Time-Series Models in Marketing: Some Recent Developments

by Marnik G. Dekimpe and Dominique M. Hanssens

This paper discusses four developments that underlie the increased use of time-series models in marketing science in this century: the expansion of marketing databases, the accelerating rate of change in the business environment, the growing interest in the marketing-finance interface, and the growth in internet capabilities. We illustrate each development with examples from recent literature and draw some conclusions for the future of time-series models in marketing.

Keywords
Time-series models, econometrics, marketing databases, turbulence, marketing accountability, internet data

1. Introduction

Ten years ago, we contributed a paper on the Past, Present and Future of Time-Series Models in Marketing to the International Journal of Research in Marketing (Dekimpe/Hanssens 2000). In that paper, we described a number of reasons why, in the past, time-series (TS) techniques received little attention from marketing model builders and users, and discussed how the development of methods designed specifically to disentangle short-run from long-run movements (such as unit-root tests, cointegration and error-correction modeling, and persistence estimation) had resulted in a renewed interest for time-series techniques in marketing.

It is fair to say that, over the last decade, time-series modeling has indeed increased its popularity in the marketing-science community. Several observations help to substantiate this claim. First, the Marketing Dynamics Conference (MDC), which features the latest developments in the area, attracts an increasing number of scholars; it will be organized for the seventh time in June 2010. Summaries of previous MDC conferences may be found in Pauwels et al. (2004) and Leeflang et al. (2009).

Second, popular marketing-modeling textbooks devote more attention to time-series techniques (e.g. Hanssens/Parsons/Schultz 2001, Chapters 5 and 6; Leeflang et al. 2000, Section 17.3). Similarly, the recent Handbook of Marketing Decision Models (edited by B. Wierenga) contained a review chapter on Time-Series Models in Marketing (Dekimpe et al. 2008), as did the MSI Publication on Assessing Marketing Strategy Performance edited by C. Moorman and D. Lehmann (Dekimpe/Hanssens 2004). Third, recent issues of the Journal of Marketing (e.g. Pauwels/Weiss 2008, Trusov/Bucklin/Pauwels 2009), the Journal of Marketing Research (e.g. Deleersnyder et al. 2009; Slotegraaf/Pauwels 2008; Srinivasan/ Vanhuele/Pauwels 2010), Marketing Science (e.g. Nijs/Srinivasan/Pauwels 2007; Pauwels/Hanssens 2007), and the International Journal of Research in Marketing (Korhelis et al. 2008; Leeflang et al. 2009) all contain multiple manuscripts applying recent time-series techniques.

Rather than reviewing all of these papers, we will focus on the four developments identified in 2000 as key drivers for the growing use of time-series models in the future (Dekimpe/Hanssens 2000, Section 4). We will assess to what extent these have indeed materialized. To recap, these drivers are (i) the expanding size of marketing data sets, (ii) the accelerating rate of change in the market environment, (iii) the opportunity to study the marketing-finance relationship, and (iv) the emergence of internet data sources.

2. The Expanding Size of Marketing Data Sets

Marketing data sets have indeed expanded in multiple directions. First, the level of temporal aggregation has become smaller and smaller. For example, weekly store-level scanning data were used in Srinivasan/Pauwels/
Nijs (2008), daily data covering multiple years were analyzed in Pauwels/Weiss (2008), while Tellis/Chandy/Thaivanich (2000) even used hourly data to study the decay of advertising effects within the same day in the context of direct-response television advertising. These smaller aggregation levels allow one to re-visit prior empirical generalizations on the size of the carry-over effect, and challenge the conventional wisdom on what level of temporal aggregation should be preferred (see e.g. Tellis/Franse 2006).

Second, data have become available on more variables. Within a single category, it is not uncommon to see information on several performance metrics, and on price, advertising, feature, and display support for multiple brands (e.g. Nijs/Srinivasan/Pauwels 2007; Steenkamp et al. 2005). Such datasets allow for the specification of extended systems of equations (e.g. VAR models) that capture all possible inter-relationships among these variables across brands. Moreover, in many instances, data on multiple categories are available, which often resulted in second-stage analyses to explain cross-category differences (e.g. Nijs et al. 2001; Pauwels/Srinivasan 2004; Srinivasan et al. 2004; Steenkamp et al. 2005). This wealth of data added a more confirmatory (hypothesis-testing) flavor to the traditionally more data-driven approach of most time-series studies.

Third, data become available across multiple countries and even continents, which allow for the study of cross-country (e.g. economic, cultural) differences in the longitudinal behavior of key marketing constructs such as consumer confidence (Lemmens et al. 2005) and advertising spending (Deleersnyder et al. 2009).

Fourth, non-traditional measures are increasingly collected at regular time intervals. As predicted in Dekimpe/Hanssens (2000, p. 188), we now observe that attitudinal variables can be matched with transactional observations, as was recently done in Srinivasan/Vanhuele/Pauwels (2010) and by Fornell et al. (2010). The former added three mindset metrics (advertising awareness, consideration and liking) to vector-autoregressive marketing response models. These attitudinal variables were found to explain almost one-third of the variance in sales. Fornell et al. (2010), in turn, added customer satisfaction information as a key predictor of subsequent consumer spending growth.

Finally, data become available over longer time spans. Several recent applications have applied time-series techniques to study business-cycle phenomena. Dekimpe et al. (2004) studied the cyclical sensitivity of 24 consumer durables, Lamey et al. (2007) studied how private-label share behaves over the business cycle, and Deleersnyder et al. (2009) studied the cyclical sensitivity of advertising spending in 37 countries from all continents. In all instances, data spanning over 20 years (or more) of data were analyzed. In all three aforementioned studies, annual data were used. More recently, however, Gijssenberg et al. (2009) study the evolution of advertising and price effectiveness using monthly spending and performance data. Using 15 years of data for over 160 brands in 37 categories, they exploit the data explosion along several of the aforementioned dimensions: longer series, at a small level of temporal aggregation, across many brands and categories.

3. The Accelerating Rate of Change in the Market Environment

Marketing environments have become more turbulent, i.e. they are subject to more shocks that alter their evolution pattern. For example, in many industries, new products enter and exit more frequently, which affects the competitive setting faced by the incumbents. Time-series models are being used to diagnose quickly the implications of such entries and exits. Pauwels/Srinivasan (2004) used separate pre- and post-entry vector-autoregressive models to study the implications of a private-label entry. This approach assumes that all parameters are subject to change, which is statistically inefficient. Moreover, if there are n new-product introductions, this would require the consideration of n+1 different regimes.

Some alternative approaches have allowed for changing parameters in parts of the model, but not in others. Dekel/Heerde et al. (2002) allowed for a structural break in the deterministic part of a univariate time-series model, but assumed that the regular noise function was unaffected. Similar structural-break unit-root tests were also used in Nijs et al. (2001) and Steenkamp et al. (2005), among others. Extensions to allow for multiple breaks at unknown points in time were introduced in Kornelis et al. (2008). Van Heerde et al. (2007) used a multivariate vector-error correction specification to assess the impact of radical new innovations on the base performance of a set of incumbents.

An alternative method to capture the increasing turbulence in many markets is to use varying-parameter models. Pauwels/Hanssens (2007), for example, used recursive unit-root and VAR models to capture changing business performance and marketing spending regimes over time, as did Pauwels/Weiss (2008) and Yoo/Hanssens (2008). Two alternative approaches that have become quite popular to account for varying parameters over time are Kalman Filtering and Dynamic Linear Models (DLM). A review of Kalman filter models, along with a discussion of multiple marketing applications, is provided in Dekimpe et al. (2008, Section 11.3). DLM models, in turn, are used in van Heerde/Mela/Manchanda (2004), and Araman/Mela/Van Heerde (2007; 2010), among others. An insightful comparison of vector-autoregressive (VAR), vector-error-correction (VECM), Kalman Filter, and Dynamic Linear Models is given in Leeflang et al. (2009, Tab. 3).

From a more substantive point of view, the recent economic/financial crisis, and the turbulence it has caused in
many markets, has received considerable research attention as well. For example, the aforementioned business-cycle studies were partly inspired by the current recession. Other papers had to explicitly control for the potentially disturbing effects of the last few data points in their data sets (i.e. the observations corresponding to the recent crisis, as in Fornell et al. 2010), conducted longitudinal studies on the impact of rising gasoline prices on consumption (Mo et al. 2009), or considered the impact of currency crises on consumption smoothing (Putt/Padmanabhan 2009). Time-series techniques are very well suited to capture the dynamic implications of such crises.

4. The Opportunity to Study the Marketing-Finance Relationship

The finance discipline, with its inherent focus on growth and the valuation of assets over time, has long been a heavy user of analytical time-series techniques. Marketing, in contrast, tended to focus on customer attitudes and behavior, and was restricted in its use of time-series techniques because of a lack of long time series. Both issues have recently changed. First, there has been a growing recognition that marketing should be held accountable for its expenditures (Ambler 2003). Second, marketing started to have access to the longitudinal data sets needed to fully exploit the power of modern time-series techniques, as discussed in Section 2 above.

Because of these developments, we have seen an explosion of studies on the interface between marketing and finance, many of them using time-series methods. Indeed, marketing decision makers are increasingly aware of their role in creating shareholder value, which calls for an evaluation of the long-term effects of their actions on both product-market response and investor response (e.g. Joshi/Hanssens 2009). Recent reviews include Hanssens/Dekimpe (2008) and Sirivasan/Hanssens (2009). Commonly used econometric methods in this research stream include stock return models (for example, Mizik/Jacobson 2004) and persistence models (for example, Pauwels et al. 2004). In the recent Special Issue of the Journal of Marketing on “Marketing Strategy Meets Wall Street”, the editors conclude that much has been learned about how the investor community incorporates important marketing assets such as brand equity and consumer satisfaction in firm stock prices. On the other hand, we know less about investors’ ability to incorporate the stock-price impact of specific marketing actions such as price changes, advertising campaigns and new-product introductions (see Hanssens/Rust/Srivastava 2009).

Given the availability of these extensive reviews, we will not discuss individual contributions in detail. However, we point out that recently, marketing researchers in this area have expanded the scope of their inquiry to include not only the level response of financial metrics to marketing investments, but also the second moment, i.e. the risk or variance involved. For example, McAlister/Srinivasan/Kim (2007) find that marketing spending lowers the systematic risk of the firm, thereby lifting its market value. Fischer/Shin/Hanssens (2009) show that, while volatile marketing spending over time may increase marketing’s impact on revenue, it also increases revenue and earnings volatility, which raises the firm’s cost of capital. We expect more research on marketing’s impact on risk and volatility to appear in the near future.

5. The Emergence of Internet Data Sources

As a final contributing factor to the expected growth of marketing time-series applications, Dekimpe/Hanssens (2000) discussed the emergence of internet data sources. As shown below, internet-generated data have proved to be particularly useful in time-series applications. Conversely, the potential insights into consumer behavior, marketing activity and competitive conduct enabled by the internet will not be unlocked without the use of advanced time-series methods. As such, we believe the internet will remain one of the most important antecedents of the influence of time-series analysis in shaping marketing thought.

Deleersnyder et al. (2002) used conventional performance metrics for newspapers (i.e. advertising revenue and subscriptions), but modeled the introduction of a free internet version as a potential structural break in a univariate time-series model. They did not find much evidence of internet-caused cannibalization. Cannibalization concerns were also the inspiration in Biyalogorski/Naik (2003). Using Kalman Filtering, they estimated the extent of cannibalization from Tower Record’s Internet Sales Division on its conventional brick-and-mortar activities. These authors did not find much evidence of such cannibalization either.

Pauwels/Weiss (2008) studied the transition from a free internet business model to a fee-based model, i.e. from offering all content for free to charging for at least some of it. Using a data set with highly temporally disaggregated data (4 years of daily observations) from an online content provider, they investigated the role of various marketing actions (e.g. search-engine referrals, targeted e-mail offerings, ...) in this transition. They showed how, in their application, managers should focus their price promotions on generating new monthly subscriptions, rather than on generating new yearly contracts (as they currently do). E-mail and search-engine referrals, in contrast, were found to be good tools to generate such annual subscriptions. Of course, more research is needed to investigate whether these findings generalize to other content providers. Trusov/Bucklin/Pauwels (2009), in turn, studied the effect of word-of-mouth (WOM) marketing on member growth at an internet social networking site, and compared it with traditional marketing vehicles. WOM referrals were found to have a substantially longer carry-over effect than traditional marketing.
actions, and to produce considerably higher response elasticities. In another application to a web hosting company, Villanueva/Yoo/Hanssens (2008) linked the way in which a customer was acquired (through marketing or through WOM) to customer equity growth, and found that WOM customers contributed nearly twice as much to the long-term value of the firm.

Several recent papers use time-series techniques to address important marketing resource allocation questions that are raised by the advent of the internet. Wiesel/ Pauwels/Arts (2009) examine how on-line and off-line marketing activities influence customers’ progression in the purchase funnel for a B2B company. They show strong cross-effects between on-line and off-line marketing and response metrics which, when taken into account in marketing planning, can significantly enhance firm profitability. Joshi/Trusov (2009) examine the relative importance of social media vs. commercial media in generating buzz around a new product, for example a motion picture release. They define conditions under which studios should use different media for promoting their new movies. Shin/Hanssens/Gajula (2009) use a web-crawling algorithm to derive daily consumer sentiment readings about different brands in a category (MP3 players) and show how these metrics are leading indicators of retail price movements for these products.

Even though the modeling approaches in these studies were developed before the widespread use of the internet (for example VAR estimation followed by persistence modeling), they are demonstrating their unique aptitude in analyzing problems and creating insight opportunities generated by this medium. In particular, these methods are good at handling data generated at a smaller level of temporal aggregation, they succeed at quantifying the role of new metrics (such as searches, referrals and e-sentiments), and they can tackle new substantive problems such as marketing resource allocation across online vs. off-line media.

6. Some Final Thoughts

Over the last decade, time-series modeling in marketing has come a long way. First, while Dekimpe/Hanssens (2000; 2004) reviewed mainly VAR/persistence-modeling applications (which were also the focus of the first Marketing Dynamics Conference at Dartmouth), we now see a more even spread between that approach and techniques such as Kalman Filtering (see Dekimpe et al. 2008 for a review), spectral analyses (e.g. Bronnenberg/Mela/Boulding 2006; Lemmens et al. 2005; 2007; 2008), and DLM modeling (see Leeﬂang et al. 2009 for an in-depth discussion). Moreover, several of these models are now increasingly embedded in the Bayesian-estimation tradition (see e.g. Fok et al. 2006; Sismeiro/Myzak/Bul- lin 2009; Van Heerde et al. 2007). For tabular overviews of these contributions, see Dekimpe/Hanssens (2000) and Leeﬂang et al. (2009).

Second, because of additional data opportunities, the range of problems tackled has been broadened. Apart from the use of conventional response metrics such as sales elasticities, we now also see studies focusing on the impact of marketing investments on long-run brand equity, customer equity and firm value. In addition, the scope of marketing drivers is expanding rapidly, and includes, for example, mindset metrics and referral (WOM) metrics.

Third, time-series applications in marketing have long focused either on forecasting (involving “horse-race” competitions among several model specifications), or were descriptive in nature (for example, in a second-stage analysis, some hypotheses on brand- and/or category-differences were tested). Recently, we see more applications that have a normative objective as well. First, more studies use the parameter estimates of time-series models (e.g. VAR or VECM models) for policy simulations (e.g. Pauwels 2004; Van Heerde et al. 2007), In recent years, these policy simulations are increasingly accompanied by explicit tests for super-exogeneity, which are needed as a safeguard against the Lucas Critique (see van Heerde et al. 2005 for an in-depth discussion). Alternatively, state-space models are especially well suited to integrate econometric analyses with normative decision-making problems faced by managers. These involve a formal dynamic optimization, as opposed to a multitude of empirical what-if simulations. Recent applications include Naik/Raman (2003) and Naik/Raman/Winer (2005), as reviewed in Dekimpe et al. (2008, Section 13.3.3).

In conclusion, recent developments in marketing databases, computational power and the need for marketing accountability have created a research environment that is especially well suited for the use of time-series modeling. Because of these trends, we are confident that the importance of time-series models in marketing will continue to grow. We even feel the best is yet to come.

References