Consider the following vectors for the explanatory variables:
\( X^*_{jt} \neq X^*_{jt} \) for \( j = k \neq 1, \)
\( X^*_{jt} = X^*_{jt} \) for \( j \neq k. \)

Assume that the \( X^*_{jt} \) vectors satisfy the sum constraint conditions derived in part (1). Let \( X^*_{jt}, \) be defined as follows:
\[ X^*_{ikt} = X^*_{ikt} + \Delta, \quad \text{with } \Delta \neq 0, \]
\[ X^*_{ikt} = X^*_{ikt} - \Delta. \]

Thus, the \( X^*_{jt} \) vectors also satisfy the sum constraint conditions. Now assume that:
\[ \beta_{ik} \neq \beta_{ik}. \]

Substituting (25) into (24) we find:
\[ 0 = \Delta (\beta_{ik} - \beta_{ik}). \]

With \( \Delta \neq 0, \) assuming \( \beta_{ik} \neq \beta_{ik} \) results in a contradiction. Thus
\[ \beta_{ij} = \beta_i^* \text{ for } j = 2, \cdots, p. \]

*Proof:* (3) \( u'\beta_i + \sum_{j=2}^{p} \beta_j c_{jt} = r_t. \)

Substituting the results of parts (1) and (2) into equation (19), we obtain:
\[ u'y_t = u'\beta_i + \sum_{j=2}^{p} \beta_j c_{jt}. \]

Given \( u'y_t = r_t, \) the proof is complete.

*If \( \beta_{ik} \) were 0 for all \( i \) and \( k, \) none of the \( X's \) would be explanatory variables. If at least one of the \( X's \) is a true explanatory variable, \( s \) and \( k \) can always be chosen such that \( \beta_{ik} \neq 0. \)
REFERENCES

1. Beckwith, Neil E. “The Response of Competing Brands to Advertising: A Multivariate Regression Test,” unpublished doctoral dissertation, Purdue University, 1970.

2. ———. “Multivariate Analysis of Sales Responses of Competing Brands to Advertising,” Journal of Marketing Research, 9 (May 1972), 168–76.

3. Hartung, Philip H. and James L. Fisher. “Brand Switching and Mathematical Programming in Market Expansion,” Management Science: Application, 2 (August 1965), 231–43.

4. Kuehn, Alfred A., Timothy W. McGuire, and Doyle L. Weiss. “Measuring the Effectiveness of Advertising,” in Science Technology and Marketing, American Marketing Association, 1966, 185–94; also in John J. Wheatley, ed., Measuring Advertising Effectiveness. Homewood, Ill.: Richard D. Irwin, 1969, 30–9.

5. Lambin, Jean-Jacques. “A Computer On-Line Marketing Mix Model,” Journal of Marketing Research, 9 (May 1972), 119–26.

6. Little, John D. C. “Models and Managers: The Concept of a Decision Calculus,” Management Science, 16 (April 1970), B466–85.

7. McGuire, Timothy W., John U. Farley, Robert E. Lucas, Jr., and Winston L. Ring. “Estimation and Inference for Linear Models in Which Subsets of the Dependent Variable are Constrained,” Journal of the American Statistical Association, 63 (December 1968), 1201–13.

8. Naert, Philippe A. and Alain Bultez. “An Aggregate Retail Outlet Location Model,” unpublished working paper, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, 1972.

9. Nakanishi, Masao. “Measurement of Sales Promotion Effect at the Retail Level: A New Approach,” unpublished working paper, Graduate School of Management, University of California, Los Angeles, 1972.

10. Schmalensee, Richard Lee. “On the Economics of Advertising,” unpublished doctoral dissertation, Massachusetts Institute of Technology, 1970.

11. Theil, Henri. Principles of Econometrics. New York: John Wiley & Sons, 1971.

12. Urban, Glen. “A Mathematical Modeling Approach to Product Line Decisions,” Journal of Marketing Research, 6 (February 1969), 40–7.

13. Zellner, Arnold. “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias,” Journal of the American Statistical Association, 57 (June 1962), 348–68.