Chapter 11
Time-Series Models in Marketing

Marnik G. Dekimpe, Philip Hans Franses, Dominique M. Hanssens, and Prasad A. Naik

11.1 Introduction

Marketing data appear in a variety of forms. A frequently occurring form is time-series data, for example, sales per week, market shares per month, the price evolution over the last few years, or historically-observed advertising-spending patterns. The main feature of time-series data is that the observations are ordered over time, and hence that it is likely that earlier observations have predictive content for future observations. Indeed, if relative prices are, say, 1.50 today, they most likely will be around 1.50 tomorrow too, or in any case, not a value of 120.

Time series can refer to a single variable, such as sales or advertising, but can also cover a vector of variables, for example sales, prices and advertising, jointly. In some instances, marketing modelers may want to build a univariate model for a time series, and analyze the series strictly as a function of its own past. This is, for example, the case when one has to forecast (or extrapolate) exogenous variables, or when the number of variables to be analyzed (e.g. the number of items in a broad assortment) is so large that building multivariate models for each of them is too unwieldy (Hanssens, Parsons and Schultz 2001). However, univariate time-series models do not handle the cause-and-effect situations that are central to marketing planning. To specify the lag structure in response models, one extends the techniques of univariate extrapolation to the case of multiple time series.

Time-series data can be summarized in time-series models. However, not all models built on time-series data are referred to as time-series models. Unlike most econometric approaches to dynamic model specification, time-series

M.G. Dekimpe
Tilburg University, Tilburg, The Netherlands and Catholic University Leuven, Leuven, Belgium
e-mail: M.G.Dekimpe@uvt.nl

The authors are indebted to Scott Neslin, Vincent Nijs and the editor, Berend Wierenga, for useful comments on an earlier version of the chapter.

B. Wierenga (ed.), *Handbook of Marketing Decision Models*,
DOI: 10.1007/978-0-387-78213-3_11, © Springer Science+Business Media, LLC 2008

modelers take a more data-driven approach. Specifically, one looks at historically-observed patterns in the data to help in model specification, rather than imposing a priori a certain structure derived from marketing or economic theory on the data. As put by Nobel Laureate Sir Clive Granger (1981, p. 121):

> It is well known that time-series analysts have a rather different approach to the analysis of economic data than does the remainder of the econometric profession. One aspect of this difference is that we admit more readily to looking at the data before finally specifying the model; in fact, we greatly encourage looking at the data. Although econometricians trained in a more traditional manner are still very much inhibited in the use of summary statistics derived from the data to help model selection, or identification, it could be to their advantage to change some of these attitudes.

This feature of looking at the data to help in model specification can be illustrated as follows. Given a hypothesized model for a time series, one can derive the properties of empirical data in case that model would truly describe the data. For example, a simple model that says that y_t only depends on y_{t-1} using the scheme $y_t = \rho\, y_{t-1} + e_t$ would imply that y_t shows a correlation with y_{t-1} of size ρ, with y_{t-2} of size ρ^2, and so on. If such a correlation structure were to be found in empirical data, one would have a first guess at what the best descriptive model could look like. If $|\rho| < 1$, the impact of past events becomes smaller and smaller, which is not the case when $\rho = 1$ (the so-called unit-root scenario, discussed in more detail in Section 11.2). A competing model with structure $y_t = e_t - \theta\, e_{t-1}$ would show a non-zero correlation between y_t and y_{t-1}, and a zero correlation between y_t and, respectively, $y_{t-2}, y_{t-3}, \ldots$ By comparing the empirically-observed correlation patterns (referred to as the empirical autocorrelation function) with the one associated theoretically with a given model structure, a model is selected that is likely to have generated the data. Other summary statistics that are useful in this respect are the partial autocorrelation function and (in case of multiple variables) the cross-correlation function (see e.g. Hanssens et al. 2001 for a review). While time-series modelers highly stimulate this "looking at the data", critics refer to this practice as data-mining, arguing that time-series models "lack foundations in marketing theory" (Leeflang et al. 2000, p. 458).

This criticism is one of the reasons why, historically, time-series models were not used that often in the marketing literature. Other reasons, described in detail in Dekimpe and Hanssens (2000), were (i) marketing scientists' traditional lack of training in time-series methods, (ii) the lack of access to user-friendly software, (iii) the absence of good-quality time-series data, and (iv) the absence of a substantive marketing area where time-series modeling was adopted as primary research tool. However, over the last few years, these inhibiting factors have begun to disappear. Several marketing-modeling textbooks now contain chapters outlining the use of time-series models (see e.g Hanssens et al. 2001; Leeflang et al. 2000), while others include an overview chapter on time-series applications in marketing (see e.g. the current volume, or Moorman and Lehmann 2004). In terms of software, several user-friendly PC-based packages have become available (e.g. Eviews), while new data sources (e.g. long series of scanner data) have considerably alleviated the data concern. In terms of the substantive marketing area, several time-series techniques have been specifically designed to disentangle short- from long-run relationships. This fits well with one of marketing's main fields of interest: to quantify the long-run impact of marketing's tactical and strategic decisions. In terms of the critique on the a-theoretic character of time-series modeling, we observe three recent developments. First, the choice of which endogenous and exogenous variables to include in the VARX (*V*ector-*A*uto*R*egressive models with *e*Xogenous variables) models is increasingly theory-driven. Second, some time-series techniques (e.g. cointegration testing for theoretically-expected equilibria) have a more confirmatory potential. Finally, following a 1995 special issue of *Marketing Science*, there is growing recognition of the value of Empirical Generalizations obtained through the repeated application of data-driven techniques on multiple data sets. We refer to Dekimpe and Hanssens (2000) for an in-depth discussion on these issues. Because of these developments, time-series models have become increasingly accepted in the marketing literature. Moreover, we see an increasing convergence between regression approaches (which often focused on obtaining unbiased estimates of marketing-mix effectiveness, but did not rely much on summary statistics derived from the data at hand to help in model specification), time-series techniques (which were used primarily to derive good forecasts or extrapolations), and structural models (which start from economic fundamentals but ignored dynamics until recently).

Time-series modelers make use of a wide array of techniques, which are discussed in detail in textbooks such as Hamilton (1994) or Franses (1998), among others. In this chapter, we will not attempt to review all of these techniques. Instead, we will focus on two domains that have recently received considerable attention in the marketing literature: (i) the use of persistence modeling to make long-run inferences (Section 11.2), and (ii) the use of state-space models and their integration with normative decision making (Section 11.3). Finally, we will discuss a number of opportunities and challenges for time-series modelers in marketing (Section 11.4).

11.2 Persistence Modeling

Long-run market response is a central concern of any marketing strategy that tries to create a sustainable competitive advantage. However, this is easier said than done, as only short-run results of marketing actions are readily available. Persistence modeling addresses the problem of long-run market-response identification by combining into one metric the net long-run impact of a chain reaction of consumer response, firm feedback, and competitor response that emerges following an initial marketing action. This marketing action could be an unexpected increase in advertising support (e.g. Dekimpe and Hanssens 1995a), a price promotion (e.g. Pauwels, Hanssens, and Siddarth 2002), or a competitive activity (e.g. Steenkamp et al. 2005), and the performance metric

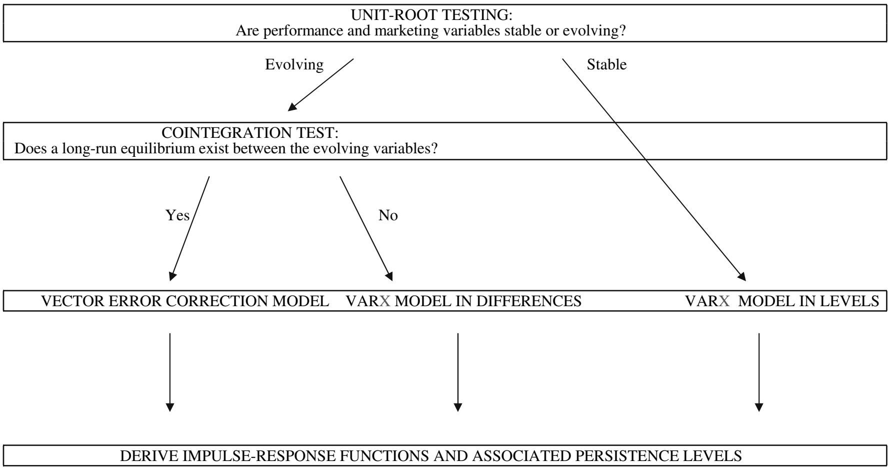

Fig. 11.1 Overview of persistence modeling procedure

could be primary (Nijs et al. 2001) or secondary (Dekimpe and Hanssens 1995a) demand, profitability (Dekimpe and Hanssens 1999), or stock prices (Pauwels, Silva-Risso, Srinivasan and Hanssens 2004), among others.

Persistence modeling is a multi-step process, as depicted in Fig. 11.1 (taken from Dekimpe and Hanssens 2004). In a first step, one applies unit-root tests to the different performance and marketing-support variables of interest to determine whether they are stable (mean or trend-stationary) or evolving. In the latter case, the series have a stochastic trend, and one has to test whether a long-run equilibrium exists between them. This is done through cointegration testing. Depending on the outcome of these preliminary (unit-root and cointegration) tests, one specifies a VARX model in the levels, a VARX model in the differences, or a Vector Error Correction Model. From these VARX models, one can derive impulse-response functions (IRFs), which trace the incremental effect of a one-unit (or one-standard-deviation) shock in one of the variables on the future values of the other endogenous variables.

Without going into mathematical details,[1] we can graphically illustrate the key concepts of the approach in Fig. 11.2 (taken from Nijs et al. 2001):

In this Figure, we depict the *incremental* primary demand that can be attributed to an initial price promotion. In the stable detergent market of Panel A, one observes an immediate sales increase, followed by a post-promotional dip. After some fluctuations, which can be attributed to factors such as purchase reinforcement, feedback rules, and competitive reactions, we observe that the incremental sales converge to zero. This does not imply that no more detergents are sold in this market, but rather that in the long run no additional sales can be attributed to

[1] We refer to Enders (1995) or Franses (1998) for excellent technical discussions on the various tests involved. Dekimpe and Hanssens (2004) review key decisions to be made in this respect.

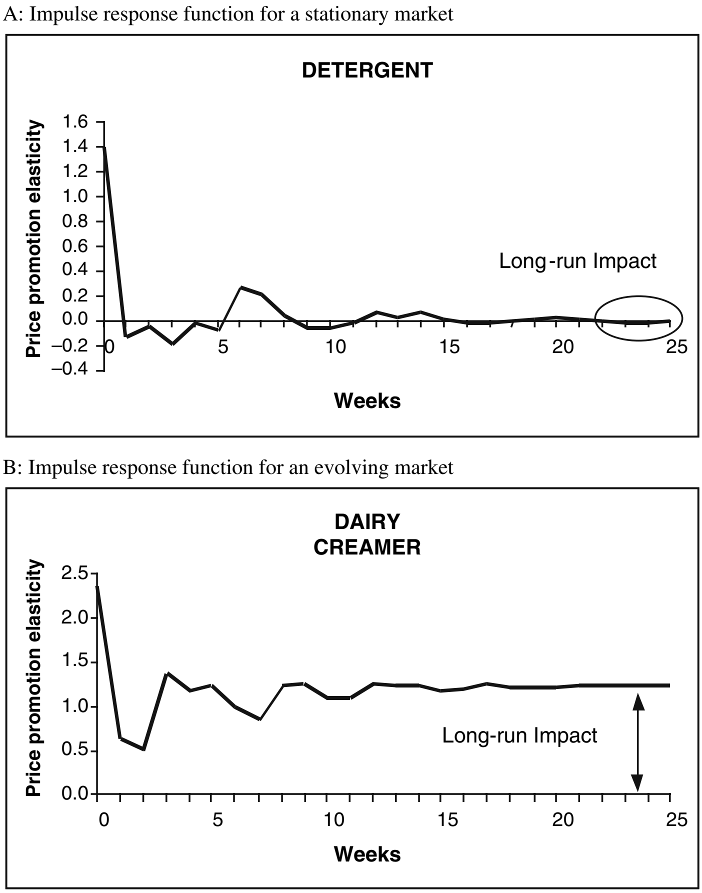

Fig. 11.2 Impulse response functions

the initial promotion. In contrast, in the evolving dairy-creamer market depicted in the bottom panel of Fig. 11.2, we see that this incremental effect stabilizes at a non-zero, or persistent, level. In that case, a long-run effect has been identified, as the initial promotion keeps on generating extra sales. This could be due to new customers who have been attracted to the category by the initial promotion and now make repeat purchases. Alternatively, existing customers may have increased their product-usage rates. From these impulse-response functions, one can derive various summary statistics, such as:

(i) the immediate performance impact of the price promotion;
(ii) the long-run or permanent (persistent) impact, i.e., the value to which the impulse-response function converges; and
(iii) the combined cumulative effect over the dust-settling period. This period is defined as the time it takes before the convergence level is obtained. For the

Figure in panel A, for example, the total effect over the dust-settling period (also referred to as the short-run effect) amounts to the area under the curve (specifically, the sum of the IRF estimates that have not yet converged to zero).

Persistence modeling offers two distinct advantages. First, it offers a clear and quantifiable distinction between short- and long-run promotional effectiveness, based on the difference between temporary and permanent movements in the data. Second, it uses a system's approach to market response, in that it combines the forces of customer response, competitive reaction, and firm decision rules. Indeed, the chain reaction of all these forces is reflected in the impulse-response functions, which are themselves derived from the multi-equation vector-autoregressive model. By incorporating such a chain reaction over time, the impulse-response function expands upon the more conventional direct & instantaneous elasticity estimates.[2]

Persistence modeling has been used extensively in the recent marketing literature, and has resulted in several strategic insights. We summarize these insights in Table 11.1, which updates Dekimpe and Hanssens (2004).

Many of these insights have been derived in a two-step modeling approach. In a first step, the procedure described in Fig. 11.1 is applied to multiple brands and/or product categories (see e.g. Nijs et al. 2001; Srinivasan et al. 2004; Steenkamp et al. 2005). In a second step, one explains the observed variability across brands or product categories in the aforementioned summary statistics (i.e. the immediate effect, the long-run effect and the dust-settling effect) through a variety of marketing-theory-based covariates.[3] These could include, for example, the advertising intensity or concentration rate in the category, or the strength and nature (private label or national brand) of the brand. However, this approach was recently criticized in Fok et al. (2006) for not appropriately accounting for the uncertainty in the first-stage parameter estimates when estimating the second-stage model. They therefore proposed a single-step Hierarchical Bayes Error Correction Model. As an added benefit, their approach offers direct estimates of a marketing instrument's short- and long-run effects. This is more parsimonious than through the aforementioned summary statistics, which are a function of many VARX parameters. A similar Error Correction Model was used in van Heerde, Helsen, and Dekimpe (2007), who investigated how short- and long-run price and advertising elasticities changed following a product-harm crisis. Both studies used a single-equation approach, however, treating all marketing-mix variables as exogenous. VARX models, in contrast, allow many of these variables to be endogenous.

---

[2] From these impulse-response functions, one can also derive a Forecast Error Variance Decomposition (FEVD) to calculate what percentage of the variation in an endogenous variable (e.g. retail price) can be attributed to contemporaneous and past changes in each of the endogenous variables (e.g. competing prices) in the system. We refer to Hanssens (1998) or Nijs et al. (2006) for an in-depth discussion on FEVD.

[3] This again helps to alleviate the criticism of being a-theoretical.

Table 11.1 Strategic insights from persistence modeling

Study	Contribution
Baghestani (1991)	Advertising has a long run impact on sales if both variables are (a) evolving and (b) in long-run equilibrium (cointegrated).
Bronnenberg, Mahajan, and Vanhonacker (2000)	Distribution coverage drives long-run market shares, especially the coverage evolution early in the life cycle.
Cavaliere and Tassinari (2001)	Advertising is not a long-run driver of aggregate whisky consumption in Italy.
Chowdhury (1994)	No long run equilibrium (cointegration) relationship is found between UK aggregate advertising spending and a variety of macro-economic variables.
Dekimpe and Hanssens (1995a)	Persistence measures quantify marketing's long-run effectiveness. Image-oriented and price-oriented advertising messages have a differential short- and long-run effect.
Dekimpe and Hanssens (1995b)	Sales series are mostly evolving, while a majority of market-share series is stationary.
Dekimpe and Hanssens (1999)	Different strategic scenarios (business as usual, escalation, hysteresis and evolving business practice) have different long-run profitability implications.
Dekimpe, Hanssens, and Silva-Risso (1999)	Little evidence of long-run promotional effects is found in FPCG markets.
Dekimpe et al. (1997)	New product introductions may cause structural breaks in otherwise stationary loyalty patterns.
Franses (1994)	Gompertz growth models with non-constant market potential can be written in error-correction format.
Franses, Kloek, and Lucas (1999)	Outlier-robust unit-root and cointegration tests are called for in promotion-intensive scanner environments.
Franses, Srinivasan, and Boswijk (2001)	Unit root and cointegration tests which account for the logical consistency of market shares.
Hanssens (1998)	Factory orders and sales are in a long-run equilibrium, but shocks to either have different long-run consequences.
Hanssens and Ouyang (2002)	Derivation of advertising allocation rules (in terms of triggering versus maintenance spending) under hysteresis conditions.
Horváth et al. (2005)	The inclusion/exclusion of competitive reaction and feedback effects affects the net unit sales effects of price reductions, as do intrafirm effects.
Horváth, Leeflang, and Otter (2002)	Structural relationships between (lagged) consumer response and (lagged) marketing instruments can be inferred through canonical correlation analysis and Wiener-Granger causality testing.
Johnson et al. (1992)	The long-run consumption of alcoholic beverages is not price sensitive.
Joshi and Hanssens (2006)	Advertising has a long-run positive effect on firm valuation.
Jung and Seldon (1995)	Aggregate US advertising spending is in long-run equilibrium with aggregate personal consumption expenditures.

Table 11.1 (continued)

Study	Contribution
Lim, Currim, and Andrews (2005)	Consumer segmentation matters in persistence modeling for price-promotion effectiveness.
McCullough and Waldon (1998)	Network and national spot advertising are substitutes.
Nijs et al. (2001)	Limited long-run category expansion effects of price promotions. The impact differs in terms of the marketing intensity, competitive structure, and competitive conduct in the industry.
Nijs, Srinivasan, and Pauwels (2006)	Retail prices are driven by pricing history, competitive retailer prices, brand demand, wholesale prices, and retailer category management considerations.
Pauwels (2004)	Restricted policy simulations allow to distinguish four dynamic forces that drive long-term marketing effectiveness: consumer response, competitor response, company inertia and company support.
Pauwels and Srinivasan (2004)	Permanent performance effects are observed from store brand entry, but these effects differ between manufacturers and retailers, and between premium-price and second-tier national brands.
Pauwels and Hanssens (2007)	Brands in mature markets go through different performance regimes, which are influenced by their marketing policies.
Pauwels et al. (2002)	The decomposition of the promotional sales spike in category-incidence, brand-switching and purchase-quantity effects differs depending on the time frame considered (short versus long run).
Pauwels et al. (2004)	Investor markets reward product innovation but punish promotional initiatives by automobile manufacturers.
Srinivasan and Bass (2000)	Stable market shares are consistent with evolving sales if brand and category sales are cointegrated.
Srinivasan, Popkowski Leszczyc, and Bass (2000)	Temporary, gradual and structural price changes have a different impact on market shares.
Srinivasan et al. (2004)	Price promotions have a differential performance impact for retailers versus manufacturers.
Steenkamp et al.(2005)	Competitive reactions to promotion and advertising attacks are often passive. This rarely involves a missed sales opportunity. If reaction occurs, it often involves spoiled arms.
Villanueva, Yoo, and Hanssens (2006)	Customers acquired through different channels have different lifetime values.
Zanias (1994)	Feedback effects occur between sales and advertising. The importance of cointegration analysis is demonstrated with respect to Granger causality testing and multi-step forecasting.

As indicated before, persistence and error-correction models have resulted in several empirical generalizations on the presence/absence of long-run marketing effects. However, these insights have remained largely descriptive. While some studies (see e.g. Pauwels 2004; van Heerde et al. 2007) have used these models for policy simulations,[4] their use for normative decision-making has remained the exception rather than the rule, and remains an important challenge for time-series modelers. The linkage with normative decision making has been made explicitly in recent applications of state-space modeling, which we review in Section 11.3. We offer somewhat more technical detail on these methods, as their usefulness for marketing has, to the best of our knowledge, not yet been covered in a review chapter.

11.3 State-Space Models, the Kalman Filter, and Normative Decision Making[5]

State-space models offer many advantages, of which we list ten at the end of Section 11.3.1. In what follows we first explain what a state-space model is; then its estimation and inference; its applications in marketing; and, finally, its role in normative analysis.

11.3.1 State Space Models

Linear state-space models are expressed by two sets of equations:

$$Y_t = Z_t\alpha_t + c_t + \varepsilon_t, \text{ and} \tag{11.1}$$

$$\alpha_t = T_t\alpha_{t-1} + d_t + \nu_t, \tag{11.2}$$

---

[4] Most applications of persistence modeling consider the impact of marketing decisions (e.g. an unexpected advertising increase, an additional promotion) that do not alter the nature of the data-generating process (see e.g. Dekimpe et al. 1999 or Srinivasan et al. 2004). As such, the implications of more drastic regime changes (e.g. a switch from EDLP to HiLo pricing strategy) tends to fall outside the scope of these studies. Still, restricted policy simulations where the data-generating process is altered were considered in Pauwels (2004), and offered many managerially useful insights. We refer to Franses (2005) or van Heerde, Dekimpe, and Putsis (2005) for an in-depth discussion on the use of time-series modeling for policy simulation, and their resulting sensitivity to the Lucas critique.

[5] A gentle introduction may be found in Meinhold and Singpurwalla (1983), who explain the Kalman filter using the language of Bayesian updating. The following recommended references are arranged in increasing level of sophistication: Harvey (1994) offers econometric applications; Shumway and Stoffer (2006) describe applied time-series models; Harrison and West (1997) provide a Bayesian perspective; Durbin and Koopman (2001) present a unifying perspective underlying Bayesian and frequentist views; Lewis (1986) explains both the normative (i.e., optimal actions) and estimation (i.e., model identification) issues; finally, Jazwinski (1970), the *pioneering* book on this topic, reveals the provenance of Kalman filters in Mathematics and Control Engineering (predating their use in statistical and social sciences).

where $\varepsilon_t \sim N(0, H_t)$, $\nu_t \sim N(0, Q_t)$, Y is a random vector (m×1) and α is random vector (n×1), where m could be greater than, less than or equal to n. The vector $Y_t = (y_{1t,} y_{2t}, \ldots y_{mt})'$ contains observed time-series such as sales of brand A, sales of brand B, and so on observed over several time periods t = 1, ..., R. Similarly, $\alpha_t = (\alpha_{1t,} \alpha_{2t}, \ldots \alpha_{nt})'$ includes multiple state-variables. A state variable α_t can play diverse roles, for example, a time-varying parameter like copy wearout in Naik, Mantrala and Sawyer (1998) or Bass et al. (2007); a construct such as goodwill or brand equity as in Sriram and Kalwani (2007); a set of variables such as market shares as in Naik, Raman and Winer (2005); a reading of a performance barometer as in Pauwels & Hanssens (2007); a random variable to capture non-stationarity and heterogeneity as in van Heerde, Mela and Manchanda (2004), or to impute missing values via the cubic spline as in Biyalogorsky and Naik (2003). A discrete-valued α_t opens up new class of models such as "Hidden Markov Models" as in Smith, Naik and Tsai (2006) or Netzer, Latting and Srinivasan (2008).

The dimensions of other matrices and vectors in the dynamic system conform to those of (Y, α). Specifically, the link matrix Z is an m x n matrix; T is an n x n transition matrix; the drift vectors (c, d) are m × 1 and n × 1, respectively; the covariance matrices H and Q have dimensions m x m and n x n, respectively. For example, in Naik and Raman (2003) integrated marketing communications (IMC) model $S_t = \lambda S_{t-1} + \sum_{i=1}^{2} \beta_i x_i + \kappa x_1 x_2 + \nu_t$, we note that the scalar state variable α = S, the 1 × 1 transition matrix T = λ, the 1 × 1 drift vector d = $\Sigma\beta_i x_i + \kappa x_1 x_2$, the transition noise $Q = \sigma_\nu^2$, $Y = S + \varepsilon$, so $Z = 1, c = 0$ and the observation noise $H = \sigma_\varepsilon^2$. In this manner, several well-known marketing models may be expressed as special cases of the state-space form (see Table 11.4).

Equation (11.2) is called the transition (or plant) equation, which captures the dynamics of the physical system explicitly. It is linked to the observed (i.e., measured) variables via equation (11.1), which is therefore called the measurement or observation equation. The vector Y is the observation vector; α is the state vector. The drift vectors (c, d) represent the effects of exogenous variables (e.g., $c_t = X_t'\beta$, $d_t = W_t'\gamma$,where X and W contain multiple variables, and (β,γ) are conformable parameter vectors). The subscript t denotes that the given quantity can change over time, indicating that it is potentially time-varying and therefore implicitly dynamic (besides the state vector, which is explicitly dynamic). Table 11.2 summarizes the names and dimensions of vector-matrices in the state-space form.

The state-space form, given by (11.1) and (11.2), is very general. For example, standard time-series models like VAR, VMA, ARIMAX are special cases (see, e.g., Durbin and Koopman 2001, Harvey 1994). In addition, structural models that capture dynamic marketing phenomena such as Brandaid, the Nerlove-Arrow model, the Vidale-Wolfe model, Tracker, Litmus, the Bass diffusion model and the IMC model have a state-space representation (see Tables 11.3 and 11.4 for details). When the state-space form is nonlinear, we express equation (11.2) more generally as $\alpha_t = T(\alpha_{t-1}) + d_t + \nu_t$, where T($\alpha$) denotes a transition function (see, e.g., the Bass diffusion model in Tables 11.3 and 11.4).

Table 11.2 Names and notation for vectors and matrices in state space models

Notation	Vector or Matrix	Name	Dimension
Y	Vector	Observation Vector	m × 1
α	Vector	State Vector	n × 1
T	Matrix	Transition Matrix	n × n
T(α)	Vector-valued function	Transition function	n × 1 outputs; n × 1 arguments
C	Vector	Drift vector (in observation)	n × 1
D	Vector	Drift vector (in transition)	m × 1
Z	Matrix	Link Matrix (from state to observation)	m × n
ε	Vector	Observation errors	m × 1
ν	Vector	Transition errors	n × 1
H	Matrix	Observation noise matrix	m × m
Q	Matrix	Transition noise matrix	n × n

The Kalman filter is a method for determining the moments (e.g., mean and covariance) of the dynamic state vector α_t, at each instant t, given the observations in Y_t. It is called a "filter" because it extracts the signal from noisy observations in Y_t via two steps. The first step, known as the *time-update*, predicts the moments of α as the system in (11.2) moves from the previous instant t-1 to the current instant t. In this time-update step, before any new observations become available, changes in the moments of α are solely due to the motion of the system in (11.2). In the second step, which is called *measurement-update*, the moments of α are updated based on the information made available in the observation vector Y, which could be noisy or incomplete (i.e., missing data) or redundant (i.e., multiple measurements on a given state variable). The exact formulae for time- and measurement- updates are given in equation (11.17) of the Appendix. Specifically, the prior mean and covariance is due to time-updating; the posterior mean and covariance is due to a measurement update. In between the prior and posterior moments in (11.17), there appears a weighting factor, known as the Kalman gain, which optimally balances (i.e., in the sense of minimizing mean squared errors) the accuracy of the dynamic model relative to the precision of actual observations. Intuitively, when observations are noisy, the filter discounts the observed data by placing a small weight; on the other hand, when model forecasts are inaccurate, the filter discounts these forecasts by relying more on the actual observed data. Thus the Kalman filter via the recursive equations in (11.17) optimally combines information from both the dynamic model and the actual observations to determine the current state of the system (i.e., the distribution of α).

Last but not least, there are many practical advantages for casting ARIMAX or any other structural dynamic model in the above state-space form:

Table 11.3 Description of dynamic marketing models

Model	The Mathematical Model	Model Description
Vidale and Wolfe (1957)	$\frac{dA}{dt} = \beta(1 - A)u - \delta A$ *Discrete Version* $A_t = (1 - \beta u_t - \delta)A_{t-1} + \beta u_t$	Over a small period of time, increase in brand awareness (A) is due to the brand's advertising effort (u), which influences the unaware segment of the market, while attrition of the aware segment occurs due to forgetting of the advertised brand.
Nerlove and Arrow (1962)	$\frac{dA}{dt} = \beta u - \delta A$ *Discrete Version* $A_t = (1 - \delta)A_{t-1} + \beta u_t$	The growth in awareness depends linearly on the advertising effort, while awareness decays due to forgetting of the advertised brand.
Brandaid (Little 1975)	$A_t = \lambda A_{t-1} + (1 - \lambda)g(u_t)$ $g(u) = \frac{u^\beta}{\phi + u^\beta}$	Brand awareness in the current period depends partly on the last period brand awareness and partly on the response to advertising effort; the response to advertising effort can be linear, concave, or S-shaped.
Tracker (Blattberg and Golanty 1978)	$A_t - A_{t-1} = (1 - e^{\alpha - \beta u_t})(1 - A_{t-1})$	The incremental awareness depends on the advertising effort, which influences the unaware segment of the market.
Litmus (Blackburn and Clancy 1982)	$A_t = (1 - e^{-\beta u_t})A^* + e^{-\beta u_t}A_{t-1}$	The current period awareness is a weighted average of the steady-state ("maximum") awareness and the last period awareness. The weights are determined by the advertising effort in period t.
Bass Model (1969)	$S_t = S_{t-1} + p(M - S_{t-1}) + q\frac{S_{t-1}}{M}(M - S_{t-1})$	Sales grow due to both the untapped market and contagion effects.
IMC Model (Naik and Raman 2003)	$S_t = \alpha + \beta_1 u_{1t} + \beta_2 u_{2t} + \kappa u_{1t} u_{2t} + \lambda S_{t-1}$	Sales grow due to not only direct effects of advertising (β_i), but also indirect effects of synergy (κ) between advertising.

i. the *exact* likelihood function can be computed to obtain parameter estimates, infer statistical significance, and select among model specifications;
ii. a *common* algorithm, based on Kalman filter recursions, can be used to analyze and estimate diverse model specifications;
iii. *multivariate* outcomes are handled as easily as univariate time-series;
iv. inter-equation *coupling* and correlations across equations can be estimated;
v. *missing values* do not require special algorithms to impute or delete data;

Table 11.4 System matrices for comparison of models

System Matrices	Vidale-Wolfe	Nerlove-Arrow	Brandaid	Tracker	Litmus	Bass model	IMC model
State Vector, α_t	$[A_t]$	$[A_t]$	$[A_t]$	$[A_t]$	$[A_t]$	$[S_t]$	$[S_t]$
Observation Vector, z	[1]	[1]	[1]	[1]	[1]	[1]	[1]
Transition Matrix, T_t	$[1-g(u_t)-\delta]$	$[1-\delta]$	$[\lambda]$	$[1-g(u_t)]$	$[1-g(u_t)]$		$[\lambda]$
Transition function, T(S)						$(1-p+q)S$ $-qS^2/M$	
Drift Vector, d_t	$[g(u_t)]$	$[g(u_t)]$	$[(1-\lambda)g(u_t)]$	$[g(u_t)]$	$[A^*g(u_t)]$	pM	g(u)
Observation Noise, H	σ_ε^2	σ_ε^2	σ_ε^2	σ_ε^2	σ_ε^2	σ_ε^2	σ_ε^2
Transition Noise, Q	σ_v^2	σ_v^2	σ_v^2	σ_v^2	σ_v^2	σ_v^2	σ_v^2
Response Function, g(x)	βx	βx					

vi. *unequally spaced* time-series observations pose no additional challenges;
vii. *unobserved* variables such as goodwill or brand equity, can be incorporated;
viii. *time-varying coefficients* and *non-stationarity* can be specified;
ix. *heterogeneity* via random coefficients can be introduced seamlessly;
x. *normative decision-making* can be integrated with econometric analyses.

Below, we briefly describe the maximum-likelihood estimation of state-space models, which are widely available in standard software packages (e.g., Eviews, SAS, GaussX, Matlab).

11.3.2 Parameter Estimation, Inference, Selection

Suppose we observe the sequence of multivariate time series $Y = \{Y_t\}$ and $X = \{X_t\}$ for $t = 1, \ldots, R$. Then, given the model equations (11.1) and (11.2), the probability of observing the entire trajectory $(Y_1, Y_2, \ldots, Y_R)$ is given by the likelihood function,

$$\begin{aligned} L(\Theta; X; Y) &= p(Y_1, Y_2, \cdots, Y_R) \\ &= p(Y_1)p(Y_2|Y_1)p(Y_3|(Y_1, Y_2)) \cdots p(Y_R|(Y_1, \cdots, Y_{R-1})) \\ &= p(Y_1|\Im_0)p(Y_2|\Im_1)p(Y_3|\Im_2) \cdots p(Y_R|\Im_{R-1}) \\ &= \prod_{t=1}^{R} p(Y_t|\Im_{t-1}). \end{aligned} \tag{11.3}$$

In equation (11.3), $p(Y_1, Y_2, \ldots, Y_R)$ denotes the joint density function, and $p(Y_t|(Y_1, \ldots; , Y_{t-1})) = p(Y_t|\Im_{t-t})$ represents the conditional density. The Appendix provides the moments of the random variable $Y_t|\Im_{t-1}$via Kalman filter recursions. In addition, the information set $\Im_{t-1} = Y_1, Y_2, \ldots, Y_{t-1}$ contains the history generated by market activity up to time t-1.

Next, we obtain the parameter estimates by maximizing the log-likelihood function with respect to Θ:

$$\hat{\Theta} = \underset{\Theta}{\mathrm{ArgMax}}\, \mathrm{Ln}(\mathrm{L}(\Theta)), \tag{11.4}$$

which is asymptotically unbiased and possesses minimum variance.

To conduct statistical inference, we obtain the standard errors by taking the square-root of the diagonal elements of the covariance matrix:

$$\mathrm{Var}(\hat{\Theta}) = \left[-\frac{\partial^2 \mathrm{Ln}(\mathrm{L}(\Theta))}{\partial\Theta\partial\Theta'} \right]^{-1}_{\Theta=\hat{\Theta}}, \tag{11.5}$$

where the right-hand side of (11.5) is the negative inverse of the Hessian matrix evaluated at the maximum-likelihood estimates (resulting from (11.4)).

Finally, for model selection, we compute the expected Kullback-Leibler (K-L) information metric and select the model that attains the smallest value on this K-L metric (see Burnham and Anderson 2002 for details). An approximation of the K-L metric is given by Akaike's information criterion, $\mathrm{AIC} = -2\mathrm{L}^* + 2\mathrm{p}$, where $\mathrm{L}^* = \max \mathrm{Ln}(\mathrm{L}(\Theta))$ and p is the number of variables in X_t. As model complexity increases, both L^* and p increase; thus, the AIC balances the tradeoff between goodness-of-fit and parsimony. However, the AIC ignores both the sample size and the number of variables in Y_t. Hurvich and Tsai (1993) provide the bias-corrected information criterion for finite samples:

$$AIC_C = -2L^* + \frac{R(Rm + pm^2)}{R - pm - m - 1}, \tag{11.6}$$

where R is the sample size, p and m are the number of variables in X and Y, respectively. To select a specific model, we compute (11.6) for different model specifications and retain the one that yields the smallest value.

11.3.3 Marketing Applications

In marketing, Xie et al. (1997) and Naik et al. (1998) pioneered the Kalman filter estimation of dynamic models. Specifically, Xie et al. (1997) studied the nonlinear but univariate dynamics of the Bass model, while Naik et al. (1998)

estimated the multivariate but linear dynamics of the modified Nerlove-Arrow model. To determine the half-life of an advertising campaign, Naik (1999) formulates an advertising model with time-varying, non-stationary effects of advertising effectiveness and then applies the Kalman filter to estimate copy and repetition wear out. His empirical results suggest that the half-life of Docker's "Nice Pants" advertising execution was about 3 months. Neelamegham and Chintagunta (1999) incorporated non-normality via a Poisson distribution to forecast box-office sales for movies. To control for the biasing effects of measurement errors in dynamic models, Naik and Tsai (2000) propose a modified Kalman filter and show its satisfactory performance on both statistical measures (e.g., means square error) and managerial metrics (e.g., budget, profit). In the context of multimedia communications, Naik and Raman (2003) design a Kalman filter to establish the existence of synergy between multiple media advertising. Biyalogorsky and Naik (2003) develop an unbalanced filter with m = 3 dependent variables and n = 2 unobserved state variables to investigate the effects of customers' online behavior on retailers' offline sales and find negligible cannibalization effects (contrary to managers' fears). They also show how to impute missing values by fitting a cubic spline smoothing via a state-space representation. To investigate the effects of product innovation, van Heerde, Mela and Manchanda (2004) deploy state space models to incorporate non-stationarity, changes in parameters over time, missing data, and cross-sectional heterogeneity, while Osinga, Leeflang and Wieringa (2007) employ state-space models to capture multivariate persistence effects.

To understand how to integrate normative decision-making with empirical state-space models, see Naik and Raman (2003) for multimedia allocation in the presence of synergy and Naik et al. (2005) for marketing-mix allocation in the presence of competition. In the context of multiple themes of advertising, Bass et al. (2006) generalize an advertising wearout model for a single ad copy developed by Naik et al. (1998). Their results indicate that copy wearout for a price-offer theme is faster than that for reassurance ads, furnishing new market-based evidence to support the notion that "hard sell" ads (e.g., informative) wear out faster than "soft sell" ads (e.g., emotional appeals). Comparing the optimal versus actual allocation of the total GRPs across the five different themes, they investigate the policy implications for re-allocating the same level of total budget. Optimal re-allocation suggests that the company increases spending on reconnect and reassurance ads at the expense of the other three themes. This re-allocation would generate an additional 35.82 million hours of calling time, which represents about 2% increase in demand.

An important question is whether or not it is possible to discover the synergy between different communication activities with traditional methods. This issue was investigated in Monte-Carlo studies by Naik, Schultz and Srinivasan (2008), who check whether regression analysis accurately estimates the true impact of marketing activities. They report the eye-opening result that *regression analysis yields substantially biased parameter estimates because market data* *contain measurement noise*. This result holds even when a *dependent* variable in dynamic advertising models is noisy. More specifically, in their simulation studies, the bias in ad effectiveness estimates range from 34 to 41%, whereas both carryover effects and cross-media synergy display downward bias of 13.6 and 27.5%, respectively. Naik and Tsai (2000) also offer similar evidence suggesting that measurement noise causes parameter biases in dynamic models. Empirical analysis based on actual market data also comport with these simulation-based findings. For example, the analyses of Toyota Corolla's multimedia campaign reveal that the estimated effects of magazine and rebate effectiveness are more than twice as large as they should be.

Given the perils of regression analysis, are there alternative approaches that managers can adopt to estimate the effects of marketing activities and synergies? Fortunately, the answer is affirmative—the Kalman filter approach described above yields unbiased estimates even in the presence of measurement noise. Naik, Schultz, and Srinivasan (2008) compare the performance of Kalman filter estimation with regression analysis under identical conditions, and they show that the Kalman filter approach yields improved estimates that are much closer to the true effects of multimedia campaign than the corresponding regression estimates.

11.3.4 Normative Decision-Making

One of the advantages of state-space modeling, as noted earlier, is that we can integrate econometric analyses with normative decision-making problems faced by managers. Below we set up such a marketing problem and illustrate how to solve it.

11.3.4.1 Managerial Decision Problem

Consider a company spending resources on two marketing activities, say television and print advertising. A brand manager faces the decision problem of determining the total budget and its allocation to these activities over time. Suppose she decides to spend effort over time as follows: $\{u_1, u_2, \ldots, u_t, \ldots\}$ and $\{v_1, v_2, \ldots, v_t, \ldots\}$. For example, "effort" can be defined specific to a context, for example, GRPs in advertising management or the number of sales calls in salesforce management. Given this specific plan $\{(u_t, v_t) : t \in (1, 2, \ldots)\}$, she generates the sales sequence $\{S_1, S_2, \ldots, S_t, \ldots\}$ and earns an associated stream of profits $\{\pi_1, \pi_2, \ldots, \pi_t, \ldots\}$. Discounting the future profits at the rate ρ, she computes the net present value $J = \sum_{t=1}^{\infty} e^{-\rho t}\pi_t(S_t, u_t, v_t)$. In other words, a media plan $(u, v) = \{(u_t, v_t){:}\ t = 1, 2, \ldots\}$ induces a sequence of sales that yields a stream of profits whose net present value is $J(u, v)$.

Formally, the budgeting problem is to find the *optimal* plan (u^*, v^*)—one that attains the maximum value J^*. To this end, the brand manager needs to determine $u^*(t)$and $v^*(t)$ by maximizing

$$J(u, v) = \int_0^\infty e^{-\rho t}\Pi(S(t), u(t), v(t))dt, \qquad (11.7)$$

where ρ denotes the discount rate, $\Pi(S, u, v) = mS - c(u, v)$ is the profit function with margin m and cost function $c(\cdot)$, and $J(u, v)$ is the performance index for any *arbitrary* multimedia policies $(u(t), v(t))$. To capture diminishing return of incremental effort, we further assume a quadratic cost function $c(u, v) = u^2 + v^2$. Below we illustrate how to derive the optimal plan using the IMC model proposed by Naik and Raman (2003).

11.3.4.2 Solution via Optimal Control Theory

In their IMC model, the sales dynamics is $S_t = \beta_1 u_t + \beta_2 v_t + \kappa u_t v_t + \lambda S_{t-1}$, where S_t is brand sales at time t, (β_1, β_2) are the effectiveness of marketing activities 1 and 2, (u_1, u_2) are dollars spent on those two activities, κ captures the synergy between them, and λ is the carryover effect. For other marketing problems, the essential dynamics would arise from the transition equation (11.2). If we have multiple transition equations in (11.2), the following approach generalizes (as we explain below). We re-express this dynamics in continuous-time as follows:

$$\frac{dS}{dt} = \beta_1 u(t) + \beta_2 v(t) + \kappa u(t)v(t) - (1 - \lambda)S(t), \qquad (11.8)$$

where dS/dt means instantaneous sales growth.

Then, to maximize our objective function in (11.7) subject to the dynamics specified in (11.8), we define the Hamiltonian function:

$$H(u, v, \mu) = \Pi(S, u, v) + \mu(\beta_1 u + \beta_2 v + \kappa uv - (1 - \lambda)S), \qquad (11.9)$$

where $\Pi(S, u, v) = mS - u^2 - v^2$ and μ is the co-state variable. We note two points; first, it is convenient to maximize H(.) in (11.9) rather than J(.) in (11.7), although the resulting solutions satisfy both these functions. Second, if we have an $n \times 1$ vector transition equation in the state space model (11.2), we would extend H(.) in (11.9) by adding additional co-state variables because each state equation has an associated co-state variable μ_j, $j = 1, \ldots, n$.

At optimality, the necessary conditions are as follows:

$$\frac{\partial H}{\partial u} = 0, \quad \frac{\partial H}{\partial v} = 0, \quad \frac{d\mu}{dt} = \rho\mu - \frac{\partial H}{\partial S}. \qquad (11.10)$$

Furthermore, these conditions are also sufficient because $H(\cdot)$ is concave in u and v. Applying the optimality conditions, we differentiate (9) with respect to u and v to get

$$\frac{\partial H}{\partial u} = 0 \quad \Rightarrow -2u + \beta_1\mu + \kappa\mu\nu = 0$$
$$\frac{\partial H}{\partial \nu} = 0 \quad \Rightarrow -2\nu + \beta_2\mu + \kappa\mu u = 0$$

Solving these algebraic equations simultaneously, we express the solutions in terms of the co-state variable:

$$u^* = \frac{\mu(2\beta_1 + \mu\beta_2\kappa)}{4 - \mu^2\kappa^2} \text{ and } v^* = \frac{\mu(2\beta_2 + \mu\beta_1\kappa)}{4 - \mu^2\kappa^2}. \tag{11.11}$$

The remaining step is to eliminate the co-state variable $\mu(t)$ by expressing it in terms of model parameters. To this end, we use the third optimality condition in (11.10):

$$\frac{d\mu}{dt} = \rho\mu - \frac{\partial H}{\partial S} \quad \Rightarrow \frac{d\mu}{dt} = -m + \mu(1-\lambda) + \rho\mu.$$

To solve this differential equation, we note that transversality conditions for an autonomous system with infinite horizon are obtained from the steady-state for state and co-state variables (Kamien and Schwartz 1991, p. 160), which are given by $\partial S/\partial t = 0$ and $\partial\mu/\partial t = 0$, respectively. Consequently, $\mu(t) = \frac{m}{(1-\lambda+\rho)}$, which we substitute in (11.11) to obtain the optimal spending plans:

$$u^* = \frac{m(\beta_2\kappa m + 2\beta_1(1+\rho-\lambda))}{4(1+\rho-\lambda)^2 - \kappa^2 m^2} \text{ and } v^* = \frac{m(\beta_1\kappa m + 2\beta_2(1+\rho-\lambda))}{4(1+\rho-\lambda)^2 - \kappa^2 m^2}. \tag{11.12}$$

From (11.12), we finally obtain the total budget $B = u^* + v^*$ as

$$B = \frac{(\beta_1 + \beta_2)m}{2(1+\rho-\lambda) - \kappa m}, \tag{11.13}$$

and the optimal media mix $\Lambda = u^*/v^*$ as

$$\Lambda = \frac{2\beta_1(1+\rho-\lambda) + m\beta_2\kappa}{2\beta_2(1+\rho-\lambda) + m\beta_1\kappa}. \tag{11.14}$$

11.3.4.3 Normative Insights

Although we can generate several propositions by analyzing comparative statics via (11.13) and (11.14), we present three main insights and implications (see Naik and Raman 2003 for their proofs and intuition).

Proposition 1 *As synergy (κ) increases, the firm should increase the media budget.*

This result sheds light on the issue of overspending in advertising. The marketing literature (see Hanssens et al. 2001, p. 260) suggests that advertisers *overspend*, i.e., the actual expenditure exceeds the optimal budget implied by normative models. However, the claim that "advertisers overspend" is likely to be exaggerated in an IMC context because the optimal budget itself is *understated* when models ignore the impact of synergy. To see this clearly, we first compute the optimal budget from (13) with synergy ($\kappa \neq 0$) and without it ($\kappa = 0$). Then, we find that the optimal budget required for managing multimedia activities in the presence of synergy is always larger than that required in its absence. Hence, in practice, if advertisers' budgets reflect their plans for integrating multimedia communications, then overspending is likely to be smaller.

Proposition 2 *As synergy increases, the firm should decrease (increase) the proportion of media budget allocated to the more (less) effective communications activity. If the various activities are equally effective (i.e., $\beta_1 = \beta_2$), then the firm should allocate the media budget equally amongst them, regardless of the magnitude of synergy.*

The counter-intuitive nature of this result is its striking feature. To understand the gist of this result, suppose that two activities have unequal effectiveness (say, $\beta_1 > \beta_2$). Then, in the absence of synergy ($\kappa = 0$), the optimal spending on an activity depends only on its own effectiveness; hence, a larger amount is allocated to the more effective activity (see Proposition 1). However, in the presence of synergy ($\kappa \neq 0$), optimal spending depends not only on its own effectiveness, but also on the spending level for the *other* activity. Consequently, as synergy increases, marginal spending on an activity increases at a rate proportional to the spending level for the other activity. Hence, even though the optimal spending levels are endogenous actions, they also affect each other due to synergy. Optimal spending on the more effective activity increases slowly, relative to the increase in the optimal spending on the less effective activity. Thus, the proportion of budget allocated to the *more* effective activity *decreases* as synergy increases.

If the two activities are equally effective, then the optimal spending levels on both of them are equal. Furthermore, as synergy increases, marginal spending on each of them increases at the *same* rate. Hence, the optimal allocation ratio remains constant at fifty percent, regardless of the increase or decrease in synergy.

To clarify this result, we present a numerical example. Consider two communications activities: TV and print advertising. Let TV ads be twice as effective as print ads; specifically, $\beta_1 = 2$ and $\beta_2 = 1$. For this illustration, we assume that $\kappa = 1$, $\rho = m = (1 - \lambda) = 0.1$. Then, substituting these values in Equations (11.13) and (11.14), we compute the optimal budget $B = 1$ and the optimal allocation Λ is 60:40. Now suppose that synergy increases from $\kappa = 1$ to $\kappa = 2$. Then, the total budget increases from $B = 1$ to $B = 1.5$, but the allocation ratio Λ becomes 55:45. In other words, the budget allocated to the more effective TV advertising decreases from 60 to 55%, and that for the less effective print advertising increases from 40 to 45%.

This finding has implications for emerging media, for example, Internet advertising. Companies should not think of Internet advertising and offline advertising (TV, Print) as *competing* alternatives. Rather, these activities possess different effectiveness levels and may benefit from integrative efforts to generate cross-media synergies. If so, the total media budget as well as its allocation to Internet advertising would grow.

Proposition 3 *In the presence of synergy, the firm should allocate a non-zero budget to an activity even if its direct effectiveness is zero.*

This result clearly demonstrates that companies must *act differently* in the context of IMC. According to extant models of advertising that ignore synergy, an advertiser should allocate a zero budget to an ineffective activity (i.e., $v^* = 0$ if $\beta_2 = 0$). In contrast, in the presence of synergy, the company benefits not only from the direct effect of an activity but also from its joint effects with *other* activities. Hence, they should *not* eliminate spending on an ineffective activity because it can enhance the effectiveness of other activities by its synergistic presence. We call this phenomenon the *catalytic influence* of an activity.

In marketing, many activities exert a catalytic influence on one another. For example, business-to-business advertising may not directly influence purchase managers to buy a company's products, but it may enhance sales call effectiveness. Another example comes from the pharmaceutical industry; product samples or collateral materials may not directly increase sales of prescription medicines, but it may enhance the effectiveness of detailing efforts (Parsons and Vanden Abeele 1981). Indeed, marketing communications using billboards, publicity, corporate advertising, event marketing, in-transit ads, merchandising, and product placement in movies arguably may not have measurable impacts on sales. Yet, advertisers spend millions of dollars on these activities. Why? The IMC framework implies that these activities, by their mere presence in the communications mix, act like catalysts, and enhance the effectiveness of other activities such as broadcast advertising or salesforce effort.

The above discussion clearly illustrated how time-series models can be linked to normative decision making. More research is needed along these lines, however, especially on how models that distinguish between short- and long-run marketing effectiveness (as described in Section 11.2) can be used to derive optimal pricing and spending policies, reflecting management's short- and long-run objectives.

11.4 Conclusion

In this paper, we reviewed two time-series approaches that have received considerable attention in the recent marketing literature: (i) persistence modeling, and (ii) state-space modeling. However, this by no means offered an exhaustive discussion of all time-series applications in marketing. Because of space limitations, we did not review the use of "more traditional" time-series techniques in marketing, such as univariate ARIMA modeling, multivariate transfer-function modeling, or Granger-causality testing. A review of these applications is given in Table 11.1 of Dekimpe and Hanssens (2000). Similarly, we did not discuss the frequency-domain approach to time-series modeling (see e.g. Bronnenberg, Mela and Boulding 2006 for a recent application on the periodicity of pricing), nor did we review recent applications of band-pass filters to isolate business-cycle fluctuations in marketing time series (see e.g. Deleersnyder et al. 2004 or Lamey et al. (2007), or the use of smooth-transition regression models to capture different elasticity regimes (see e.g. Pauwels, Srinivasan and Franses 2007). Indeed, the use of time-series techniques in marketing is expanding rapidly, covering too many techniques and applications to be fully covered in detail in a single chapter.

Referring to the expanding size of marketing data sets, the accelerating rate of change in the market environment, the opportunity to study the marketing-finance relationship, and the emergence of internet data sources, Dekimpe and Hanssens argued in 2000 that "for time-series modelers in marketing, the best is yet to come." (p. 192). In a recent *Marketing Letters* article, Pauwels et al. (2004) identified a number of remaining challenges, including ways to (i) capture asymmetries in market response, (ii) allow for different levels of temporal aggregation between the different variables in a model, (iii) cope with the Lucas Critique, (iv) handle the short time series often encountered when working at the SKU level, and (v) incorporate Bayesian inference procedures in time-series modeling. In each of these areas, we have already seen important developments. For example, Lamey et al. (2007) developed an asymmetric growth model to capture the differential impact of economic expansions and recessions on private-label growth, and Ghysels, Pauwels and Wolfson 2006 introduced Mixed Data Sampling (MIDAS) regression models in marketing to dynamically relate hourly advertising to daily sales; see also Tellis and Franses (2006) who derive for some basic models what could be the optimal level of aggregation. Tests for the Lucas critique are becoming more widely accepted in marketing (see e.g. Franses 2005, van Heerde et al. 2005, 2007). Krider et al. (2005) developed graphical procedures to test for Granger causality between short time series, and Bayesian procedures are increasingly used to estimate error-correction specifications (see e.g Fok et al. 2006, van Heerde et al. 2007).

In sum, the diffusion of time-series applications in marketing has started. We hope the current chapter will contribute to this process.

Appendix: Moments of the Conditional Density $\mathbf{p(Y_t|\vartheta_{t-1})}$

This appendix provides the moments of the conditional density $p(Y_t|\vartheta_{t-1})$. We recall that the observation equation is $Y_t = Z_t\alpha_t + c_t + \varepsilon_t$, the transition equation is $\alpha_t = T_t\,\alpha_{t-1} + d_t + \nu_t$, and error terms are distributed as $\varepsilon_t \sim N(0, H_t)$ and $\nu_t \sim N(0,Q_t)$. Since the error terms are distributed normally and both the transition and observation equations are linear in the state variables α_t, the random variable $(Y_t|\vartheta_{t-1})$is normally distributed (because the sum of normal random variables is normal.)

Let $\hat{Y}_t$ denote the mean and f_t be the variance of the normal random variable $(Y_t|\vartheta_{t-1})$. By taking the expectation of observation equation, we obtain

$$\begin{aligned}\hat{Y}_t &= E[Y_t|\vartheta_{t-1}] \\ &= E[Z_t\alpha_t + c_t + \varepsilon_t|\vartheta_{t-1}] \\ &= Z_t E[\alpha_t|\vartheta_{t-1}] + c_t + 0 \\ &= Z_t a_{t|t-1} + c_t,\end{aligned} \tag{11.15}$$

where $a_{t|t-1}$ is the mean of the state variable $\alpha_t|\vartheta_{t-1}$ Similarly, the variance of $(Y_t|\vartheta_{t-1})$is

$$\begin{aligned}f_t &= \mathrm{Var}[Y_t|\vartheta_{t-1}] \\ &= \mathrm{Var}[Z_t\alpha_t + \varepsilon_t|\vartheta_{t-1}] \\ &= Z_t \mathrm{Var}[\alpha_t|\vartheta_{t-1}]Z_t' + \mathrm{Var}[\varepsilon_t|\vartheta_{t-1}] \\ &= Z_t P_{t|t-1} Z_t' + H_t,\end{aligned} \tag{11.16}$$

where $P_{t|t-1}$ is the covariance matrix of state variable $\alpha_t|\vartheta_{t-1}$.

Next, we obtain the evolution of mean vector and covariance matrix of α_t via the celebrated Kalman recursions (see, e.g., Harvey 1994 for details):

$$\begin{aligned}&\text{Prior mean :} && a_{t|t-1} = T_t a_{t-1} + d_t \\ &\text{Prior covariance :} && P_{t|t-1} = T_t P_{t-1} T_{t-1}' + Q_t \\ &\text{Kalman Gain Factor :} && K_t = P_{t|t-1} Z_t' f_t^{-1} \\ &\text{Posterior mean :} && a_{t|t} = a_{t|t-1} + K_t(Y_t - \hat{Y}_t) \\ &\text{Posterior covariance :} && P_{t|t} = P_{t|t-1} - K_t Z_t P_{t|t-1}.\end{aligned} \tag{11.17}$$

Finally, we apply recursions in (11.17) for each t, $t = 1, \ldots, R$ to obtain $a_{t|t-1}$ and $P_{t|t-1}$, starting with a diffused initial prior on $\alpha_0 \sim N(a_0, P_0)$. For example, given (a_0, P_0), we get $(a_{1|0}, P_{1|0})$ and thus $(a_{1|1}, P_{1|1})$; now given $(a_{1|1}, P_{1|1})$, we get $(a_{2|1}, P_{2|1})$ and thus $(a_{2|2}, P_{2|2})$; and so on. Knowing $a_{t|t-1}$ and $P_{t|t-1}$ for each t, we determine the moments of $(Y_t|\vartheta_{t-1})$ via equations (11.15) and (11.16). The initial mean vector, a_0, is estimated by treating it as hyper-parameters in the likelihood function.

References

Baghestani, H. 1991. Cointegration Analysis of the Advertising-Sales Relationship. *Journal of Industrial Economics* **34** 671–681.

Bass, F., N. Bruce, S. Majumdar, B.P.S. Murthi. 2007. A Dynamic Bayesian Model of Advertising Copy Effectiveness in the Telecommunications Sector. *Marketing Science* **26**(March–April) 179–195.

Biyalogorsky, E., P.A. Naik. 2003. Clicks and Mortar: The Effect of Online Activities on Offline Sales. *Marketing Letters* **14**(1) 21–32.

Blackburn, J.D., K.J. Clancy. 1982. LITMUS: A New Product Planning Model. A.A. Zoltners, *Marketing Planning Models*, North-Holland Publishing Company, New York, NY, 43–62.

Blattberg, R., J. Golanty. 1978. TRACKER: An Early Test Market Forecasting and Diagnostic Model for New Product Planning. *Journal of Marketing Research* **15**(May) 192–202.

Bronnenberg, B.J., V. Mahajan, W.R. Vanhonacker. 2000. The Emergence of Market Structure in New Repeat-Purchase Categories: The Interplay of Market Share and Retailer Distribution. *Journal of Marketing Research* **37**(1) 16–31.

Bronnenberg, B.J., C. Mela, W. Boulding. 2006. The Periodicitiy of Pricing. *Journal of Marketing Research* **43** 477–493.

Burnham, K. P., D.R. Anderson. 2002. *Model selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2 nd Edition, Springer-Verlag, New York, USA.

Cavaliere, G., G. Tassinari. 2001. Advertising Effect on Primary Demand: A Cointegration Approach. *International Journal of Advertising* **20** 319–339.

Chowdhury, A.R. 1994. Advertising Expenditures and the Macro-Economy: Some New Evidence. *International Journal of Advertising* **13** 1–14.

Dekimpe, M.G., D.M. Hanssens. 1995a. The Persistence of Marketing Effects on Sales. *Marketing Science* **14**(1) 1–21.

Dekimpe, M.G., D.M. Hanssens. 1995b. Empirical Generalizations about Market Evolution and Stationarity. *Marketing Science* **14**(3,2) G109-G121.

Dekimpe, M.G., D.M. Hanssens. 1999. Sustained Spending and Persistent Response: A New Look at Long-Term Marketing Profitability. *Journal of Marketing Research* **36**(4) 397–412.

Dekimpe, M.G., D.M. Hanssens . 2000. Time-series Models in Marketing: Past, Present and Future. *International Journal of Research in Marketing* **17**(2–3) 183–193.

Dekimpe, M.G., D. M. Hanssens. 2004. Persistence Modeling for Assessing Marketing Strategy Performance. in Moorman, C., D.R. Lehmann, *Assessing Marketing Strategy Performance*. Marketing Science Isntitute, Cambridge MA, 69–93.

Dekimpe, M.G., D.M. Hanssens, J.M. Silva-Risso. 1999. Long-Run Effects of Price Promotions in Scanner Markets. *Journal of Econometrics* **89**(1/2) 269–291.

Dekimpe, M.G., J.-B.E.M. Steenkamp, M. Mellens, P.Vanden Abeele. 1997. Decline and Variability in Brand Loyalty. *International Journal of Research in Marketing* **14**(5) 405–420.

Deleersnyder, B., M.G. Dekimpe, M. Sarvary, P.M. Parker. 2004. Weathering Tight Economic Times: The Sales Evolution of Consumer Durables over the Business Cycle.*Quantitative Marketing and Economics* **2**(4) 347–383.

Durbin, J., S.J. Koopman 2001. *Time Series Analysis by State Space Methods*. Oxford University Press: London, UK.

Enders, W. 1995. *Applied Econometric Time Series*.John Wiley, New York, N.Y.

Fok, D., C. Horváth, R. Paap, P.H. Franses 2006. A Hierarchical Bayes Error Correction Model to Explain Dynamic Effects of Price Changes. *Journal of Marketing Research* **43** 443–461.

Franses, P.H. 1994. Modeling New Product Sales; An Application of Cointegration Analysis. *International Journal of Research in Marketing* **11** 491–502.

Franses, P.H. 1998. *Time Series Models for Business and Economic Forecasting*. Cambridge University Press, Cambridge.

Franses, P.H. 2005. On the Use of Econometric Models for Policy Simulations in Marketing. *Journal of Marketing Research* **42**(1) 4–14.

Franses, P.H., T. Kloek, A. Lucas. 1999. Outlier Robust Analysis of Long-Run Marketing Effects for Weekly Scanner Data. *Journal of Econometrics*. **89**(1/2) 293–315.
Franses, P.H., S. Srinivasan, H. Peter Boswijk. 2001. Testing for Unit Roots in Market Shares. *Marketing Letters* **12**(4) 351–364.
Ghysels, E., K.H. Pauwels, P.J. Wolfson. 2006. The MIDAS touch: Linking Marketing to Performance at Different Frequencies.*Working Paper*.
Granger, C.W.J. 1981. Some Properties of Time Series Data and their Use in Econometric Model Specification. *Journal of Econometrics* **16**(1) 121–130.
Hamilton, J.D. 1994 *Time Series Analysis*.Princeton University Press, Princeton, N.J.
Hanssens, D.M. 1998. Order Forecasts, Retail Sales, and the Marketing Mix for Consumer Durables. *Journal of Forecasting* **17**(3) 327–346.
Hanssens, D.M., M. Ouyang. 2002. Hysteresis in Market Response: When is Marketing Spending an Investment? *Review of Marketing Science* 419
Hanssens, D.M., L.J. Parsons, R.L. Schultz. 2001. *Market Response Models: Econometric and Time Series Analysis*.Kluwer Academic Publishers, Boston, MA.
Harrison, J., M. West. 1997. *Bayesian Forecasting and Dynamic Models*. 2 nd edition, Springer: New York, NY.
Harvey, A.C. 1994. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press: New York
Horváth, C., P.S.H. Leeflang, J. Wieringa, D.R. Wittink. 2005. Competitive Reaction- and Feedback Effects Based on VARX Models of Pooled Store Data. *International Journal of Research in Marketing* **22**(4) 415–426.
Horváth, C., P.S.H. Leeflang, P.W. Otter. 2002. Canonical Correlation Analysis and Wiener-Granger Causality Tests: Useful Tools for the Specification of VAR Models. *Marketing Letters,* **13**(91) 53–66.
Hurvich, C., C.-L. Tsai. 1993. A Corrected Akaike Information Criterion for Vector Autoregressive Model Selection. *Journal of Time Series Analysis* **14**(3) 271–279.
Jazwinski, A.H. 1970. *Stochastic Process and Filtering Theory*. Academic Press: New York, NY.
Johnson, J.A., E.H. Oksanen, M.R. Veal, D. Frets. 1992. Short-run and Long-run Elasticities for Canadian Consumption of Alcoholic Beverages: An Error-Correction Mechanism/Cointegration Approach. *The Review of Economics and Statistics* **74**, 64–74.
Joshi, A., D.M. Hanssens. 2006. Advertising Spending and Market Capitalization. *Working Paper*. UCLA Anderson Graduate School of Management.
Jung, C., B.J. Seldon. 1995. The Macroeconomic Relationship between Advertising and Consumption. *Southern Economic Journal* **61**(3) 577–588.
Kamien, M.I., N. Schwartz. 1991. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*. 2 nd edition, North Holland, Amsterdam..
Krider, R.E., T. Li, Y. Liu, C. B. Weinberg. 2005. The Lead-Lag Puzzle of Demand and Distribution: A Graphical Method Applied to Movies. *Marketing Science* **24**(4) 635–645.
Lamey, L., B. Deleersnyder, M.G. Dekimpe, J-B.E.M. Steenkamp. 2007. How Business Cycles Contribute to Private-Label Success: Evidence from the U.S. and Europe. *Journal of Marketing* **71**(January) 1–15.
Leeflang, P.S.H., D.R. Wittink, M. Wedel, P.A. Naert. 2000. *Building Models of Marketing Decisions*. Kluwer Academic Publishers, Boston, MA.
Lewis, F.. 1986. *Optimal Estimation: With Introduction to Stochastic Control Theory*. John Wiley & Sons, New York, NY.
Lim, J., I. Currim, R.L. Andrews. 2005. Consumer Heterogeneity in the Longer-term Effects of Price Promotions. *International Journal of Research in Marketing* **22**(4) 441–457.
Little, J.D.C. 1975. Brandaid: A Marketing Mix Model, Part 1-Structure. *Operations Research* **23** 628–655.
McCullough, B.D., T. Waldon. 1998. The Substitutability of Network and National Spot Television Advertising. *Quarterly Journal of Business and Economics* **37**(2) 3–15.

Meinhold, R.J., N.D. Singpurwalla. 1983. Understanding the Kalman Filter. *The American Statistician* **37**(2) 123–127.

Moorman, C., D.R. Lehmann. 2004. *Assessing Marketing Strategy*.Marketing Science Institute, Cambridge MA.

Naik, P.A. 1999. Estimating the Half-life of Advertisements. *Marketing Letters* **10**(3) 351–362.

Naik, P.A., K. Raman. 2003. Understanding the Impact of Synergy in Multimedia Communications. *Journal of Marketing Research* **40**(3) 375–388.

Naik, P.A., C.-L. Tsai. 2000. Controlling Measurement Errors in Models of Advertising Competition. *Journal of Marketing Research* **37**(1) 113–124.

Naik, P.A., M.K. Mantrala, A. Sawyer. 1998. Planning Pulsing Media Schedules in the Presence of Dynamic Advertising Quality. *Marketing Science* **17**(3) 214–235.

Naik, P., K. Raman, R. Winer. 2005. Planning Marketing-Mix Strategies in the Presence of Interactions. *Marketing Science* **24**(1) 25–34.

Naik, P.A., D.E. Schultz and S. Srinivasan. 2007. Perils of Using OLS to Estimate Multimedia Communications Effects. *Journal of Advertising Research* 257–269.

Neelamegham, R., P. Chintagunta. 1999. A Bayesian Model to Forecast New Product Performance in Domestic and International Markets, *Marketing Science* **18**(2) 115–136.

Nerlove, M., K. Arrow. 1962. Optimal Advertising Policy Under Dynamic Conditions. *Econometrica* **29**(May) 129–42.

Netzer, O., J. Lattin, V. Srinivasan. 2008. A Hidden Markov Model of Customer Relationship Dynamics. *Marketing Science* **27**(March–April) 185–204.

Nijs, V.R., M.G. Dekimpe, J.-B.E.M. Steenkamp, D.M. Hanssens. 2001. The Category Demand Effects of Price Promotions. *Marketing Science* **20**(1) 1–22.

Nijs, V., S. Srinivasan, K. Pauwels. 2006. Retail-Price Drivers and Retailer Profits. *Marketing Science* forthcoming.

Osinga, E.C., P.S.H. Leeflang, J. Wieringa. 2007. The Persistence of Marketing Effects over the Brand's Life Cycle. *Working Paper*. University of Groningen.

Parsons, L.J., P.V. Abeele. 1981. Analysis of Sales Call Effectiveness. *Journal of Marketing Research* **18**(1) 107–113.

Pauwels, K. 2004. How Dynamic Consumer Response, Dynamic Competitor Response and Expanded Company Action Shape Long-Term Marketing Effectiveness. *Marketing Science* **23**(4) 596–610.

Pauwels, K., I. Currim, M.G. Dekimpe, E. Ghysels, D.M. Hanssens, N. Mizik, P. Naik. 2004. Modeling Marketing Dynamics by Time Series Econometrics. *Marketing Letters* **15**(4) 167–183.

Pauwels, K., and D.M. Hanssens. 2007. Performance Regimes and Marketing Policy Shifts in Mature Markets. *Marketing Science*, **26**(May–June) 293–311.

Pauwels, K., D.M. Hanssens, S. Siddarth. 2002. The Long-Term Effect of Price Promotions on Category Incidence, Brand Choice and Purchase Quality. *Journal of Marketing Research* **39**(4) 421–439.

Pauwels, K., J. Silva-Risso, S. Srinivasan, D.M. Hanssens. 2004. The Long-Term Impact of New-Product Introductions and Promotions On Financial Performance and Firm Value. *Journal of Marketing* **68**(4) 142–156.

Pauwels, K., S. Srinivasan. 2004. Who Benefits from Store Brand Entry? *Marketing Science* **23**(3) 364–390.

Pauwels, K., S. Srinivasan, P.H. Franses. 2007. When Do Price Thresholds Matter in Retail Categories", *Marketing Science* **26**(1) 83–100.

Shumway, R.H., D.S. Stoffer. 2006. *Time Seies Analysis and Its Applications*. Springer: New York, NY.

Smith, A., P. Naik, C.-L. Tsai. 2006. Markov-switching Model Selection Using Kullback-Leibler Divergence. *Journal of Econometrics* **134**(2) 553–577.

Srinivasan, S., F.M. Bass. 2000. Cointegration Analysis of Brand and Category Sales: Stationarity and Long-run Equilibrium in Market Shares. *Applied Stochastic Models in Business and Industry* **16** 159–177.

Srinivasan, S., P.T.L. Popkowski Leszczyc, F.M. Bass. 2000. Market Share Response and Competitive Interaction: The Impact of Temporary, Permanent and Structural Changes in Prices. *International Journal of Research in Marketing* **17**(4) 281–305.

Srinivasan, S., K. Pauwels, D.M. Hanssens, M.G. Dekimpe. 2004. Do promotions increase profitability and for whom. *Management Science* **50**(5) 617–629.

Sriram, S., M. Kalwani. 2007. Optimal Advertising and Promotion Budgets in Dynamic Markets with Brand Equity as a Mediating Variable. *Management Science* **53**(1) 46–60.

Steenkamp, J.-B.E.M., V.R. Nijs, D.M. Hanssens, M.G. Dekimpe. 2005. Competitive Reactions to Advertising and Promotion Attacks. *Marketing Science* **24**(1) 35–54.

Tellis, G.J., P.H. Franses. 2006. Optimal Data Interval for Estimating Advertising Response. *Marketing Science* **25** 217–229.

Van Heerde, H.J., M.G. Dekimpe, W.P. Putsis, Jr. 2005. Marketing Models and the Lucas Critique. *Journal of Marketing Research* **42**(1) 15–21.

Van Heerde, H.J., K. Helsen, M.G. Dekimpe. 2007. The Impact of a Product-Harm Crisis on Marketing Effectiveness. *Marketing Science* **26**(2) 230–245.

Van Heerde, H.J., C.F. Mela, P. Manchanda. 2004. The Dynamic Effect of Innovation on Market Structure, *Journal of Marketing Research* **41**(2) 166–183.

Vidale, M.L., H.B. Wolfe. 1957. An Operations-Research Study of Sales Response to Advertising. *Operations Research* **5**(June) 370–381.

Villanueva, J., S. Yoo, D.M. Hanssens. 2006. Customer Acquisition Channels and Customer Equity: A Long-Run View. *Working Paper*. UCLA Anderson Graduate School of Management.

Xie, J., M. Song, M. Sirbu, Q. Wang. 1997. Kalman Filter Estimation of New Product Diffusion Models. *Journal of Marketing Research* **34**(3) 378–393.

Zanias, G.P. 1994. The Long Run, Causality, and Forecasting in the Advertising-Sales Relationship. *Journal of Forecasting* **13** 601–610.