Modeling Potentially Time-Varying Effects of Promotions on Sales

Philip Hans Franses, Richard Paap, Philip A. Sijthoff

ERIM Report Series *Research in Management*	
ERIM Report Series reference number	ERS-2001-05-MKT
Publication	January 2001
Number of pages	27
Email address corresponding author	Franses@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) Rotterdam School of Management / Faculteit Bedrijfskunde Erasmus Universiteit Rotterdam PoBox 1738 3000 DR Rotterdam, The Netherlands Phone: # 31-(0) 10-408 1182 Fax: # 31-(0) 10-408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS **R**ESEARCH **I**NSTITUTE OF **M**ANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	A commonly applied modeling tool for the analysis of promotional effects on weekly sales data is a linear regression model. Usually, such a model includes 0/1 dummy variables for promotions, where weeks with a promotion get a value of 1. When these variables are included in a model with parameters which are constant over time, the market researcher implicitly makes two important but rather restrictive assumptions. The first is that anytime a dummy variable takes a value of 1 and the relevant parameter is significant, there is a non-zero effect of promotion on sales. The second is that this effect is constant across all weeks. In many practical cases however, one may conjecture that the effects of promotion are not constant over time. Therefore, we propose a new and rather parsimonious econometric model for the purpose of measuring the effects of promotions, while allowing for time-variation in these effects. The main idea is that promotions can (but not necessarily) lead to positive and suddenly large values of sales in the same week, and that they can perhaps lead to large negative values in the week there-after, if there is a, what is called, post-promotion dip. We discuss representation and interpretation of the model, and we outline the maximum likelihood parameter estimation method. Simulation results suggest that the estimation method is quite reliable and that the distribution of the estimator is approximately normal. We illustrate the model in substantial detail on two sets of empirical data in order to indicate its practical usefulness	
Library of Congress Classification (LCC)	5001-6182	Business
	5410-5417.5	Marketing
	HF 5438.5	Sales Promotion
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 31	Marketing
	C 44	Statistical Decision Theory
	M 39	Marketing and Advertising: Other
European Business Schools Library Group (EBSLG)	85 A	Business General
	280 G	Managing the marketing function
	255 A	Decision theory (general)
	280 K	Marketing Science (Quantitative analysis)
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.40	Marketing
	85.03	Methoden en technieken, operations research
	85.40	Marketing
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Marketing / Besliskunde	
	Promotie (Public Relations), Econometrische Modellen	
Free keywords	Sales, Promotions, Time-Varying Effects, Censored Regression	
Other information		

Modeling Potentially Time-Varying Effects of Promotions on Sales

Philip Hans Franses*
Econometric Institute and
Department of Marketing and Organization
Erasmus University Rotterdam

Richard Paap
Rotterdam Institute for Business Economic Studies
Erasmus University Rotterdam

Philip A. Sijthoff
Econometric Institute
Erasmus University Rotterdam

*We thank Martijn van der Molen for research assistance. The data and Gauss programs used in this paper can be obtained from the corresponding author. The address for correspondence is P.H. Franses, Econometric Institute H11-15, Erasmus University Rotterdam, P.O.Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: franses@few.eur.nl

Modeling Potentially Time-Varying Effects of Promotions on Sales

Abstract

A commonly applied modeling tool for the analysis of promotional effects on weekly sales data is a linear regression model. Usually, such a model includes 0/1 dummy variables for promotions, where weeks with a promotion get a value of 1. When these variables are included in a model with parameters which are constant over time, the market researcher implicitly makes two important but rather restrictive assumptions. The first is that anytime a dummy variable takes a value of 1 and the relevant parameter is significant, there is a non-zero effect of promotion on sales. The second is that this effect is constant across all weeks.

In many practical cases however, one may conjecture that the effects of promotion are not constant over time. Therefore, we propose a new and rather parsimonious econometric model for the purpose of measuring the effects of promotions, while allowing for time-variation in these effects. The main idea is that promotions can (but not necessarily) lead to positive and suddenly large values of sales in the same week, and that they can perhaps lead to large negative values in the week thereafter, if there is a, what is called, post-promotion dip. We discuss representation and interpretation of the model, and we outline the maximum likelihood parameter estimation method. Simulation results suggest that the estimation method is quite reliable and that the distribution of the estimator is approximately normal. We illustrate the model in substantial detail on two sets of empirical data in order to indicate its practical usefulness

Key words: Sales, Promotions, Time-Varying Effects, Censored Regression

1 Introduction

Promotions like features and display are key marketing-mix instruments. From a managerial point of view it is important to understand their impact on, for example, sales, brand choice and interpurchase times, see Gupta (1988), Blattberg, Briesch and Fox (1995), Blattberg and Neslin (1990), Kumar and Leone (1988) and Moriarty (1985), to mention just a few of the many contributions. Recent studies in, say, the last five years, on the impact of promotions tend to focus on at least two issues. The first issue concerns the potentially different short-run and long-run effects on sales and market shares, see for example Dekimpe and Hanssens (1995a,b), and on brand choice, see for example Papatla and Krishnamurti (1996) and Paap and Franses (2000). The second important issue appears to be the potentially time-varying impact of promotions, thereby allowing for, what is called, postpromotion dips, see Foekens et al. (1999), Neslin and Schneider-Stone (1996) and Van Heerde et al. (2000), among others. In the present paper we aim to contribute to the literature on this second issue, that is, we also address this issue of potentially time-varying effects of promotions, where we particularly focus on weekly scanner data on sales.

A typical modeling tool for the analysis of promotional effects on weekly sales data is a linear regression model, see Leeflang et al. (2000). An advanced version of such a model is the SCAN*PRO model, proposed by Wittink et al. (1988). Usually, the regression model includes 0/1 dummy variables for display and feature promotions, where weeks with a promotion get a value of 1. When these variables are included in a regression model, and when it is additionally assumed that their effect can be summarized by a parameter which is constant over time, then one implicitly makes two important assumptions. The first is that anytime a dummy variable takes a value of 1 and the associated parameter is statistically significant, there is a non-zero effect on sales. The second is that this effect is constant over time. An interesting approach to handle the second assumption is put forward in Foekens, Leeflang and Wittink (1999), where the promotion parameters are made a function of explanatory variables. Generally, their approach, however, cannot relax the first assumption, that is, it does not explicitly include the possibility that promotions can have zero effect in some weeks. Additionally, one would want to allow for the potential presence of for example post-promotion dips, which are also not necessarily of similar size over time. In the present paper we therefore propose a new and parsimonious econometric model, which does allow for substantial flexibility concerning the potential time-varying effects of promotion. It should already be stressed that this flexibility does not come at the cost of many parameters. In fact, in our empirical illustration we will see that a few extra parameters are sufficient to allow for substantial flexibility. The main idea of our model is that promotions can (but not necessarily will) lead to positive and suddenly large values of sales in the same week, and that they can lead to large negative values in the week thereafter, which then can be associated with a post-promotion dip.

The outline of our paper is as follows. In Section 2, we propose the new model. We discuss its representation and its interpretation. In Section 3, we discuss a method for parameter estimation, which is based on maximum likelihood [ML]. Various technicalities are relegated to an appendix. Using simulations we examine the practical precision of this method. We find that for samples of 125 data points or more, the ML method is quite reliable and that the distribution of the estimator is approximately normal. In Section 4, we consider our new model for 36 sets of series, containing sales, prices and display and feature promotion. We do this in order to investigate if the model can be useful in general. For various data sets we find that the model yields plausible results, although in several cases the analysis suffers from a shortage of useful data points. To illustrate the potential relevance of our model, we zoom in on the specific results for two data sets. Finally, in Section 5, we conclude with some remarks and various topics for further research.

2 A New Model

In this section we propose a new, and essentially nonlinear, econometric model for describing potentially time-varying effects of promotions on weekly sales. To save notation, we explicitly focus on extending and modifying a rather stylized regression type model. We are aware of the fact that there are various ways of extending our model and we will indicate so in the section with concluding remarks, but for the present paper the stylized model suffices. In Section 2.1, we discuss the model representation. In Section 2.2, we elaborate on the interpretation of the model.

2.1 Representation

Consider a sales variable y_t (transformed by taking natural logarithms denoted as "log"), which is weekly observed for $t = 1, 2, \ldots, T$, and suppose there are observations on explanatory variables like the log of price and the log of advertising collected in z_t and that there are 0/1 dummy variables on promotional activities collected in x_t. The motivation for taking logs is that the parameters in the regression model can be interpreted as elasticities. For further reference, the sets of variables z_t and x_t can contain a column of ones in order to capture mean effects. A rather stylized regression-type model for sales is now given by

$$y_t = \rho y_{t-1} + z_t' \beta + x_t' \gamma + \varepsilon_t, \tag{1}$$

where it is usually assumed that $\varepsilon_t \sim \text{NID}(0, \sigma_\varepsilon^2)$. The inclusion of only one lag of the sales variable (that is, y_{t-1}) is for notational convenience, where it should be mentioned that in the empirical section below we find that first-order dynamics are sufficient for our exemplary data. The model in (1) imposes that the effect of promotions is always equal to γ. To allow for a post-promotion dip, one can include x_{t-1} variables in (1), but then still, one would impose a constant effect of these lagged variables.

To make model (1) more flexible, that is, to allow for time-varying effects of promotion, Foekens et al. (1999) make the parameter γ a function of various variables. A possible drawback of this approach is that one has to choose those variables. Additionally, this approach does not explicitly allow for zero effects. Another way of introducing flexibility is to abstain from a linear specification and to consider semi- or non-parametric methods, see Van Heerde et al. (2000). A possible drawback here is that the estimation results may become hard to interpret.

In this paper we propose a model which modifies (1) such that (i) there are time-varying effects of promotions, that (ii) there are also time-varying sizes of post-promotion dips, and (iii) that promotions can have zero effects, while ensuring that the model parameters can still easily be interpreted. More precise, we propose to modify (1) into

$$y_t = \rho y_{t-1} + z_t'\beta + v_t + w_t + \varepsilon_t, \tag{2}$$

where v_t and w_t are defined by

$$v_t = \begin{cases} x_t'\gamma_1 + u_{1,t} & \text{if } x_t'\gamma_1 + u_{1,t} \geq 0 \\ 0 & \text{if } x_t'\gamma_1 + u_{1,t} < 0 \end{cases} \tag{3}$$

and

$$w_t = \begin{cases} x_{t-1}'\gamma_2 + u_{2,t} & \text{if } x_{t-1}'\gamma_2 + u_{2,t} \leq 0 \\ 0 & \text{if } x_{t-1}'\gamma_2 + u_{2,t} > 0, \end{cases} \tag{4}$$

where we assume that $u_{1,t} \sim \text{NID}(0, \sigma_{u_1}^2)$ and $u_{2,t} \sim \text{NID}(0, \sigma_{u_2}^2)$, and where it is assumed that ε_t, $u_{1,t}$ and $u_{2,t}$ are mutually uncorrelated. The v_t and w_t variables model the possible current and lagged effect of the promotions collected in x_t on current sales. Current promotion activity has no effect, that is, v_t equals zero unless a linear combination of promotion variables $(x_t'\gamma_1)$ exceeds a stochastic threshold $-u_{1,t}$, in which case v_t measures a positive effect of promotion. Similarly, promotion activity in the previous week has no effect on current sales, that is, w_t is zero unless a linear combination of one-week lagged variables $(x_{t-1}'\gamma_2)$ is below a stochastic threshold $-u_{2,t}$, in which case w_t measures the negative effect due to a post-promotion dip.

To understand the main idea of the model, it can be useful to examine two polar cases. When v_t is always positive and w_t is always negative, the model in (2)-(4) reduces to the linear model in (1) with the extra regressor x_{t-1}. Second, when promotions never have an effect (positive or negative), the variables v_t and w_t do not enter (1) for any t , and hence the variables cannot be observed. In both cases, one would encounter estimation problems in practice, as some of the parameters in (3) and (4) are not identified. The interesting case of course is the case where v_t and w_t are sometimes relevant, thereby indicating time-varying effects of promotions and the time-varying occurrence of post-promotion dips.

2.2 Unconditional Inference

The model in (2)-(4) can be interpreted in various ways. The introduction of error terms $u_{1,t}$ and $u_{2,t}$ in (3) and (4), respectively, allows for uncertainty about the potential effects of promotions, additional to the usual standard deviation of the parameter estimate in a linear regression model. As a useful consequence of these error terms, we are now able to compute the following joint probabilities, that is

$$\begin{aligned}
\Pr[v_t = 0, w_t = 0|x_t, x_{t-1}; \theta] &= \Pr[v_t = 0|x_t; \theta] \Pr[w_t = 0|x_{t-1}; \theta] \\
\Pr[v_t > 0, w_t = 0|x_t, x_{t-1}; \theta] &= \Pr[v_t > 0|x_t; \theta] \Pr[w_t = 0|x_{t-1}; \theta] \\
\Pr[v_t > 0, w_t < 0|x_t, x_{t-1}; \theta] &= \Pr[v_t > 0|x_t; \theta] \Pr[w_t < 0|x_{t-1}; \theta] \\
\Pr[v_t = 0, w_t < 0|x_t, x_{t-1}; \theta] &= \Pr[v_t = 0|x_t; \theta] \Pr[w_t < 0|x_{t-1}; \theta],
\end{aligned} \tag{5}$$

where $\theta = \{\rho, \beta, \sigma_\varepsilon, \gamma_1, \sigma_{u_1}, \gamma_2, \sigma_{u_2}\}$ summarizes the model parameters, and where we have used that $u_{1,t}$ and $u_{2,t}$ are uncorrelated. The first probability for example can be interpreted as the probability of no effect of promotion and no post-promotion dip in week t, given the explanatory dummy variables in x_t and x_{t-1}, while for example the third probability concerns the probability of observing both effects. Obviously, it holds that the sum of the probabilities in (5) equals 1.

The probability that promotions in week t do not have an effect on sales is given by

$$\begin{aligned}
\Pr[v_t = 0|x_t; \theta] &= \Pr[x_t'\gamma_1 + u_{1,t} < 0|x_t; \theta] \\
&= \int_{-\infty}^{-x_t'\gamma_1} \frac{1}{\sigma_{u_1}} \phi(u_{1,t}/\sigma_{u_1}) du_{1,t} = \Phi\left(\frac{-x_t'\gamma_1}{\sigma_{u_1}}\right) = \Phi_{1,t},
\end{aligned} \tag{6}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the probability density function (pdf) and cumulative distribution function (cdf) of the standard normal distribution, respectively. Obviously, the probability that a promotion in week t has a positive effect is given by $(1-\Phi_{1,t})$. Similarly, one can derive the probability of the absence of a post-promotion dip in week t, that is,

$$\begin{aligned}
\Pr[w_t = 0|x_t; \theta] &= \Pr[x_{t-1}'\gamma_2 + u_{2,t} > 0|x_{t-1}; \theta] \\
&= \int_{-x_{t-1}'\gamma_2}^{\infty} \frac{1}{\sigma_{u_2}} \phi(u_{2,t}/\sigma_{u_2}) du_{2,t} = 1 - \Phi\left(\frac{-x_{t-1}'\gamma_2}{\sigma_{u_2}}\right) = 1 - \Phi_{2,t}
\end{aligned} \tag{7}$$

and hence the probability of the presence of such a dip equals $\Phi_{2,t}$.

As v_t is a censored variable, the expected effect of a promotion in week t does not equal $x_t'\gamma_1$. In fact, we can derive that it equals

$$\begin{aligned}
\mathrm{E}[v_t|x_t;\theta] &= \mathrm{E}[v_t|v_t=0,x_t;\theta]\Pr[v_t=0|x_t;\theta] + \mathrm{E}[v_t|v_t\neq 0,x_t;\theta]\Pr[v_t\neq 0|x_t;\theta] \\
&= 0 + \left(x_t'\gamma_1 + \sigma_{u_1}\frac{\phi_{1,t}}{1-\Phi_{1,t}}\right)(1-\Phi_{1,t}) \\
&= x_t'\gamma_1(1-\Phi_{1,t}) + \sigma_{u_1}\phi_{1,t},
\end{aligned} \tag{8}$$

with $\phi_{1,t} = \phi(-x_t'\gamma_1/\sigma_{u_1})$, see Maddala (1983, p. 365) and Gourieroux en Monfort (1995, p. 483, property B.45) . Likewise, the expected size of the post-promotion dip effect equals

$$\begin{aligned}
&\mathrm{E}[w_t|x_{t-1};\theta] \\
&= \mathrm{E}[w_t|w_t=0,x_{t-1};\theta]\Pr[w_t=0|x_{t-1};\theta] + \mathrm{E}[w_t|w_t\neq 0,x_{t-1};\theta]\Pr[w_t\neq 0|x_{t-1};\theta] \\
&= 0 + \left(x_{t-1}'\gamma_2 + \sigma_{u_2}\frac{-\phi_{2,t}}{\Phi_{2,t}}\right)\Phi_{2,t} \\
&= x_{t-1}'\gamma_2\Phi_{2,t} + \sigma_{u_2}\phi_{2,t},
\end{aligned} \tag{9}$$

where $\phi_{2,t} = \phi(-x_{t-1}'\gamma_2/\sigma_{u_2})$. The above expressions all indicate that the model allows for a time-varying probability of an effect of promotion and for a time-varying size of these effects.

3 Parameter Estimation and Conditional Inference

In this section we discuss the maximum likelihood [ML] method for estimating the parameters in our new model. Next, we evaluate the small sample performance of this method through simulation experiments. Finally, we discuss the construction of residuals and conditional inference, which involves the observed sales variable.

3.1 Estimation by Maximum Likelihood

To derive that likelihood function, we first consider the pdf of y_t given y_{t-1}, z_t, and given v_t and w_t. As our model can be written as

$$y_t|y_{t-1},z_t,v_t,w_t \sim \mathrm{NID}(\rho y_{t-1} + z_t\beta + v_t + w_t, \sigma_\varepsilon^2), \tag{10}$$

we have

$$f(y_t|y_{t-1},z_t,v_t,w_t;\theta) = \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma_\varepsilon^2}(y_t-\rho y_{t-1}-z_t'\beta-v_t-w_t)^2\right). \tag{11}$$

The pdf of y_t given y_{t-1}, x_t and x_{t-1} is

$$
\begin{aligned}
&f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta) = \\
&\quad \Pr[v_t = 0, w_t = 0|x_t, x_{t-1}; \theta] f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t=0, w_t=0} + \\
&\quad \Pr[w_t = 0, x_{t-1}; \theta] \int_{-x_t'\gamma_1}^{\infty} \frac{1}{\sigma_{u_1}} \phi(u_{1,t}/\sigma_{u_1}) f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t=x_t'\gamma_1+u_{1,t}, w_t=0} du_{1,t} + \\
&\quad \Pr[v_t = 0, x_t; \theta] \int_{-\infty}^{-x_{t-1}'\gamma_2} \frac{1}{\sigma_{u_2}} \phi(u_{2,t}/\sigma_{u_2}) f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t=0, w_t=x_{t-1}'\gamma_2+u_{2,t}} du_{2,t} + \\
&\quad \int_{-\infty}^{-x_{t-1}'\gamma_2} \int_{-x_t'\gamma_1}^{\infty} \frac{\phi(u_{1,t}/\sigma_{u_1})}{\sigma_{u_1}} \frac{\phi(u_{2,t}/\sigma_{u_2})}{\sigma_{u_2}} f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t=x_t'\gamma_1+u_{1,t}, w_t=x_{t-1}'\gamma_2+u_{2,t}} du_{1,t} du_{2,t},
\end{aligned} \tag{12}
$$

which is relevant for the likelihood function.

The overall log-likelihood is given by

$$
\ell(Y|Z, X; \theta) = \sum_{t=2}^{T} \ln f(y_t|y_{t-1}, x_t, x_{t-1}; \theta), \tag{13}
$$

where $Y = (y_1, \ldots, y_T)$, $Z = (z_1, \ldots, z_T)$ and $X = (x_1, \ldots, x_T)$. The ML estimator $\hat{\theta}$ is obtained by maximizing this log-likelihood function for the parameter vector θ. This can be done using a standard numerical optimization routine like Newton-Raphson or the BHHH algorithm. To facilitate the computation of the likelihood, we can write the integrals in (12) in terms of pdf and cdfs of a (bivariate) normal distribution, see Appendix A.

3.2 Small Sample Properties

Before we turn to empirical illustrations of our new model, it seems wise to see to what extent the ML estimation routine outlined in the above subsection yields reliable results. More precise, we are interested as to whether the estimator $\hat{\theta}$ is consistent and whether the asymptotic distribution of the estimator is approximately normal. Additionally, we want to analyze the properties of the estimator in relatively small samples, including the typical sample size encountered in marketing, which amounts to about 2 to $2\frac{1}{2}$ years of weekly data. For this purpose, we generate data from the following data generating process [DGP], that is,

$$
y_t = \mu + \rho y_{t-1} + v_t + w_t + \varepsilon_t \tag{14}
$$

with $\varepsilon_t \sim \text{NID}(0, \sigma^2_\varepsilon)$ and

$$v_t = \begin{cases} \gamma_{1,1} + \gamma_{1,2}x_t + u_{1,t} & \text{if } \gamma_{1,1} + \gamma_{1,2}x_t + u_{1,t} \geq 0 \\ 0 & \text{if } \gamma_{1,1} + \gamma_{1,2}x_t + u_{1,t} < 0 \end{cases} \tag{15}$$

and

$$w_t = \begin{cases} \gamma_{2,1} + \gamma_{2,2}x_{t-1} + u_{2,t} & \text{if } \gamma_{2,1} + \gamma_{2,2}x_{t-1} + u_{2,t} \leq 0 \\ 0 & \text{if } \gamma_{2,1} + \gamma_{2,2}x_{t-1} + u_{2,t} > 0 \end{cases} \tag{16}$$

with $u_{i,t} \sim \text{NID}(0, \sigma^2_{u_i})$ for $i = 1, 2$, and where $x_t \sim \text{NID}(0, 1)$. The parameters in the DGP appear in the second column of Table 1. These parameters are chosen such that they imply that in about half the sample we have non-zero observations on the two censored variables.

Insert Table 1 about here.

We generate 1000 times 125, 250 and 1000 observations on the various variables, and we estimate the parameters in the models, where we take the true parameters as the starting values. Table 1 provides the relevant simulation results. The first conclusion that can be drawn from the third and fourth column of the table going from the first panel (125 observations) to the third panel (1000 observations) is that the ML estimator becomes increasingly more precise with increasing sample size, thereby suggesting consistency.

The final 6 columns of Table 1 deal with the distributional properties of the ML estimator, where we focus on the so-called z-scores. We compute the tail probabilities of these z-scores for different percentiles of the standard normal distribution. We observe that, even for a sample as small as 125 observations, the distribution of the z-scores is rather close to the normal distribution. When we increase the sample size, we see an ever closer match and hence the estimator seems asymptotically normally distributed.

In sum, these simulation results suggest that we may safely use the ML estimation routine in practice. Of course, the smaller the sample, the more cautious one should be. However even in the case of 125 observations, we find reasonably reliable results.

3.3 Residuals and Fit

The new model contains unobserved variables and this makes that the construction of residuals is a little different than for the standard regression case. Residuals in (2) can be defined as the difference between y_t and its conditional mean, that is

$$\begin{aligned}\hat{\varepsilon}_t &= y_t - \mathrm{E}[y_t|y_{t-1}, x_t, x_{t-1}; \hat{\theta}] \\ &= y_t - \hat{\rho} y_{t-1} - z_t'\hat{\beta} - \mathrm{E}[v_t|x_t; \hat{\theta}] - \mathrm{E}[w_t|x_{t-1}; \hat{\theta}].\end{aligned} \tag{17}$$

where $\mathrm{E}[v_t|x_t;\theta]$ and $\mathrm{E}[w_t|x_{t-1};\theta]$ are defined in (8) and (9). The fit of the model, that is, $\mathrm{E}[y_t|y_{t-1}, x_t, x_{t-1}; \hat{\theta}]$, is thus given by

$$\hat{\rho} y_{t-1} + z_t'\hat{\beta} + \mathrm{E}[v_t|x_t; \hat{\theta}] + \mathrm{E}[w_t|x_{t-1}; \hat{\theta}], \tag{18}$$

which involves the computation of a few integrals.

3.4 Conditional Inference

In Section 2 we have discussed inference on the unobserved components of our model v_t and w_t. This inference is unconditional on the history and current value of y_t. If we however know the value of y_t we can make conditional inference on the probability of the presence of a promotion effect and/or of a post-promotion dip. For example, the probability that there was no promotion effect and no post-promotion dip in week t, given y_t and y_{t-1}, is

$$\begin{aligned}&\Pr[v_t = 0, w_t = 0|y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta] = \\ &\qquad \frac{\Pr[v_t = 0, w_t = 0|x_t, x_{t-1}; \theta] f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t=0, w_t=0}}{f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta)},\end{aligned} \tag{19}$$

which is equal to the first part of the sum in (12) divided by the pdf of y_t. Likewise we obtain

$$\begin{aligned}
&\Pr[v_t > 0, w_t = 0|y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta] = \\
&\quad \frac{\Pr[w_t = 0, x_{t-1}; \theta] \int_{-x_t'\gamma_1}^{\infty} \frac{1}{\sigma_{u_1}} \phi(u_{1,t}/\sigma_{u_1}) f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t = x_t'\gamma_1 + u_{1,t}, w_t = 0} du_{1,t}}{f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta)} \\
&\Pr[v_t = 0, w_t < 0|y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta] = \\
&\quad \frac{\Pr[v_t = 0, x_t; \theta] \int_{-\infty}^{-x_{t-1}'\gamma_2} \frac{1}{\sigma_{u_2}} \phi(u_{2,t}/\sigma_{u_2}) f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t = 0, w_t = x_{t-1}'\gamma_2 + u_{2,t}} du_{2,t}}{f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta)} \\
&\Pr[v_t > 0, w_t > 0|y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta] = \\
&\quad \frac{\int_{-\infty}^{-x_{t-1}'\gamma_2} \int_{-x_t'\gamma_1}^{\infty} \frac{\phi(u_{1,t}/\sigma_{u_1})}{\sigma_{u_1}} \frac{\phi(u_{2,t}/\sigma_{u_2})}{\sigma_{u_2}} f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t = x_t'\gamma_1 + u_{1,t}, w_t = x_{t-1}'\gamma_2 + u_{2,t}} du_{1,t} du_{2,t}}{f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta)},
\end{aligned} \tag{20}$$

indicating the conditional probability of the presence of promotional and post-promotional effects in week t given y_t and y_{t-1}.

The above formulas can also be used to obtain marginal and conditional probabilities. For example, the marginal conditional probability that there was no post-promotion dip in week t is

$$\begin{aligned}&\Pr[w_t = 0|y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta] = \\ &\quad \Pr[v_t = 0, w_t = 0|y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta] + \Pr[v_t > 0, w_t = 0|y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta]. \quad (21)\end{aligned}$$

As a second example, the conditional probability that a promotion in week t had an effect, given that there was no post-promotion dip in week t, is

$$\Pr[v_t > 0|w_t = 0, y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta] = \frac{\Pr[v_t > 0, w_t = 0|, y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta]}{\Pr[w_t = 0, y_t, y_{t-1}, z_t, x_t, x_{t-1}; \theta]}. \quad (22)$$

In (8) and (9) we have derived the expected value of w_t and v_t, where we did not condition on y_t and y_{t-1}. Hence, these expected values can be used to predict the effects of promotion and post-promotion dips in week t. However, once (log) sales has been observed, we can estimate these effects by conditioning on y_t. The expected effect of promotions x_t in week t, conditional on y_t and y_{t-1}, is

$$\begin{aligned}&\mathrm{E}[v_t|x_t, x_{t-1}, z_t, y_t, y_{t-1}; \theta] = \\ &\frac{\Pr[w_t = 0, x_{t-1}; \theta] \int_{-x_t'\gamma_1}^{\infty} v_t \frac{1}{\sigma_{u_1}} \phi(u_{1,t}/\sigma_{u_1}) f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t = x_t'\gamma_1 + u_{1,t}, w_t = 0} du_{1,t}}{f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta)} + \\ &\frac{\int_{-\infty}^{-x_{t-1}'\gamma_2} \int_{-x_t'\gamma_1}^{\infty} v_t \frac{\phi(u_{1,t}/\sigma_{u_1})}{\sigma_{u_1}} \frac{\phi(u_{2,t}/\sigma_{u_2})}{\sigma_{u_2}} f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t = x_t'\gamma_1 + u_{1,t}, w_t = x_{t-1}'\gamma_2 + u_{2,t}} du_{1,t} du_{2,t}}{f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta)}.\end{aligned}$$
(23)

Note that this expression can be simplified using the results (31), (32) and (36) in Appendix A. To compute the expectation of a truncated univariate distribution, we refer to Maddala (1993, p. 365). To compute the expectation of a truncated bivariate distribution we can use the results in Rosenbaum (1961), see also Maddala (1993, p. 386).

Along similar lines, we can compute the expected value of the post-promotion dip in week t, conditional on y_t and y_{t-1}, which is given by

$$\begin{aligned}
&\mathrm{E}[w_t|x_t, x_{t-1}, z_t, y_t, y_{t-1}; \theta] = \\
&\frac{\Pr[v_t = 0, x_t; \theta] \int_{-\infty}^{-x'_{t-1}\gamma_2} w_t \frac{1}{\sigma_{u_2}} \phi(u_{2,t}/\sigma_{u_2}) f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t=0, w_t=x'_{t-1}\gamma_2+u_{2,t}} du_{2,t}}{f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta)} + \\
&\frac{\int_{-\infty}^{-x'_{t-1}\gamma_2} \int_{-x'_t\gamma_1}^{\infty} w_t \frac{\phi(u_{1,t}/\sigma_{u_1})}{\sigma_{u_1}} \frac{\phi(u_{2,t}/\sigma_{u_2})}{\sigma_{u_2}} f(y_t|y_{t-1}, z_t, v_t, w_t; \theta)|_{v_t=x'_t\gamma_1+u_{1,t}, w_t=x'_{t-1}\gamma_2+u_{2,t}} du_{1,t} du_{2,t}}{f(y_t|y_{t-1}, z_t, x_t, x_{t-1}; \theta)}.
\end{aligned} \tag{24}$$

The expressions in (23) and (24) are to be used to estimate the size of the contributions of current and lagged promotions to sales.

4 Application

In this section we investigate to what extent the newly developed econometric model is useful to describe weekly scanner data. We take the following strategy. First, we estimate the model for a large number of data sets. Next we highlight the results for two specific cases.

We have 36 A.C. Nielsen data sets containing weekly sales on brands in product categories such as peanut butter, canned tuna, toilet tissues, and ketchup. These data sets all concern national brands. We also have actually paid prices for each week. Additionally, in each of the 36 cases, we have 0/1 dummy variables for a feature promotion and for a display promotion. All data sets cover 124 weeks in the years 1986-1988. Full details on these data sets can be obtained from the authors. We transform the sales and price data by taking natural logarithms, and denote the resultant variables as y_t and p_t. The promotion variable are denoted as f_t en d_t, respectively.

Some preliminary experimentation involving the analysis of residual autocorrelations, model fit and normality of the residuals of linear models like (1) reveals that the following specification turns out the most useful, that is,

$$y_t = \mu + \rho y_{t-1} + \beta_1 \Delta p_t + v_t + w_t + \varepsilon_t \tag{25}$$

where $\Delta p_t = p_t - p_{t-1}$, $\varepsilon_t \sim \text{NID}(0, \sigma_\varepsilon^2)$ and where we take

$$v_t = \begin{cases} \gamma_{1,1} + \gamma_{1,2} f_t + \gamma_{1,3} d_t + u_{1,t} & \text{if } \gamma_{1,1} + \gamma_{1,2} f_t + \gamma_{1,3} d_t + u_{1,t} \geq 0 \\ 0 & \text{if } \gamma_{1,1} + \gamma_{1,2} f_t + \gamma_{1,3} d_t + u_{1,t} < 0 \end{cases} \tag{26}$$

and

$$w_t = \begin{cases} \gamma_{2,1} + \gamma_{2,2} f_{t-1} + \gamma_{2,3} d_{t-1} + u_{2,t} & \text{if } \gamma_{2,1} + \gamma_{2,2} f_{t-1} + \gamma_{2,3} d_{t-1} + u_{2,t} \leq 0 \\ 0 & \text{if } \gamma_{2,1} + \gamma_{2,2} f_{t-1} + \gamma_{2,3} d_{t-1} + u_{2,t} > 0 \end{cases} \tag{27}$$

with $u_{i,t} \sim \text{NID}(0, \sigma_{u_i}^2)$ for $i = 1, 2$. In words, our model assumes that log sales depends on one-week lagged log sales, on the rate of change of the price (Δp_t), and on two unobserved variables v_t and w_t, which take positive values if only current promotions have a large enough positive effect and which take negative values only if there is a large enough post-promotion dip, respectively.

To evaluate the potential usefulness of this model we compare it with a model which would appear if v_t is always positive and w_t is always negative, that is, the linear model

$$y_t = \mu + \alpha y_{t-1} + \beta_1 \Delta p_t + \beta_2 f_t + \beta_3 d_t + \beta_4 f_{t-1} + \beta_5 d_{t-1} + \varepsilon_t. \tag{28}$$

In this model the promotional activities affect sales (y_t) in a linear way. Note that the linear model (28) is not nested in the nonlinear model (25)-(27) and hence we cannot rely on standard likelihood ratio tests to compare both specifications. In order to compare the models, we therefore use the familiar Akaike and Schwarz model selection criteria, defined by

$$AIC = -2\frac{\ell(Y|Z, X; \hat{\theta})}{T} + \frac{2k}{T}$$

and

$$BIC = -2\frac{\ell(Y|Z, X; \hat{\theta})}{T} + \frac{k}{T}\log T,$$

respectively, where k denotes the number of parameters. The model with the minimum value on these criteria is to be preferred.

4.1 General Results

For each of the 36 cases, we estimate the general model in (25)-(27). In 12 of the 36 cases, we were not able to find a proper maximum of the likelihood function. As suggested before, this can mean that either the promotions have no effect at all or that the promotions always have an effect. A detailed look at some estimation results for the corresponding linear models suggests that the second cause appears most plausible.

For the 24 remaining cases, we were able to obtain sensible parameter estimates. However, in about 18 of the 24 cases, we find that the estimated standard errors of (some of) the parameters in the post-promotion dip component w_t are relative large. This suggests that the censored variable w_t contributes to the sales very infrequently or that the effect of one of the marketing-mix variable in w_t is completely dominated by the effect of the other. If we remove the post-promotion dip component from the model and hence consider models with only current effects, we obtain much better results. Also, when we leave out one of the promotion variables, as it seldom takes a value of 1, we can obtain quite some improvement. To save space, we do not want to pay specific attention to all these cases and model variations. We will only zoom in on two illustrative cases in the next subsection.

In sum, the overall conclusion from our analysis concerning 36 cases seems to be that our new econometric model can be useful in about two out of three cases. It must be stressed though that one may encounter large confidence intervals for some parameters, which is simply due to the fact that the censored variables appear in the sales equation very infrequently, and hence this should not be taken as evidence against the usefulness of the model as discussed above.

4.2 Two Specific Results

We now turn to a more detailed discussion of the estimation results for two particular cases. The first case concerns Kippy's peanut butter and the second concerns C&H sugar. The sales data (in dollars) are given in Figures 1 and 2, and it is clear that there can be sharp effects of promotions. Even after taking natural logs of these sales data, those pronounced peaks persist. The estimation results of the nonlinear model, that is, (25)-(27) and the linear model in (28) appear in Table 2. When we compare the values of the model selection criteria in this table, we see that the minimum values are obtained for the nonlinear model.

Insert Figures 1 and 2 about here.

Insert Table 2 about here.

The estimation results in Table 2 can be interpreted as follows. For Kippy we find for the nonlinear model that a feature promotion has an effect of size 0.906, at least, when it has an effect, whereas the linear model indicates an effect of 0.574. Hence, assuming that this promotion always has an effect and always the same seems to underestimate its effectiveness. A similar finding also holds for C&H, where the nonlinear model finds significant contribution of f_t and d_t of size 0.553 and 0.587, whereas the linear model only finds values of 0.253 and 0.449. For C&H we obtain also some interesting results concerning a post-promotion dip. The nonlinear model indicates no such dip for feature (an insignificant effect of 0.103), but substantial evidence of such a dip for display promotion (-0.484). In contrast, the linear model finds a post-promotion dip effect of -0.297 for display, but a contrasting effect of feature. Finally, notice that it holds for both brands that the variances in the censored equations are about 2.5 to 10 as large as the variance of ε_t. This suggests that there is substantial uncertainty about the effects of promotions, thereby indicating that the linear model underestimates the standard errors of the promotion parameters.

Insert Figures 3, 4, 5 and 6 about here.

As the C&H example gives the most significant results for the nonlinear model, we continue with this example. In Figures 3 and 4, we display for C&H the fitted values of v_t and w_t in each of their censored regressions. From Figure 3 it can be observed that there are several observations for which the fitted value of v_t is rather large, and the same holds for w_t in Figure 4, where large and negative values occur quite frequently. The conditional probability of a nonzero v_t is given in Figure 5, together with scaled dummy variables for feature and display. It can be observed that these probabilities can be large, even weeks after a promotion, thereby suggesting that sometimes promotions can have a longerlasting effect. This should entirely be due to the estimated residuals in the censored regressions. Furthermore, these probabilities clearly are not only non-zero in the case the dummies take a value of 1. The conditional probabilities of nonzero w_t in Figure 6 indicate the presence of many post-promotion dips for this particular brand. Again, it can be seen that it sometimes may take a while for this dip to become noticeable.

In Figures 5 and 6 we presented the conditional probabilities that there are effects of promotions, and these are not informative for the size of the effects. For the size, we should evaluate the expressions in (23) and (24) and the results are given in Figures 7 and 8. These graphs clearly show that the model also allows for time-variation in the size of the (post-)promotional effects.

5 Concluding Remarks

For managers it is important to understand the effects of their marketing-mix instruments on for example sales. It is widely conjectured that promotions do not always have a positive effect in the same week or a negative effect the week thereafter, and also that if promotions have an effect, it is not always of the same size. In this paper we translated these conjectures into a new and parsimonious econometric model. The model is rather flexible, and its boundary cases are linear models. The model parameters can easily and rather precisely be estimated using maximum likelihood. The model can lead to various interpretations concerning the probabilities and size of promotion effects, where all these aspects can vary substantially over time. The model should be very useful to managers as our empirical analysis indicated its potential usefulness in two out of three cases and as our detailed analysis of two specific cases showed that a linear model either underestimates or misspecifies the effects of promotions.

There are at least two limitations of our study, and each in turn suggests a fruitful area for further research. The first limitation concerns the model itself. We included in the model parts for the promotion effects only dummy variables for the promotion of the own brand. Naturally one would expect that there are competitive effects. Hence, one could extend the model by including a measure of promotional intensity in the entire product category. It can also be useful to consider a multivariate extension of the current model. A second limitation concerns the empirical analysis. We only looked at within-sample estimation results. Perhaps alternative insights could be gained by also looking at out-of-sample forecasting performance of the new nonlinear model.

A Appendix

In this appendix we show how the pdf in the log-likelihood function (13) can be expressed in terms of pdfs and cdfs of (bivariate) normal distributions.

The first term of the RHS of (12) can simply be written as

$$\Phi(-x_t'\gamma_1/\sigma_{u_1})(1-\Phi(-x_{t-1}'\gamma_2/\sigma_{u_2}))\frac{1}{\sigma_\varepsilon}\phi((y_t-\rho y_{t-1}-x_t'\beta)/\sigma_\varepsilon), \tag{29}$$

where we use (6) and (7). Upon using the expression for $\Pr[w_t = 0|x_{t-1};\theta]$, the second term of the RHS of (12) can be simplified by using

$$\begin{aligned}
&\int_{-x_t'\gamma_1}^{\infty}\frac{1}{\sigma_{u_1}}\phi(u_{1,t}/\sigma_{u_1})\frac{1}{\sigma_\varepsilon}\phi((u_{1,t}-c_1)/\sigma_\varepsilon)du_{1,t}\\
&\quad=\int_{-x_t'\gamma_1}^{\infty}\frac{1}{\sigma_u\sqrt{2\pi}}\exp\left(-\frac{1}{2}(u_{1,t}/\sigma_u)^2\right)\frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\exp\left(-\frac{1}{2}((u_{1,t}-c_1)/\sigma_\varepsilon)^2\right)du_{1,t} \qquad (30)\\
&\quad=\frac{\sigma_1}{\sigma_{u_1}\sigma_\varepsilon\sqrt{2\pi}}\exp\left(\frac{c_1^2}{2}(\sigma_1^2\sigma_\varepsilon^{-4}-\sigma_\varepsilon^{-2})\right)\int_{-x_t'\gamma_1}^{\infty}\frac{1}{\sigma_1}\phi((u_{1,t}-\sigma_\varepsilon^{-2}\sigma_1^2c_1)/\sigma_1)du_{1,t},
\end{aligned}$$

where $c_1 = y_t-\rho y_{t-1}-z_t'\beta-x_t'\gamma_1$ and $\sigma_1^2 = \frac{\sigma_{u_1}^2\sigma_\varepsilon^2}{\sigma_{u_1}^2+\sigma_\varepsilon^2}$. Hence, the second term can be written as

$$(1-\Phi(-x_{t-1}'\gamma_2/\sigma_{u_2}))\frac{\sigma_1}{\sigma_{u_1}\sigma_\varepsilon\sqrt{2\pi}}\exp\left(\frac{c_1^2}{2\sigma_\varepsilon^2}(\sigma_1^2/\sigma_\varepsilon^2-1))\right)\Phi((\sigma_\varepsilon^{-2}\sigma_1^2c_1+x_t'\gamma_1)/\sigma_1). \tag{31}$$

In a similar way the third term of the RHS of (12) can be written as

$$\Phi(-x_t'\gamma_1/\sigma_{u_1})\frac{\sigma_2}{\sigma_{u_2}\sigma_\varepsilon\sqrt{2\pi}}\exp\left(\frac{c_2^2}{2\sigma_\varepsilon^2}(\sigma_2^2/\sigma_\varepsilon^2-1)\right)(1-\Phi((\sigma_\varepsilon^{-2}\sigma_2^2c_2+x_{t-1}'\gamma_2)/\sigma_2)), \tag{32}$$

where $c_2 = y_t-\rho y_{t-1}-z_t'\beta-x_{t-1}'\gamma_2$ and $\sigma_2^2 = \frac{\sigma_{u_2}^2\sigma_\varepsilon^2}{\sigma_{u_2}^2+\sigma_\varepsilon^2}$.

The final and fourth term of the RHS in (12) is slightly more difficult to handle. We first write

$$\frac{\phi(u_{1,t}/\sigma_{u_1})}{\sigma_{u_1}}\frac{\phi(u_{2,t}/\sigma_{u_2})}{\sigma_{u_2}} = \frac{1}{2\pi}|\Sigma_u|^{-\frac{1}{2}}\exp(-\frac{1}{2}u_t'\Sigma_u^{-1}u_t) \tag{33}$$

with $\Sigma_u = \text{diag}(\sigma_{u_1}^2,\sigma_{u_2}^2)$ and $u_t = (u_{1,t},u_{2,t})'$ and

$$\begin{aligned}
f(y_t|y_{t-1},v_t,w_t;\theta)|_{v_t=x_t'\gamma_1+u_{1,t},w_t=x_{t-1}'\gamma_2+u_{2,t}} &= \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\exp(-\frac{1}{2}(c_3-u_{1,t}-u_{2,t})^2)\\
&= \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\exp(-\frac{1}{2}(u_t-a)'A(u_t-a))
\end{aligned} \tag{34}$$

with $A = \sigma_\varepsilon^{-2} E$ with E a (2×2) matrix of ones, $c_3 = y_t - \rho y_{t-1} - z_t'\beta - x_t'\gamma_1 - x_{t-1}'\gamma_2$ and $a = (c_3, 0)'$. If we use the result in Paap & Franses (1999), that is,

$$\begin{aligned} u_t'\Sigma_u^{-1}u_t + (u_t - a)'A(u_t - a) &= u_t'\Sigma_u^{-1}u' + u_t'Au_t - a'Au_t - u_t'Aa + a'Aa \\ &= u_t'(\Sigma_u^{-1} + A)u_t + a'Au_t - u_t'Aa + a'Aa \\ &= (u_t - b)'(\Sigma_u^{-1} + A)(u_t - b) - b'(\Sigma_u^{-1} + A)b + a'Aa \end{aligned} \tag{35}$$

with $b = (\Sigma_u^{-1} + A)^{-1}Aa$, we can write the fourth term of the RHS of (12) as

$$\begin{aligned} &\int_{-\infty}^{-x_{t-1}'\gamma_2}\int_{-x_t'\gamma_1}^{\infty} \frac{1}{2\pi}|\Sigma_u|^{-\frac{1}{2}} \frac{1}{\sigma_\varepsilon\sqrt{2\pi}} \exp(-\frac{1}{2}(u_t'\Sigma_u^{-1}u_t + (u_t - a)'A(u_t - a))) = \\ &\qquad \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}|\Sigma_u|^{-\frac{1}{2}}|\Sigma_u^{-1} + A|^{-\frac{1}{2}} \exp(\frac{1}{2}(b'(\Sigma_u^{-1} + A)b - \frac{1}{2}a'Ba) \\ &\quad \int_{-\infty}^{-x_{t-1}'\gamma_2}\int_{-x_t'\gamma_1}^{\infty} \frac{1}{2\pi}|\Sigma_u^{-1} + A|^{\frac{1}{2}} \exp(-\frac{1}{2}(u_t - b)'(\Sigma_u^{-1} + A)(u_t - b))du_{1,t}du_{2,t}, \end{aligned} \tag{36}$$

where the latter part corresponds to the cdf of a bivariate normal distribution with mean b and covariance matrix $(\Sigma_u^{-1} + A)^{-1}$.

Table 1: Properties of the ML estimator[a]

parameter	true value	$E[\hat{\theta}]$	$E[(\hat{\theta}-\theta)^2]$	nominal size z-scores[b] left tail 0.01	left tail 0.05	left tail 0.10	right tail 0.10	right tail 0.05	right tail 0.01
				125 observations					
μ	0.50	0.49	0.056	0.04	0.08	0.13	0.16	0.11	0.06
ρ	0.50	0.49	0.006	0.01	0.05	0.09	0.14	0.07	0.02
σ_ε	0.50	0.42	0.022	0.00	0.02	0.04	0.27	0.21	0.12
$\gamma_{1,1}$	0.00	-0.02	0.181	0.06	0.11	0.15	0.04	0.01	0.00
$\gamma_{1,2}$	0.50	0.54	0.049	0.00	0.02	0.06	0.13	0.10	0.05
$\gamma_{2,1}$	0.00	0.10	0.243	0.00	0.01	0.02	0.12	0.08	0.05
$\gamma_{2,2}$	-0.50	-0.57	0.063	0.05	0.08	0.12	0.07	0.02	0.00
σ_{u_1}	0.50	0.51	0.048	0.02	0.07	0.12	0.01	0.00	0.00
σ_{u_2}	0.50	0.52	0.050	0.01	0.05	0.10	0.01	0.00	0.00
				250 observations					
μ	0.50	0.51	0.033	0.03	0.08	0.12	0.11	0.07	0.04
ρ	0.50	0.49	0.003	0.01	0.06	0.09	0.13	0.07	0.02
σ_ε	0.50	0.45	0.013	0.02	0.05	0.07	0.23	0.17	0.10
$\gamma_{1,1}$	0.00	0.01	0.097	0.06	0.12	0.17	0.06	0.02	0.00
$\gamma_{1,2}$	0.50	0.52	0.023	0.00	0.03	0.09	0.16	0.12	0.07
$\gamma_{2,1}$	0.00	-0.00	0.085	0.00	0.02	0.05	0.14	0.09	0.05
$\gamma_{2,2}$	-0.50	-0.51	0.021	0.05	0.11	0.16	0.07	0.03	0.01
σ_{u_1}	0.50	0.50	0.024	0.02	0.09	0.15	0.07	0.03	0.02
σ_{u_2}	0.50	0.50	0.027	0.01	0.08	0.15	0.08	0.04	0.01
				1000 observations					
μ	0.50	0.50	0.005	0.01	0.05	0.10	0.10	0.06	0.01
ρ	0.50	0.50	0.001	0.00	0.04	0.09	0.12	0.06	0.01
σ_ε	0.50	0.50	0.002	0.01	0.05	0.09	0.11	0.06	0.02
$\gamma_{1,1}$	0.00	-0.00	0.018	0.02	0.08	0.13	0.10	0.04	0.01
$\gamma_{1,2}$	0.50	0.51	0.005	0.01	0.04	0.08	0.13	0.07	0.02
$\gamma_{2,1}$	0.00	0.00	0.017	0.00	0.04	0.09	0.12	0.06	0.02
$\gamma_{2,2}$	-0.50	-0.50	0.005	0.02	0.07	0.12	0.09	0.04	0.00
σ_{u_1}	0.50	0.50	0.006	0.01	0.05	0.10	0.09	0.05	0.01
σ_{u_2}	0.50	0.49	0.005	0.01	0.04	0.08	0.09	0.04	0.00

[a] The DGP is given in (14)-(16). The number of replications is 1000.

[b] The z-scores are defined as $(\hat{\theta}-\theta)/\hat{\sigma}(\hat{\theta})$, where $\hat{\sigma}(\hat{\theta})$ denotes the estimated standard error of $\hat{\theta}$. The cells report the empirical size.

Table 2: Estimation results for Kippy and C&H.

	Kippy				C&H			
	nonlinear[a]		linear[b]		nonlinear[a]		linear[b]	
parameter	$\hat{\theta}$[c]	$\hat{\sigma}(\hat{\theta})$[d]	$\hat{\theta}$[c]	$\hat{\sigma}(\hat{\theta})$	$\hat{\theta}$[c]	$\hat{\sigma}(\hat{\theta})$	$\hat{\theta}$[c]	$\hat{\sigma}(\hat{\theta})$
μ	2.310**	0.237	1.919**	0.336	2.163**	0.158	2.749**	0.356
ρ	0.443**	0.052	0.507**	0.081	0.472**	0.039	0.312**	0.088
β_1	-4.047**	1.039	-3.264**	1.240	-6.197**	0.585	-5.877**	0.768
β_2			0.574**	0.230			0.253**	0.101
β_3			0.272	0.231			0.449**	0.135
β_4			-0.077	0.226			0.266**	0.109
β_5			-0.105	0.233			-0.297**	0.139
σ_ε	0.327**	0.070			0.038**	0.010		
$\gamma_{1,1}$	-0.593	0.600			-0.166*	0.095		
$\gamma_{1,2}$	0.906*	0.463			0.553**	0.124		
$\gamma_{1,3}$	0.205	0.359			0.587**	0.152		
$\gamma_{2,1}$	0.626	0.572			0.015	0.089		
$\gamma_{2,2}$	-0.446	0.760			0.103	0.127		
$\gamma_{2,3}$	0.987	0.677			-0.484**	0.201		
σ_{u_1}	1.000**	0.270			0.440**	0.061		
σ_{u_2}	1.427**	0.443			0.421**	0.066		
AIC	2.374		2.485		0.848		1.127	
BIC	2.647		2.645		1.121		1.287	
$\ell(Y\|Z,X;\hat{\theta})$	-133.998		-145.809		-40.138		-62.300	

[a] The model is given in (25)-(27).
[b] The model is given in (28).
[c] Maximum likelihood estimate. ** and * denote significant at 5% and 10 % level, respectively.
[d] Estimated standard error.

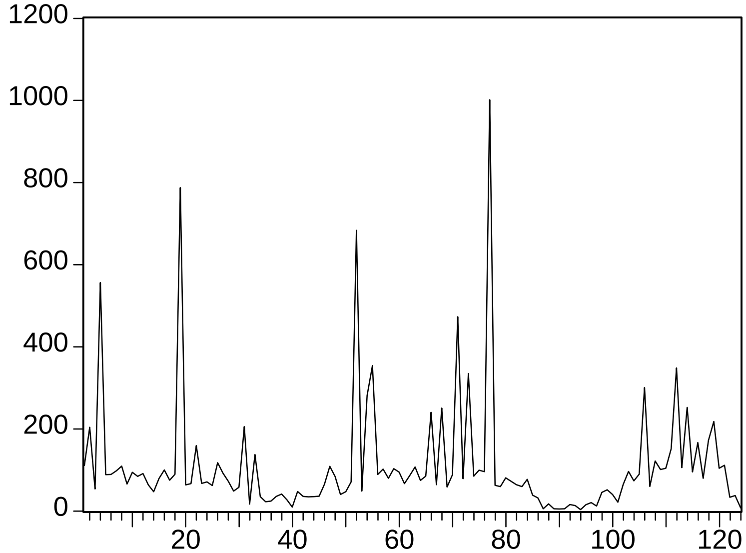

Figure 1: Weekly sales ($T = 124$) of Kippy Peanut Butter (in dollars)

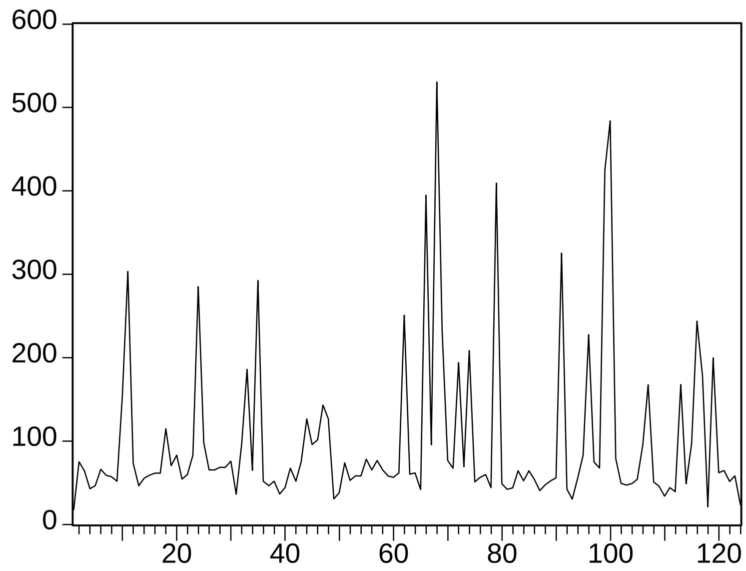

Figure 2: Weekly sales ($T = 124$) of C&H sugar (in dollars)

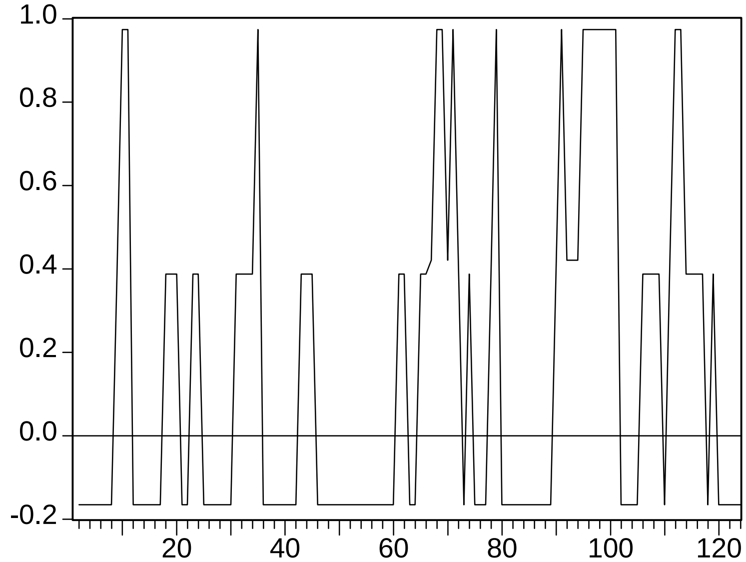

Figure 3: Fitted values of v_t in the censored latent regression $(x_t'\hat{\gamma}_1)$ for C&H

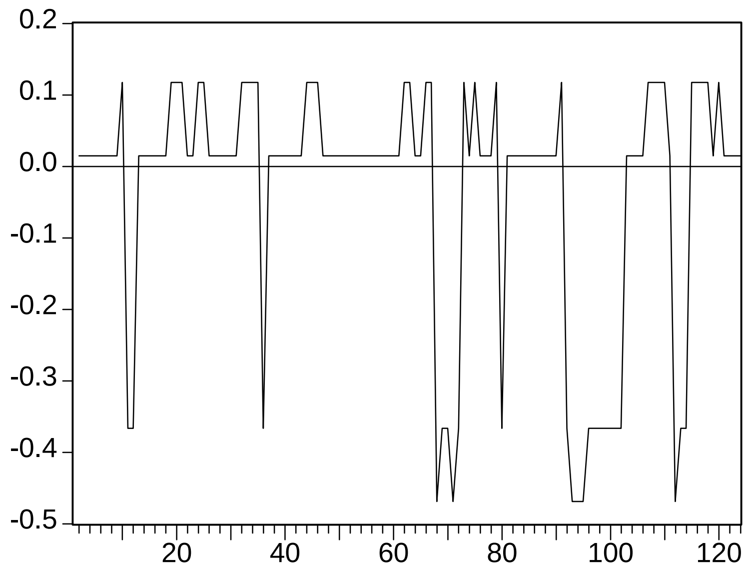

Figure 4: Fitted values of w_t in the censored latent regression $(x_{t-1}'\hat{\gamma}_2)$ for C&H

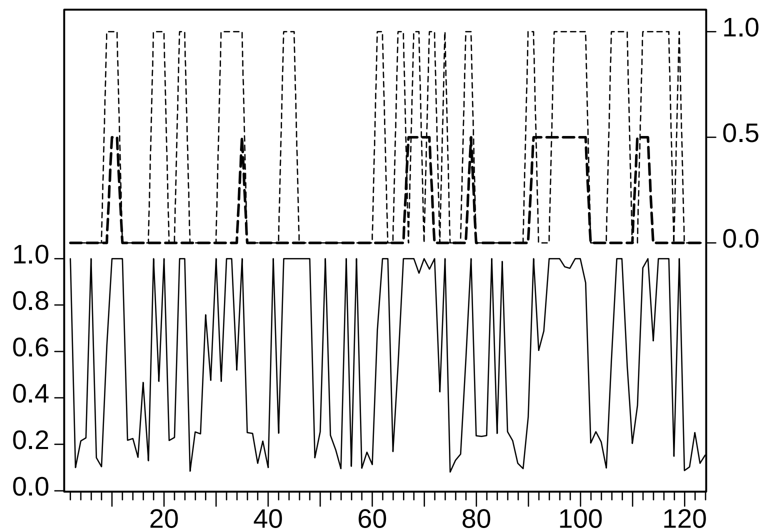

Figure 5: The conditional probability $\Pr[v_t > 0|y_t, y_{t-1}, x_t, x_{t-1}; \hat{\theta}]$ (straight line), and the display (short dashes) and scaled feature dummies (bold long dashes) for C&H

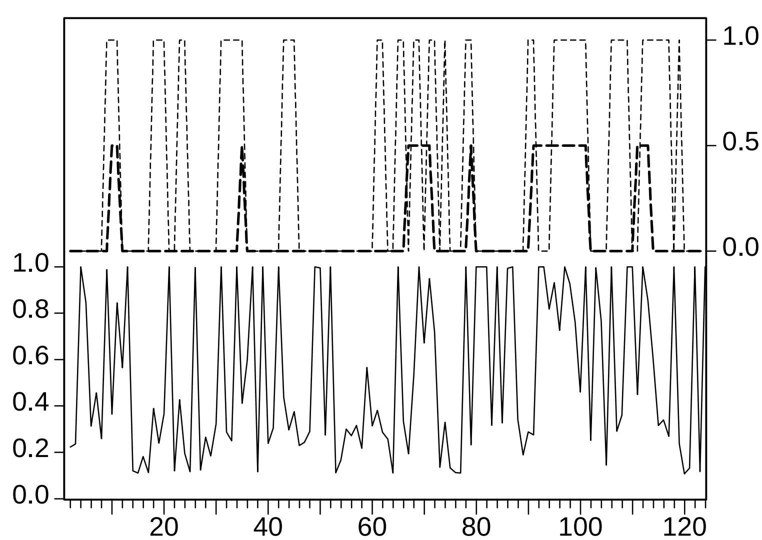

Figure 6: The conditional probability $\Pr[w_t < 0|y_t, y_{t-1}, x_t, x_{t-1}; \hat{\theta}]$ (straight line), and the display (short dashes) and scaled feature dummies (bold long dashes) for C&H

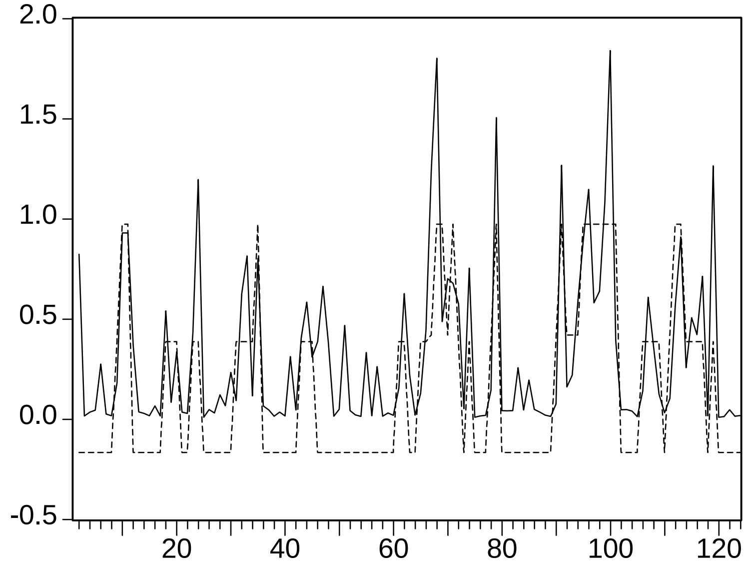

Figure 7: The conditional expectation $\mathrm{E}[v_t|y_t, y_{t-1}, x_t, x_{t-1}; \hat{\theta}]$ (straight line), and $x_t'\hat{\gamma}_1$ (short dashes) for C&H

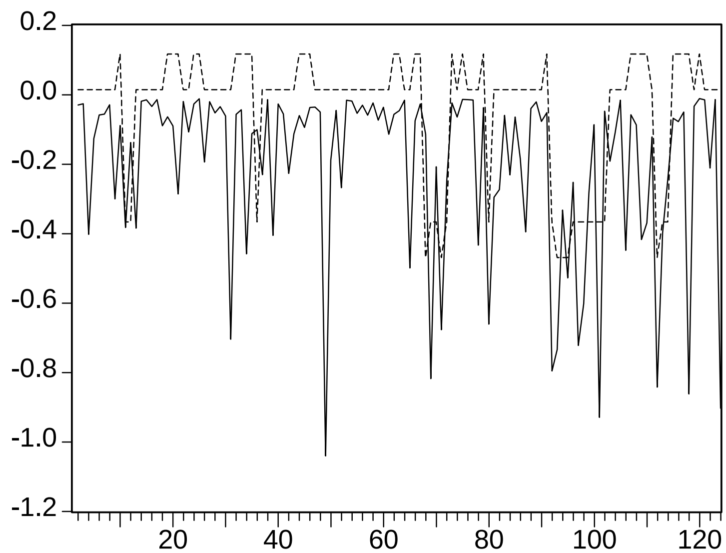

Figure 8: The conditional expectation $\mathrm{E}[w_t|y_t, y_{t-1}, x_t, x_{t-1}; \hat{\theta}]$ (straight line), and $x_{t-1}'\hat{\gamma}_2$ (short dashes) for C&H

References

Blattberg, Robert C., Richard Briesch and Edward J. Fox (1995), How Promotions Work, *Marketing Science*, 14, G122-G132.

Blattberg, Robert C. and Scott A. Neslin (1990), *Sales Promotion, Concepts, Methods and Strategies*, Englewood Cliffs NJ: Prentice Hall.

Dekimpe, Marnik G., and Dominique M. Hanssens (1995a),The Persistence of Marketing Effects on Sales, *Marketing Science*, 14, 1-21.

Dekimpe, Marnik G., and Dominique M. Hanssens (1995b), Empirical Generalizations about Market Evolution and Stationarity, *Marketing Science*, 14, G109-G121.

Foekens, Eijte W., Peter S.H. Leeflang and Dick R. Wittink (1999), Varying Parameter Models to Accommodate Dynamic Promotion Effects, *Journal of Econometrics*, 89, 249-268.

Gourieroux, Christian and Alain Monfort (1995), *Statistics and Econometric Models*, vol. 2, Cambridge: Cambridge University Press.

Gupta, Sunil (1988), Impact of Sales Promotions on When, What, and How Much to Buy, *Journal of Marketing Research*, 25, 342-355.

Kumar, V. and Robert P. Leone (1988), Measuring the Effect of Retail Store Promotion on Brand and Store Substitution, *Journal of Marketing Research*, 25, 178-185.

Leeflang, Peter S.H., Dick R. Wittink, Michel Wedel and Philippe A. Naert (2000), *Building Models for Marketing Decisions*, International Series in Quantitative Marketing, Boston: Kluwer Academic Publishers.

Maddala, George. S. (1983), *Limited-Dependent and Qualitative Variables in Econo-*

metrics, Econometric Society Monographs No. 3, Cambridge: Cambridge University Press.

Moriarty, Mark M. (1985), Retail Promotional Effects on Intra- and Interbrand Sales Performance, *Journal of Retailing*, 61, 27-47.

Neslin, Scott A, and Linda G. Schneider-Stone (1996), Consumer Inventory Sensitivity and the Postpromotion Dip, *Marketing Letters*, 7, 77-94.

Paap, Richard and Philip Hans Franses (1999), Do the US and Canada Have a Common Nonlinear Cycle in Unemployment?, *Econometric Institute Report*, 9907/A, Erasmus University Rotterdam.

Paap, Richard and Philip Hans Franses (2000), A Dynamic Multinomial Probit Model for Brand Choice with Different Long-Run and Short-Run Effects of Marketing-Mix Variables, *Journal of Applied Econometrics*, 15, 717-744.

Papatla, Purushottam and Lakshman Krishnamurthi (1996), Measuring the Dynamic Effects of Promotions on Brand Choice, *Journal of Marketing Research*, 33, 20-35.

Rosenbaum, S., 1961, Moments of a Truncated Bivariate Normal Distribution, *Journal of the Royal Statistical Association B*, 23, 405-408.

Van Heerde, Harald J., Peter S.H. Leeflang and Dick R. Wittink (2000), The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data, *Journal of Marketing Research*, 37, 383-395.

Wittink, Dick R., Michael J. Addona, William J. Hawkes, and John C. Porter (1988), SCAN*PRO, the Estimation, Validation, and Use of Promotional Effects Based on Scanner Data, Working paper, Johnson Graduate School of Management, Cornell University.

Publications in the Report Series Research[*] in Management

ERIM Research Program: "Decision Making in Marketing Management"

2001
Predicting Customer Potential Value. An application in the insurance industry
Peter C. Verhoef & Bas Donkers
ERS-2001-01-MKT

Modeling Potenitally Time-Varying Effects of Promotions on Sales
Philip Hans Franses, Richard Paap & Philip A. Sijthoff
ERS-2001-05-MKT

2000
Impact of the Employee Communication and Perceived External Prestige on Organizational Identification
Ale Smidts, Cees B.M. van Riel & Ad Th.H. Pruyn
ERS-2000-01-MKT

Forecasting Market Shares from Models for Sales
Dennis Fok & Philip Hans Franses
ERS-2000-03-MKT

The Effect of Relational Constructs on Relationship Performance: Does Duration Matter?
Peter C. Verhoef, Philip Hans Franses & Janny C. Hoekstra
ERS-2000-08-MKT

Informants in Organizational Marketing Research: How Many, Who, and How to Aggregate Response?
Gerrit H. van Bruggen, Gary L. Lilien & Manish Kacker
ERS-2000-32-MKT

The Powerful Triangle of Marketing Data, Managerial Judgment, and Marketing Management Support Systems
Gerrit H. van Bruggen, Ale Smidts & Berend Wierenga
ERS-2000-33-MKT

Consumer Perception and Evaluation of Waiting Time: A Field Experiment
Gerrit Antonides, Peter C. Verhoef & Marcel van Aalst
ERS-2000-35-MKT

Broker Positions in Task-Specific Knowledge Networks: Effects on Perceived Performance and Role Stressors in an Account Management System
David Dekker, Frans Stokman & Philip Hans Franses
ERS-2000-37-MKT

Modeling Unobserved Consideration Sets for Household Panel Data
Erjen van Nierop, Richard Paap, Bart Bronnenberg, Philip Hans Franses & Michel Wedel
ERS-2000-42-MKT

[*] A complete overview of the ERIM Report Series Research in Management:
www.erim.eur.nl, under the heading 'Publications'

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Decision Making in Marketing Management
F&A Financial Decision Making and Accounting
STR Strategic Renewal and the Dynamics of Firms, Networks and Industries

A Managerial Perspective on the Logic of Increasing Returns
Erik den Hartigh, Fred Langerak & Harry Commandeur
ERS-2000-48-MKT

The Mediating Effect of NPD-Activities and NPD-Performance on the Relationship between Market Orientation and Organizational Performance
Fred Langerak, Erik Jan Hultink & Henry S.J. Robben
ERS-2000-50-MKT

Sensemaking from actions: Deriving organization members' means and ends from their day-to-day behavior
Johan van Rekom, Cees B.M. van Riel & Berend Wierenga
ERS-2000-52-MKT