International Journal of Forecasting 17 (2001) 121–128
www.elsevier.com/locate/ijforecast

Forecasting market shares from models for sales

Dennis Fok[a,*], Philip Hans Franses[b]

[a]Erasmus Research Institute of Management, Erasmus University Rotterdam, Faculty of Economics, H16-24, P.O. Box 1738, NL-3000 DR Rotterdam, Netherlands
[b]Econometric Institute, Department of Marketing and Organization, Erasmus University Rotterdam, Rotterdam, Netherlands

Abstract

Dividing forecasts of brand sales by a forecast of category sales, when they are generated from brand specific sales–response models, renders biased forecasts of the brands' market shares. In this note we propose as an alternative a simulation-based method which results in unbiased forecasts of market shares. An application of this forecasting technique to a five brand tuna fish market illustrates its practical relevance.

Keywords: Sales models; Market shares; Forecasting

1. Introduction

Market researchers often focus on modeling and forecasting marketing performance measures, such as brand choice and interpurchase times at the individual household level and sales and market shares at the aggregated level (see e.g. Leeflang, Wittink, Wedel and Naert (2000) and Franses and Paap (2001) for recent surveys). Household-specific data usually concerns cross sections or panels, while aggregated data often concerns weekly or monthly time series observations. Cross-sectional or panel data have the advantage that the effects of marketing instruments can be observed at the individual level, while a potential disadvantage is that often one needs to account for unobserved heterogeneity across households. In contrast, time series observations do not suffer from such heterogeneity, but there it is only possible to draw inference at the aggregate level and one has to take account of possibly complicated dynamic patterns.

Here the focus is on forecasting market shares at the brand level. Indeed, market shares can be of particular interest, as shares automatically imply that a manager can evaluate the sales performance relative to the performance of the product category. Also, market shares are less sensitive to the impact of growth and seasonal fluctuations. One obvious question is whether one should construct a quantitative

*Corresponding author. Tel.: +31-104-088-947; fax: +31-104-089-031.
E-mail address: dfok@few.eur.nl (D. Fok).

0169-2070/01/$ – see front matter
PII: S0169-2070(00)00075-3

model for market shares using, for example, the familiar attraction model (see Brodie & Kluyver, 1987; Kumar, 1994), or that one should construct a model for sales using for example the SCAN*PRO model (see Wittink, 1987; Wittink, Addona, Hawkes & Porter, 1988; Foekens, Leeflang & Wittink, 1994; Van Heerde, Leeflang & Wittink, 2000). In the latter case one can then use these models to generate sales forecasts and, given these, forecasts of market shares.

In this note we will confine ourselves to the question how one can generate forecasts for shares given models for sales, as this turns out not to be a trivial exercise. We will indicate that simply dividing brand sales forecasts by category sales forecasts, which seems to be common practice, yields biased forecasts for market shares. Hence, one needs to resort to an alternative method. We propose a simulation-based method to obtain unbiased forecasts. Simulation-based methods have become increasingly more common in models for marketing performance measures such as brand choice and interpurchase times (see e.g. Allenby & Rossi, 1999; Bronnenberg, Mahajan & Vanhonacker, 2000).

The outline of this paper is as follows. In Section 2, we discuss two methods for forecasting market shares given models for sales. The first method is the above-mentioned division of forecasts, which will be called the naive method, and the second is the more appropriate simulation based method, denoted by SB. In Section 3, we illustrate the practical relevance of the SB method for an example concerning five brands of tuna fish. In Section 4, we conclude with some remarks.

2. Forecasting market shares

Suppose that there are I brands in a certain product category, and suppose the availability of weekly scanner data on sales and various explanatory variables. One possible form of a sales model assumes a multiplicative specification to relate explanatory variables such as promotion and price to current sales (see, for example, Wittink et al., 1988), although other forms are of course also possible (see Leeflang et al., 2000). The sales of brand i, $i = 1, \ldots, I$ at time t, $t = 1, \ldots, T$, denoted by $S_{i,t}$, are then modeled as:

$$S_{i,t} = \exp(\mu_i + \varepsilon_{i,t}) \prod_{j=1}^{I} \prod_{k=1}^{K} \exp(x_{k,j,t})^{\beta_{k,j,i}} \qquad (1)$$

where $\varepsilon_t \equiv (\varepsilon_{1,t}, \ldots, \varepsilon_{I,t})' \sim N(0, \Sigma)$ and where $x_{k,j,t}$ denotes the k-th explanatory variable (for example, price or advertising) for brand j at time t and $\beta_{k,j,i}$ is the corresponding coefficient for brand i. The normality assumption for the error terms is not strictly necessary, although it does facilitate straightforward estimation of the model parameters. The parameter μ_i is a brand-specific constant. The error process $\varepsilon_{i,t}$ is usually assumed to be only correlated across brands and not over time, that is, $\varepsilon_{i,t}$ is assumed to be independent of $\varepsilon_{j,t-1}$, $j = 1, \ldots, I$. Finally, by using the exponential transformation, one ensures that sales forecasts are always positive.

To capture lagged structures in (1), one can include lagged sales in the specification. The most general autoregressive structure follows from the inclusion of lagged sales of all brands. In that case, when a P-th order autoregressive structure is used, the model becomes:

$$S_{i,t} = \exp(\mu_i + \varepsilon_{i,t}) \prod_{j=1}^{I} \prod_{k=1}^{K} \exp(x_{k,j,t})^{\beta_{k,j,i}} \prod_{j=1}^{I} \prod_{p=1}^{P} S_{j,t-p}^{\alpha_{p,j,i}} \qquad (2)$$

To estimate the parameters, the model usually is linearized by taking natural logarithms of the sales. The resulting equations form an I-dimensional vector autoregressive model with ex-