

Logical Consistency and Sum-Constrained Linear Models

Frederieke van Perlo-ten Kleij Ton Steerneman* Ruud H. Koning[†]

September 25, 2006

ABSTRACT

A topic that has received quite some attention in the seventies and eighties is logical consistency of sum-constrained linear models. Loosely defined, a sum-constrained model is logically consistent if the restrictions on the parameters and explanatory variables are such that the sum constraint is automatically satisfied. The literature on logical consistency, however, has not been unambiguous. The main reason for this is that a rigorous definition of logical consistency has not been given in the (marketing) literature. Inspired by an extensive discussion by Koehler and Wildt (1981), we therefore present two closely related definitions of logical consistency. By using an elegant and direct approach, we derive necessary and sufficient conditions for the sum-constrained linear model to be logically consistent. We summarize the results on this topic up until now, and we clarify some differences and obscurities. We also show that some generally accepted results of McGuire and Weiss (1976) are not correct.

Keywords: Restrictions on parameters and regressors in linear models with sum-constrained dependent variables.

*This author has passed away between the first version of this paper and the final manuscript.

[†]Corresponding author. Address: Department of Econometrics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands. Email: r.h.koning@rug.nl

1 INTRODUCTION

Sum-constrained linear models are linear models in which the dependent variables add up to a fixed number. Such models naturally occur in various fields of research. In demand analysis the amounts spent on the categories of consumer goods and services that are distinguished add up to total expenditure, in production theory the cost shares of the various factors of production add up to unity, in marketing analysis the probabilities that a specific brand will be chosen add up to unity, in international trade the flows of exports from a specific country to different destinations add up to total exports, and so on (De Boer and Harkema, 1984). In their 1968 article, McGuire (1968) et al. examined the restrictions on the parameters and explanatory variables implied by a sum-constrained model, such that the sum constraint is automatically satisfied. This article formed the basis for the concept ‘logical consistency’, a term that originates from the marketing literature on sum-constrained linear models. The idea of logical consistency is that parameters and regressors in a model with a sum-constrained dependent variable cannot move freely, but are in some way restricted, such that the sum constraint is automatically satisfied.

According to Koehler and Wildt (1981), this particular relationship between constraints on variables on the one hand and constraints on model parameters on the other hand had been largely ignored. This was apparent from the fact that a comprehensive examination and rigorous definition of logical consistency had not been provided in the marketing literature. In order to give a thorough discussion on the concept of logical consistency, we start in section 2 by giving two (closely related) definitions of logical consistency. These definitions are a mixture between some of the definitions of Koehler and Wildt. We elaborate on the differences in interpretation of these two definitions and we illustrate these differences by means of some examples. In section 3, necessary and sufficient conditions will be given for the sum-constrained linear model to be ‘logically consistent’. These results are based on an underexposed article by Weverbergh, Naert and Bultez (1981). Although the results of Weverbergh, Naert and Bultez encompassed all the preceding results on logical consistency, these authors only considered the case where the number of observations exceeds the number of explanatory variables plus one, an assumption which may not be realistic in practical applications. Besides making the definitions of logical consistency operational in an elegant and direct way, our main contribution in this section is to generalize the results for an arbitrary number of observations. Section 4 puts some well-known results of, for example, Naert and Bultez (1973) and McGuire and Weiss (1976) in perspective, by showing how they relate to the results of section 2 and 3. In section 5 we provide several examples to illustrate the usefulness of our approach. One of the examples deals with a case not covered by the generally accepted results of McGuire and Weiss.

The main purpose of this paper is to give two (related) definitions of logical consistency of a sum-constrained linear model and to show how these definitions lead to empirically useful restrictions on model variables and parameters. As a by-product, we also correct the mistakes which have been made concerning logical consistency. Our message is not to encourage practitioners to start using the linear model in case the dependent variable is sum-constrained. We are well aware of the fact that there are more appropriate models for this, such as the class of attraction models (Nakanishi and Cooper 1974; Bell, Keeney and Little 1975; Bultez and Naert 1975). We think, however, that the main results on logical consistency by McGuire and Weiss (1976), which have been often referred to in the past, need to be corrected. The results of McGuire and Weiss (1976), as well as those of Naert and Bultez (1973), have often been used in empirical applications, within marketing (Naert and Weverbergh 1981; Leeftang and Reuyl 1984; Ghosh, Neslin and Schoemaker 1984), but also outside marketing (De Boer, Harkema and Soede 1996; De Boer and Martinez 1999).

2 TWO DEFINITIONS OF LOGICAL CONSISTENCY

Consider the general sum-constrained linear model, defined by the following system of equations

$$y_{ti} = \alpha_i + \sum_{j=1}^{k_i} z_{tij} \beta_{ij} + \varepsilon_{ti} \quad (t = 1, \dots, T; i = 1, \dots, n) \quad (1)$$

subject to

$$\sum_{i=1}^n y_{ti} = m_t \quad (t = 1, \dots, T), \quad (2)$$

where y_{ti} denotes the t th observation on the dependent variable in the i th category, α_i is an intercept term, z_{tij} represents the t th observation on the j th element of a set of k_i explanatory variables which are supposed to be specific for the i th dependent variable, β_{ij} is an unknown parameter to be estimated, ε_{ti} represents a zero-mean disturbance, and n denotes the number of categories distinguished. We define $y_t = (y_{t1}, \dots, y_{tn})'$, $z_{ti} = (z_{ti1}, \dots, z_{tik_i})'$, $\varepsilon_t = (\varepsilon_{t1}, \dots, \varepsilon_{tn})'$, $\alpha = (\alpha_1, \dots, \alpha_n)'$, $\beta_i = (\beta_{i1}, \dots, \beta_{ik_i})'$, and $\beta = (\beta'_1, \dots, \beta'_n)'$. Then this model can also be written as

$$y_t = \alpha + X_t \beta + \varepsilon_t \quad (3)$$

subject to

$$l'_n y_t = m_t, \quad (4)$$

where $[X_t = \begin{pmatrix} z'_{t1} & 0 & \dots & 0 \\ 0 & z'_{t2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & z'_{tn} \end{pmatrix}]$ so that X_t is $n \times k$ with $k = \sum_{i=1}^n k_i$. We assume that the dis-

turbance terms ε_t are independently and identically distributed as $\mathcal{N}_n(0, \Omega)$. By summing (3) over i and taking expectations, it follows that

$$l'_n(\alpha + X_t \beta) = m_t \quad \text{for } t = 1, \dots, T, \quad (5)$$

and therefore

$$l'_n \varepsilon_t = 0. \quad (6)$$

Two important observations are the following.

- (i) Because the ε_t are i.i.d. $\mathcal{N}_n(0, \Omega)$, it follows from equation (6) that the covariance matrix Ω of the disturbances must be sum-constrained, that is, $\Omega l_n = 0$.
- (ii) The regressors X_t and the parameters α and β must obviously be restricted in some way in order to satisfy equation (5) for all t .

Observation (i) implies that the covariance matrix is singular, which means that the vectors of disturbances are linearly dependent. When estimating the model, this needs to be taken into account, see, for example, Naert and Weverbergh (1985), Gaver, Horsky and Narasimhan (1988) and De Boer and Harkema (1997). Observation (ii) is related to what has been referred to as ‘logical consistency’ in the literature. In our definition of logical consistency, we are only concerned with the restrictions on the explanatory variables and the parameters which are necessary to guarantee that the sum constraint (4) is automatically satisfied. In market share models, for example, equation (4) must hold with $m_t = 1$ for all t . Besides these equality restrictions, the market shares y_{ti} are also between zero and one. We will not take such inequality restrictions into account, and in the context of market share models, we will only focus on the condition that market shares must sum to one. Because of (6), we therefore do not further consider the disturbance term.

It is also desirable that equation (4) holds for the estimated and the predicted values of the dependent variable. In fact, this is how logical consistency has been implicitly defined most often in the marketing literature (see, e.g., Weverbergh, Naert and Bultez 1981 and Naert and Bultez 1973). It is not clear from the literature whether the estimated or the predicted values of the dependent variable are to be sum-constrained. This difference in interpretation (estimated versus predicted) is one of the reasons why we introduce two different definitions of logical consistency.

DEFINITION 2.1 (LOGICAL DATA-CONSISTENCY) *The sum-constrained model defined by (3) and (4) is logically data-consistent if, for a given set of values of the explanatory variables $X_t, t = 1, \dots, T$, the parameters α and β are such that the sum constraint (4) is automatically satisfied.*

DEFINITION 2.2 (LOGICAL PREDICTOR-CONSISTENCY) *The sum-constrained model defined by (3) and (4) is logically predictor-consistent if the parameters α and β and all possible values of the explanatory variables X_t are such that the sum constraint (4) is automatically satisfied.*

If a model is logically predictor-consistent, it is also logically data-consistent. The converse is not always true. The main difference between the two definitions is that in definition 2.1, we only consider the observed regressors X_t contained in the data used to estimate the model, rather than the larger set of all admissible values of X_t . Definition 2.2 allows for unobserved or future information $\{X_T, X_{T+1}, \dots\}$, whereas definition 2.1 only takes observed information into account. If the only one objective of the sum-constrained model is to explain the linear relationship between the regressors and the dependent variable for a given set of observations, then definition 2.1 can be used. In practice, however, a model is often used to make predictions, so that the model should also hold for future observations. Indeed, otherwise extrapolation would not make any sense. In that case, definition 2.2 should be used. In our opinion, definition 2.2 is therefore also the most relevant of the two. An exception is the case where the matrix of regressors is fixed, for example in designed experiments or in replication studies. In this case, the matrix of regressors is constant over time, and there is no difference in adopting definition 2.1 or 2.2.

Definition 2.2 implies that the restrictions on α and β are not allowed to depend on T , an idea which was already suggested by McGuire et al. (1968). This also means that if the model holds for $t = 1, \dots, T$, then the parameters are not allowed to change if we add another observation $t = T + 1$. The significance of this assumption comes to the fore, for example, in prediction or replication studies. According to definition 2.1, the restrictions on α and β may depend on T . The sum constraint (4) must be satisfied for the observed data X_t , replication or prediction is not considered relevant in this entirely data-driven definition of logical consistency.

In section 3, we come back to this difference in interpretation of logical consistency. If $T > k + 1$, we will show that the two definitions will often not lead to different conditions for logical consistency. If $T < k + 1$, however, there can be a difference, a detail that has been neglected in the literature on logical consistency.

3 NECESSARY AND SUFFICIENT CONDITIONS

In this section, we derive necessary and sufficient conditions for the sum-constrained linear model to be logically data-consistent, by using a similar approach as Weverbergh, Naert and Bultez (1981). The work of these authors was in turn based on earlier work of Weverbergh (1976). They explicitly assumed that $T > k + 1$. We show that the main results still hold if $T \leq k + 1$. We also derive necessary and sufficient conditions for the sum-constrained linear model to be logically predictor-consistent, a somewhat more difficult situation. In practice, the two definitions often lead to the same conditions if $T > k + 1$. If $T \leq k + 1$, some subtleties emerge, and we show that the necessary and sufficient conditions for a model to be logically data-consistent are not sufficient for the model to be logically predictor-consistent.

In vector notation, the sum-constrained linear model defined by (1) and (2) is represented by the following system of linear equations

$$y_i = \alpha_i \iota_T + Z_i \beta_i + \varepsilon_i \quad (i = 1, \dots, n) \quad (7)$$

and

$$\sum_{i=1}^n y_i = m, \quad (8)$$

where y_i , ε_i and m denote $T \times 1$ vectors with elements y_{ti} , ε_{ti} , and m_t respectively, Z_i denotes a $T \times k_i$ matrix of observations on a set of k_i explanatory variables which are specific for the dependent variable in the i th category, and β_i denotes a $k_i \times 1$ vector of parameters β_{ij} . Although the case $m_t = 1$ for all t is the most important case encountered in practice, the case of varying m_t is also interesting, for example, in macro-economic systems, where consumption plus savings must be equal to income, which varies over time (Weverbergh, Naert and Bultez 1981). According to Weverbergh (1976), the case of constant m_t can be easily generalized to the case of variable m_t by normalizing the equations. That is, divide (1) by m_t , so that the constant term is replaced by the regressor $1/m_t$ with parameter α_i , and the dependent variables

sum to one again. However, this normalization causes heterogeneity of the disturbance term which in turn causes other difficulties. We therefore start with the most general case of non-constant m_t and we will show how the results simplify in case the m_t are constant.

As already discussed in section 2, by summing equation (8) over i and taking expectations, it follows that

$$\sum_{i=1}^n (\alpha_i \iota_T + Z_i \beta_i) = m \quad (9)$$

and

$$\sum_{i=1}^n \varepsilon_i = 0.$$

Here, we assumed that the regressors are nonstochastic, or at least independent of the disturbances. If we define the $T \times (k + 1)$ matrix Z by

$$Z = (Z_1 \quad Z_2 \quad \cdots \quad Z_n \quad \iota_T) \quad (10)$$

and the $(k + 1) \times 1$ vector $\beta^* = (\beta'_1, \dots, \beta'_n, \iota'_n \alpha)'$, where $\alpha = (\alpha_1, \dots, \alpha_n)'$, then (9) can also be written as

$$Z\beta^* = m. \quad (11)$$

Equation (11) is the basis for all our results on logical consistency. We can interpret (11) in two ways. Given a set of regressors Z , the vector of parameters β^* must satisfy particular constraints in order to satisfy (11). On the other hand, given a set of parameters, the regressors Z must be constrained to make sure that equation (11) holds. In practice, the data will often be given, and the question is how the parameters should be constrained such that (11) will be satisfied. However, in the data collection process, it is probably taken into account that particular attention has to be paid to the choice of independent variables. In the context of ‘logical consistency’, we are looking for (necessary and sufficient) conditions on both the regressors and the parameters simultaneously to guarantee that the dependent variables are sum-constrained. The interaction between the number of constraints on regressors and parameters is apparent from equation (11). If there are many restrictions on the parameters, then the regressors have somewhat more freedom to move in order to satisfy (11), and vice versa. In the following, we will make a distinction between logical data-consistency and logical predictor-consistency, as defined in section 2.

3.1 LOGICAL DATA-CONSISTENCY

Let $Z = (Z_1, \dots, Z_T)'$, where the $(k + 1) \times 1$ vectors $Z_t, t = 1, \dots, T$ denote the rows of the matrix Z . Note the slight abuse of notation with respect to Z_i and Z_t , where the former denotes the $T \times k_i$ matrix of regressors and the latter denotes the t th row of Z . We assume that the index (category versus time) clarifies the meaning of the notation. According to equation (11), logical data-consistency requires $Z'_t \beta^* = m_t$ for $t = 1, \dots, T$.

If the rank of Z equals $k + 1$, then β^* is uniquely determined, so that equation (11) has a nontrivial solution if and only if $\text{rank}(Z) < k + 1$. Moreover, equation (11) has a solution if and only if

$$\text{rank}(Z) = \text{rank}(Z, m), \quad (12)$$

or, in other words, m must be part of the column space of Z . Summarizing, we proved the following theorem.

THEOREM 3.1 *Necessary and sufficient conditions relating constraints on parameters and on explanatory variables, such that the sum-constrained linear model defined by (7) and (8) is logically data-consistent, are given by*

$$Z\beta^* = m, \quad (13)$$

and

$$\text{rank}(Z) = \text{rank}(Z, m), \quad (14)$$

with Z as defined in (10), and $\beta^* = (\beta'_1, \dots, \beta'_n, \iota'_n \alpha)'$.

If $T > k + 1$, then the rank of Z is smaller than or equal to the number of columns of Z . Equation (13) shows that in this case, the rank of Z must be smaller than $k + 1$, for if Z is of full (column) rank, that is $\text{rank}(Z) = k + 1$, then β^* is uniquely determined according to theorem 3.1. Therefore, for the sum-constrained model to be meaningful in this case, regressors should be dependent, for example because they are sum-constrained across equations (e.g., advertising shares which appear in market share models as explanatory variables), or because they are the same across equations (so-called homogeneous regressors to be described in section 4). It is important to note that theorem 3.1 does not exclude sum-constrained or homogeneous regressors on beforehand. This is in contrast to the results of McGuire and Weiss (1976), who claimed that regressors must be either sum-constrained or homogeneous. In section 5, we will give an example of a logically (data- or predictor-) consistent sum-constrained model with a regressor that is neither homogeneous, nor sum-constrained.

In most applications, $m = \iota_T$, that is, the dependent variables sum to one. The following corollary deals with the somewhat more general case where $m = c\iota_T$.

COROLLARY 3.1 *Necessary and sufficient conditions relating constraints on parameters and on explanatory variables, such that the sum-constrained linear model defined by (7) and (8), with $m = c\iota_T$, is logically data-consistent, are given by*

$$\beta^* = \Lambda\lambda, \quad (15)$$

where $\beta^* = (\beta'_1, \dots, \beta'_n, \iota'_n\alpha - c)'$, Λ is an $(k + 1) \times v$ matrix, whose columns form a basis for the null space of Z as defined in (10), and where λ is a $v \times 1$ vector of proportionality factors.

With $\beta^* = (\beta'_1, \dots, \beta'_n, \iota'_n\alpha - c)'$, equation (9) can also be written as

$$Z\beta^* = 0. \quad (16)$$

Therefore, the vector β^* must be a solution to the system of homogeneous linear equations

$$Za = 0, \quad (17)$$

where a is a vector of dimension $(k + 1) \times 1$. If $\beta^* = \Lambda\lambda$ is a linear combination of vectors satisfying (17), then β^* also satisfies (16), which proves the sufficiency of the conditions. To prove necessity, suppose there exists a solution β^* . Because β^* is a solution of (16), it must belong to the null space of Z . Therefore, β^* can be written as a linear combination of the columns of Λ .

Corollary 3.1 resembles the main result of Weverbergh, Naert and Bultez (1981). Whereas these authors stated that this result only holds if $T > k + 1$, we emphasize that corollary 3.1 applies both if $T > k + 1$ and $T \leq k + 1$.

The following corollary shows that we can derive the number of parameter restrictions from the rank of the matrix Z . A proof of this corollary can be found in Weverbergh, Naert and Bultez (1981), our version of this proof is somewhat easier.

COROLLARY 3.2 *Let $p = \text{rank}(Z)$, with Z as defined in (10). Logical data-consistency of the model defined by (7) and (8) implies that there is one sum restriction on the constant terms, and that there are $p - 1$ restrictions on the k parameters β_{ij} .*

PROOF Let $T : V \rightarrow W$ be a linear transformation, where V and W are vector spaces, and let $\mathcal{K}(T)$ denote the kernel or the null space of T , and $\mathcal{R}(T)$ the range or image of T . If V is finite-dimensional, then it is well known that

$$\dim\mathcal{K}(T) + \dim\mathcal{R}(T) = \dim V,$$

see, e.g., Friedberg, Insel and Spence (1992). Applied to corollary 3.1, it follows that for the linear transformation $Z : \mathbb{R}^{k+1} \rightarrow \mathbb{R}^T$, the equality

$$v + \text{rank}(Z) = k + 1 \quad (18)$$

must hold. If we define $p = \text{rank}(Z)$, then it follows from corollary 3.1 that there is one sum restriction on the constant terms, and that there are $p - 1$ restrictions on the k parameters β_{ij} . The number of degrees of freedom on the vector β^* is therefore equal to v (e.g., Weverbergh 1976).

Observe that corollary 3.2 once more reflects the balance between the number of restrictions on the regressors (the rank of Z) and the parameters.

3.2 LOGICAL PREDICTOR-CONSISTENCY

We showed above that a model is logically data-consistent if and only if the T restrictions $Z_t' \beta^* = m_t$ are satisfied for $t = 1, \dots, T$. This does not guarantee, however, that $Z_{T+1}' \beta^* = m_{T+1}$ will also hold if we add an observation for $t = T + 1$ (irrespective of $T \leq k + 1$ or $T > k + 1$). Therefore, the model need not be logically predictor-consistent. However, we can derive a result analogous to theorem 3.1. Define

$$Z = \begin{pmatrix} Z_{\text{obs}} \\ Z_{\text{pred}} \end{pmatrix}, \quad (19)$$

where the $T \times (k + 1)$ matrix $Z_{\text{obs}} = (Z_1, \dots, Z_T)'$ contains the observed data, and the $s \times (k + 1)$ matrix $Z_{\text{pred}} = (Z_{T+1}, \dots, Z_{T+s})'$ contains the unobserved data, used for prediction. Moreover, Z_{obs} and Z_{pred} are defined in a similar way as in (10). The number of unobserved observations s can be arbitrarily large, the unobserved part of the matrix of regressors Z_{pred} is supposed to capture all possible values of the regressors. Similarly, define

$$m = \begin{pmatrix} m_{\text{obs}} \\ m_{\text{pred}} \end{pmatrix}. \quad (20)$$

For logical predictor-consistency, equation (11) must hold for both Z_{obs} (so that the model is logically data-consistent) and for Z_{pred} , so that the sum constraint is also satisfied for predicted values of the dependent variable. Moreover, the restrictions on β^* which guarantee that $Z_t' \beta^* = m_t$ for the observed data $t = 1, \dots, T$ are not allowed to change if we add the restrictions $Z_t' \beta^* = m_t$ for $t = T + 1, \dots, T + s$ for the unobserved data. This means that the rows Z_{T+1}, \dots, Z_{T+s} must be linear combinations of the rows Z_1, \dots, Z_T . We have now proven the following theorem.

THEOREM 3.2 *Necessary and sufficient conditions relating constraints on parameters and on explanatory variables, such that the sum-constrained linear model defined by (7) and (8) is logically predictor-consistent, are given by*

$$\begin{pmatrix} Z_{\text{obs}} \\ Z_{\text{pred}} \end{pmatrix} \beta^* = \begin{pmatrix} m_{\text{obs}} \\ m_{\text{pred}} \end{pmatrix}, \quad (21)$$

and

$$\text{rank}(Z) = \text{rank}(Z_{\text{obs}}) = \text{rank}(Z_{\text{obs}}, m_{\text{obs}}) = \text{rank}(Z, m), \quad (22)$$

with Z and m as defined in (19), resp. (20), and $\beta^* = (\beta_1', \dots, \beta_n', \iota_n' \alpha)'$.

The discussion whether $T \leq k + 1$ or $T > k + 1$ is not relevant in this context, because the matrix Z does not only capture the values of the observed regressors, but also all possible future observations. Therefore, we can always assume that the number of rows of Z exceeds the number of columns.

An interesting question is when the conditions implied by theorem 3.2 differ from those implied by theorem 3.1. As long as values of the unobserved regressors Z_{T+1}, \dots, Z_{T+s} and the constants m_{T+1}, \dots, m_{T+s} are a linear combination of the observations Z_1, \dots, Z_T resp. the constant terms m_1, \dots, m_T , the restrictions on the parameters implied by theorem 3.2 will be the same as the conditions that follow from theorem 3.1. If $T > k + 1$, we will show in section 5 that in most situations adding a row to the matrix Z will not change its rank. In this case, logical data-consistency and logical predictor-consistency lead to the same conditions. Note, however, that even if $T > k + 1$, it is possible that adding a row to the matrix Z changes its rank. In case $T < k + 1$, it becomes more likely that adding an observation to the matrix Z leads to an additional restriction on the parameters, so that theorem 3.1 will give conditions different from those of theorem 3.2.

In case $m = c \iota_T$, we can derive a result analogous to corollary 3.1. The proof of corollary 3.3 is similar to the proof of corollary 3.1.

COROLLARY 3.3 *Necessary and sufficient conditions relating constraints on parameters and on explanatory variables, such that the sum-constrained linear model defined by (7) and (8), with $m = c \iota_T$, is logically predictor-consistent, are given by*

$$\beta^* = \Lambda \lambda, \quad (23)$$

where $\beta^* = (\beta_1', \dots, \beta_n', \iota_n' \alpha - c)'$, Λ is an $(k + 1) \times v$ matrix, whose columns form a basis for the null space of Z as defined in (19), and where λ is a $v \times 1$ vector of proportionality factors.

As long as the values of the unobserved regressors Z_{T+1}, \dots, Z_{T+s} are a linear combination of the observations Z_1, \dots, Z_T , corollary 3.3 leads to the same conditions on the parameters as corollary 3.1. Or, stated differently, as long as the rank of Z as defined in (10) does not change if we add additional observations. In this case, the null space of Z will remain the same, so that β^* as defined in (23) will not change either.

In many applications of sum-constrained models, the null space of the matrix Z is apparent, and corollary 3.1 or corollary 3.3 can be used to find the conditions on the parameters to ensure logical data-consistency and logical predictor-consistency respectively. In section 5 we give a few examples to demonstrate how to exploit these corollaries. The corollaries are particularly illustrative if $T > k + 1$, since in this case, we can explicitly derive the conditions on β^* . For the case $T \leq k + 1$, we discuss the difficulties, and the distinction between the two definitions of logical consistency as discussed above will then also become clear.

4 RELATION TO THE EXISTING LITERATURE

As far as we know, McGuire (1968) et al. were the first authors who examined the restrictions on parameters and explanatory variables implied by a sum-constrained model, such that the sum constraint is automatically satisfied. This article formed the basis for the concept ‘logical consistency’. According to Weverbergh, Naert and Bultez (1981), logically consistent models are models specified in such a way that logical constraints defining the range of variation of the dependent variable are automatically satisfied by the values predicted by the model. Naert and Bultez (1973) stated that a logically consistent market share model should predict market shares between zero and one which sum to one. Naert and Bultez (1973) came up with necessary and sufficient conditions for a linear model to predict sum-constrained dependent variables. These authors derived that regressors must be sum-constrained for the linear model to be logically consistent. McGuire and Weiss (1976) observed that there was an error of omission in the proof of Naert and Bultez, because these authors overlooked the possibility of so-called homogeneous regressors. A homogeneous variable has the property that the ratio of this variable in any pair of equations is constant across all observations, although the value of this ratio is identical for different pairs of equation. McGuire and Weiss therefore corrected this error, yet also came to the conclusion that the sum constraint on the predicted market shares induces constraints on both parameters and explanatory variables. Their generally accepted results showed that a logically consistent sum-constrained linear model only allows homogeneous and sum-constrained regressors. Although the results of McGuire and Weiss apply to the most common types of sum-constrained models, their results are not as complete as they claim, a fact recognized by Weverbergh, Naert and Bultez (1981). From the discussions of Naert and Bultez and of McGuire and Weiss, Weverbergh, Naert and Bultez concluded that there must be a relation between the number of constraints on the parameters and the number of constraints on the explanatory variables when dealing with sum-constrained linear models. In this underexposed paper by Weverbergh, Naert and Bultez, necessary and sufficient conditions were derived for logical consistency of the sum-constrained linear model. These authors, however, also failed to give a proper definition of logical consistency, and derived results only for the case where the number of observations exceeds the number of explanatory variables plus one.

5 SOME EXAMPLES

In this section, we give some examples that illustrate the use of corollary 3.1 to determine the relations between the constraints on parameters and explanatory variables implied by logical data-consistency. We address the differences between $T \leq k + 1$ and $T > k + 1$. Subsequently, we consider what happens if we want to adopt definition 2.2 of logical consistency, so that we need to use corollary 3.3. We show when differences in interpretation of logical consistency may arise. The first two examples are derived from Weverbergh, Naert and Bultez (1981) and are intended to demonstrate the most common situations of homogeneous and sum-constrained regressors. The third example shows that it is possible to consider other types of regressors, and this is therefore a counterexample to the results of McGuire and Weiss.

EXAMPLE 1

A well-known application of a sum-constrained model is the market share response function where the market share is a function of lagged market share and advertising share (Beckwith 1972, Leeflang and Reuyl 1984). Consider three brands whose market shares y_{ti} are affected by two variables z_{tij} , $i = 1, 2, 3$, $j = 1, 2$, and $t = 1, \dots, T$. Let z_{ti1} denote the advertising share of brand i in the market at time t , and let z_{ti2} denote lagged market share $m_{t-1,i}$, $i = 1, 2, 3$. Let Z_{ij} denote the j th column of Z_i , $i = 1, 2, 3$, $j = 1, 2$. Then we can write

$$y_i = \alpha_i \iota_T + Z_{i1}\beta_{i1} + Z_{i2}\beta_{i2} + \varepsilon_i$$

for $i = 1, 2, 3$ with

$$\sum_{i=1}^3 y_i = \iota_T,$$

because market shares should sum to one. The regressors satisfy the restrictions

$$\sum_{i=1}^3 Z_{i1} = \sum_{i=1}^3 Z_{i2} = \iota_T. \quad (24)$$

Therefore, our matrix Z becomes

$$Z = (Z_{11} \ Z_{12} \ Z_{21} \ Z_{22} \ Z_{31} \ Z_{32} \ \iota_T). \quad (25)$$

Because of the sum constraints (24), the rank of Z equals at most five. In most cases, the rank of Z will indeed equal five as long as the number of rows (T) exceeds the number of columns ($k + 1$). We will assume that the rank of Z is equal to five. According to corollary 3.1, a necessary and sufficient condition for the model to be logically data-consistent is that the vector

$$\beta^* = (\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \beta_{31}, \beta_{32}, (\alpha_1 + \alpha_2 + \alpha_3 - 1))'$$

is of the form $\beta^* = \Lambda\lambda$, where Λ is an $7 \times \nu$ matrix, whose columns form a basis for the null space of Z , and where λ is a $\nu \times 1$ vector of proportionality factors. We know from corollary 3.2 that $\nu = 7 - 5 = 2$, and it is easy to check that the columns of the matrix

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix} \quad (26)$$

form a basis for the null space of Z , because

$$Z\Lambda = \left(\sum_{i=1}^3 Z_{i1} - \iota_T, \sum_{i=1}^3 Z_{i2} - \iota_T \right) = (0, 0).$$

Therefore,

$$\beta^* = \Lambda\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_1 \\ \lambda_2 \\ \lambda_1 \\ \lambda_2 \\ -\lambda_1 - \lambda_2 \end{pmatrix}, \quad (27)$$

so that we have the following restrictions on the parameters:

$$\beta_{1j} = \beta_{2j} = \beta_{3j} = \lambda_j$$

for $j = 1, 2$, and

$$\sum_{i=1}^3 \alpha_i + \lambda_1 + \lambda_2 - 1 = 0.$$

Note that the number of restrictions on the β_{ij} equals $\text{rank}(Z) - 1 = 4$, and that there is one restriction on the sum of the constant terms, as mentioned in section corollary 3.2.

If the rank of Z is smaller than five because $T < 5$, similar arguments can be used to derive necessary and sufficient conditions for logical data-consistency. That is, find the vectors spanning the null space of Z , and use a linear combination of these vectors to find the restrictions on the vector of parameters. If, for example, $T = 3$, so that Z only has three rows (which we assume to be independent), then the dimension of the null space equals $7 - 3 = 4$ according to (18). The matrix Λ then contains four columns. The first two columns are equal to the columns of the matrix Λ as defined in (26), yet there will be two additional columns. The parameter vector β^* now is a linear combination of four vectors (the vectors spanning the null space of Z) and will therefore be less restricted than the vector we found in (27), which is a combination of two vectors.

To derive necessary and sufficient conditions for the model to be logically predictor-consistent, observe that future observations of the regressors z_{tij} will also be sum-constrained, i.e., $\sum_i z_{tij} = 1, i = 1, \dots, n, j = 1, 2$, and $t = T + 1, \dots, T + s$. Adding rows to the matrix Z as defined in (25) will therefore not change the rank of Z in case $T > 5$. Therefore, the null space of the matrix Z will stay the same, so that the matrix Λ found in (26) can also be used in corollary 3.3. Therefore, the restrictions on the parameters will be the same as for logical data-consistency. In case $T < 5$, we showed above that the necessary and sufficient conditions for logical data-consistency will become less stringent. These conditions will then be neither sufficient nor necessary for logical predictor-consistency.

EXAMPLE 2

Now consider a variation of example 1 with the second explanatory variable (lagged market share) replaced by another variable which is the same for each brand, that is, $Z_{i2} = Z_{*2}, i = 1, 2, 3$. The variable Z_{*2} is a so-called homogeneous variable, as described in section 4. The matrix Z is now equal to

$$Z = (Z_{11} \quad Z_{*2} \quad Z_{21} \quad Z_{*2} \quad Z_{31} \quad Z_{*2} \quad \iota_T), \quad (28)$$

so that Z now has a rank of at most four. If we assume that the rank of Z is exactly equal to four (so that $T \geq 4$), then it is easy to see that the columns of the matrix

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}$$

form a basis for the null space of Z , for

$$Z\Lambda = \left(\sum_{i=1}^3 Z_{i1} - \iota_T, Z_{*2} - Z_{*2}, Z_{*2} - Z_{*2} \right) = (0, 0, 0).$$

There will be one restriction on the sum of the constant terms, and $\text{rank}(Z) - 1 = 3$ restrictions on the β_{ij} . Because

$$\beta^* = \Lambda \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 + \lambda_3 \\ \lambda_1 \\ -\lambda_2 \\ \lambda_1 \\ -\lambda_3 \\ -\lambda_1 \end{pmatrix}, \quad (29)$$

it follows that the parameters are restricted as follows:

$$\begin{aligned} \beta_{11} = \beta_{21} = \beta_{31} &= \lambda_1 \\ \beta_{12} + \beta_{22} + \beta_{32} &= 0, \end{aligned}$$

and

$$\sum_{i=1}^3 \alpha_i + \lambda_1 - 1 = 0.$$

As in example 1, if $T < 4$, then the rank of Z will also be smaller than four, and β^* will become less restricted. If, for example, $T = 2$ such that $\text{rank}(Z) = 2$, then the dimension of the null space equals $\nu = 7 - 2 = 5$, so that the matrix Λ will have two additional columns. The parameter vector β^* now is a linear combination of five vectors and will therefore be less restricted than the vector we found in (29), which is a combination of three vectors.

As far as logical predictor-consistency is concerned, a similar situation as in example 1 occurs. For additional observations $z_{tij}, t = T + 1, \dots, T + s$, the first regressor will still be sum-constrained, i.e. $\sum_{i=1}^n z_{ti1} = 1$, and the second regressor will also be homogeneous, i.e. $z_{ti2} = z_{t*2}$ for all i . It is easy to check that in this case, the rank of Z as defined in (28) will not change if we add these future observations. Therefore, the necessary and sufficient conditions found for logical data-consistency in case $T > 4$ are also necessary and sufficient for logical predictor-consistency. In case $T < 4$, the conditions for logical data-consistency will again differ from the conditions for logical predictor-consistency.

EXAMPLE 3

Consider the model of example 2, yet with advertising share replaced by the expenditures on advertising, so that we do not have the restriction $Z_{11} + Z_{21} + Z_{31} = \iota_T$ anymore. That is, we have two regressors, the first one is allowed to move freely, the second one is homogeneous. The matrix Z is the same as in example 2, yet the matrix Λ is different. Because the first regressor is not sum-constrained anymore, the rank of Z will be at most five. If we assume that it equals five, then the dimension of the null space of Z equals two, and we must delete the first column of Λ as in example 3. We can see that in this case

$$\beta^* = \Lambda \lambda = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda_1 + \lambda_2 \\ 0 \\ -\lambda_1 \\ 0 \\ -\lambda_2 \\ 0 \end{pmatrix},$$

so that $\beta_{i1} = 0$ for $i = 1, 2, 3$, which implies that advertising share cannot be a part of the model in this way. This example therefore shows that the conditions from corollary 3.1 sometimes indicate which regressors are not possible.

The above shows that the first regressor should also be restricted in some way. Therefore, suppose that the advertising share of the third brand is always equal to the sum of the advertising shares of the first two brands, that is,

$$Z_{11} + Z_{21} = Z_{31}. \quad (30)$$

Such a situation could occur in practice if a manufacturer has a certain advertising budget for a couple of brands which should be divided among these brands. The matrix Z is again the same as in example 2, yet with the new restrictions (30) on the regressors, the rank of Z will now be equal to four at most. If we assume that the rank equals four, then the dimension of the null space of Z equals three, and it is easy to see that the columns of the matrix

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

form a basis for the null space of Z . Therefore,

$$\beta^* = \Lambda \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 + \lambda_3 \\ \lambda_1 \\ -\lambda_2 \\ -\lambda_1 \\ -\lambda_3 \\ 0 \end{pmatrix}, \quad (31)$$

it follows that the parameters are restricted as follows:

$$\begin{aligned} \beta_{11} = \beta_{21} = -\beta_{31} &= \lambda_1 \\ \beta_{12} + \beta_{22} + \beta_{32} &= 0, \end{aligned}$$

and

$$\sum_{i=1}^3 \alpha_i - 1 = 0.$$

The advertising share restricted as in (30) is a regressor that is neither homogeneous nor sum-constrained, according to the definitions of section 2. Therefore, this example clearly illustrates the fact that for a model to be logically data-consistent or logically predictor-consistent, the regressor need not be homogeneous or sum-constrained. The discussion regarding the difference between logical data-consistency and logical predictor-consistency is analogous to the discussions in example 1 and example 2.

6 CONCLUSIONS

In this paper, we presented two (related) definitions of logical consistency of a sum-constrained linear model, namely, logical data-consistency and logical predictor-consistency. We derived necessary and sufficient conditions corresponding to each of the two definitions and showed where and when differences may occur. Weverbergh, Naert and Bultez (1981) derived comparable results, yet only considered the case $T > k + 1$. These authors also did not give a proper definition of logical consistency. We showed that the results also hold in case $T \leq k + 1$, but that in this case some caution is needed concerning the interpretation of logical consistency. We also presented the conditions of McGuire and Weiss (1976), which, as from that moment, were generally accepted to be necessary and sufficient for ‘logical consistency’. In both of our definitions of logical consistency, the conditions of McGuire and Weiss are sufficient, yet not always necessary. Moreover, we presented an example that clearly shows that the conditions of McGuire and Weiss are not necessary at all. We also demonstrated by means of an example that the conditions for logical consistency sometimes can be used to show that a particular regressor cannot be a part of the model. As far as the difference between logical data-consistency and logical predictor-consistency is concerned, we showed that in most cases, the conditions will be the same if $T > k + 1$. In case $T < k + 1$, the conditions for logical predictor-consistency are often more restrictive than those for logical data-consistency.

REFERENCES

- [1] Beckwith, N.E. (1972), "Multivariate analysis of sales responses of competing brands to advertising", *Journal of Marketing Research*, 9, 168-176
- [2] Bell, D.E., Keeney, R.L. and Little, J.D.C. (1975), "A market share theorem", *Journal of Marketing Research*, 12, 136-141
- [3] Bultez, A.V. and Naert, P.A. (1975), "Consistent sum-constrained models", *Journal of the American Statistical Association*, 70, 529-535
- [4] De Boer, P.M.C., and Harkema, R. (1984), "Maximum likelihood estimation of sum-constrained linear models when samples are small", *Kwantitatieve Methoden*, 16, 97-108
- [5] De Boer, P.M.C. and Harkema, R. (1997), "Sum-constrained models with undersized samples", *Statistica Neerlandica*, 51, 72-89
- [6] De Boer, P.M.C., Harkema, R. and Soede, A.J. (1996), "Maximum likelihood estimation of market share models with large numbers of shares", *Applied Economics Letters*, 3, 45-48
- [7] De Boer, P.M.C. and Martinez, C. (1999), "Spanish imports on manufactures and the European Union: an empirical assessment of the entry", *Empirica*, 26, 141-149
- [8] Friedberg, S.H., Insel, A. and Spence, L.E. (1992), *Linear Algebra*, Prentice Hall, Singapore
- [9] Gaver, K.M., Horsky, D. and Narasimhan, C. (1988), "Invariant estimators for market share systems and their finite sample behavior", *Marketing Science*, 7, 169-186
- [10] Ghosh, A., Neslin, S. and Schoemaker, R. (1984), "A comparison of market share models and estimation procedures", *Journal of Marketing Research*, 11, 202-210
- [11] Koehler, G.J. and Wildt, A.R. (1981), "Specification and estimation of logically consistent linear model", *Decision Sciences*, 12, 1-31
- [12] Leeflang, P.S.H. and Reuyl, J.C. (1984), "On the predictive power of market share attraction model", *Journal of Marketing Research*, 11, 211-215
- [13] McGuire, T.W. and Weiss, D.L. (1976), "Logically consistent market share models II", *Journal of Marketing Research*, 13, 296-302
- [14] McGuire, T.W., Farley, J.U., Lucas, R. E. and Ring, L. W. (1968), "Estimation and inference for linear models in which subsets of the dependent variable are constrained", *Journal of the American Statistical Association*, 63, 1201-1213
- [15] Nakanishi, M. and Cooper, L. G., (1974), "Parameter estimation for a multiplicative competitive interaction model—least squares approach", *Journal of Marketing Research*, 11, 303-311
- [16] Naert, P.A. and Bultez, A. (1973), "Logically consistent market share models", *Research Notes and Communications, Journal of Marketing Research*, 10, 334-340
- [17] Naert P.A. and Weverbergh M. (1981), "On the prediction power of market share attraction models", *Journal of Marketing Research*, 18, 146-153
- [18] Naert, P.A. and Weverbergh, M. (1985), "Market share specification, estimation and validation: toward reconciling seemingly divergent views", *Journal of Marketing Research*, 22, 453-46
- [19] Weverbergh, M. (1976), "Restrictions on linear sum-constrained models: A generalization", *University of Antwerp, Centre for Managerial Economics and Econometrics*, Antwerpen, Working paper
- [20] Weverbergh, M. and Naert, P. and Bultez, A. (1981), "Logically consistent market share models revisited", *Belgisch Tijdschrift voor Statistiek, Informatica en Operationeel Onderzoek*, 21, 3-37