Intern. J. of Research in Marketing 29 (2012) 134–147

Contents lists available at SciVerse ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Modeling coexisting business scenarios with time-series panel data: A dynamics-based segmentation approach

Catarina Sismeiro [a,*], Natalie Mizik [b], Randolph E. Bucklin [c]

[a] Imperial College Business School, Imperial College, London, United Kingdom
[b] Kenan-Flagler School of Business, University of North Carolina, Chapel Hill, United States
[c] UCLA Anderson School of Management, United States

ARTICLE INFO

Article history:
First received in 26, November 2009 and was under review for 4½ months
Available online 17 March 2012

Area Editor: Harald J. Van Heerde

Keywords:
Sales force
Segmentation
Marketing-mix effectiveness
Econometric methods
Time-series modeling

ABSTRACT

At a given point in time, individual consumers may be in different stages of the product adoption or consumption cycle. As a result, different types of behavioral patterns may coexist within a single product market. Existing segmentation approaches typically do not address long-term dynamics in customer response and do not adequately capture this phenomenon. We develop an approach for modeling the coexistence of multiple dynamic behavioral patterns (business scenarios) within a single product market. We apply this approach to physician panel data on drug prescriptions and direct-to-physician promotions. We find markedly different responses across physician segments. For firms that track customer-level marketing activity and sales over time, market segmentation based on dynamic scenarios can provide a new tool for efficient targeting. The proposed approach is straightforward to implement and is scalable to very large samples and continuous testing.

1. Introduction

The concept of business scenarios in the context of time-series modeling was first discussed by Dekimpe and Hanssens (1999). In their study, the authors presented four possible alternative scenarios for a given market (Fig. 1): "business as usual" (in which both performance and marketing variables are stationary), "escalation" (in which only the marketing variables are evolving), "hysteresis" (in which only the performance measure is evolving), and "evolving business practice" (with evolving performance and marketing variables). Depending which scenario is detected in the market under analysis, alternative formulations for vector autoregressive models would be appropriate, which, when estimated, lead to different strategic conclusions.

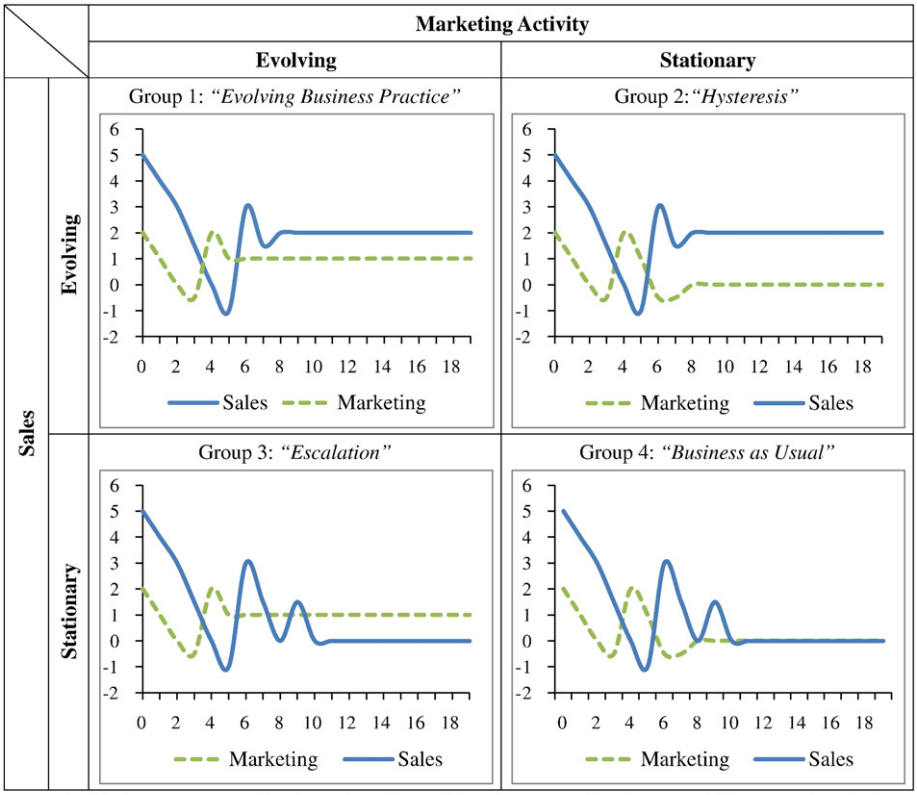
In this paper, we propose that, due to differences in customers and firm behavior, multiple business scenarios may coexist within a single product market. Just as products, industries, and markets may be at different stages of their life cycles, individual consumers may be at different stages of their consumption life cycles and may be subject to targeted marketing actions that differ significantly (in terms of intensity and scope) from those aimed at other consumers. Several prominent theories provide support for such a phenomenon. For example, new product diffusion and adoption theories rely on the existence of distinct consumer segments that learn about and adopt the product at different points in time (Bass, 1969; Rogers, 2003). Firms have access to a broad set of tools to customize marketing actions at the individual level in terms of intensity, message, and even the medium used.

Despite the likely coexistence of multiple business scenarios in a market, the modeling of distinct dynamic consumer responses and dynamic firm behavior has not been incorporated into existing segmentation methodologies. With this paper, we develop and illustrate an approach that uses time-series panel data to investigate whether different dynamic business scenarios might exist concurrently across a firm's customer base and what the data imply strategically. If these business scenarios do coexist, some customers (or segments) might present the firm with "business as usual," while others present the perils of "escalation" or the opportunities of "hysteresis" or "coevolution" scenarios. To study this phenomenon, we relax the common, though typically implicit, assumption that the time-series properties of the data are uniform across panelists.

Our approach allows us to effectively identify distinct dynamic patterns at an individual level and to investigate the differences in response dynamics across customers. It consists of two steps. In the first step, we test for the order of integration in the data using unit-root tests at the disaggregate level. We conduct two sets of unit-root tests, one for the outcome variable and one for the marketing covariates. Then, using the test results, we assign individuals to one of four groups. In the second step, we specify an appropriate panel vector autoregressive model (PVAR) for each group, and we estimate

* Corresponding author. Tel.: +44 02075949108.
E-mail address: c.sismeiro@imperial.ac.uk (C. Sismeiro).

0167-8116/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijresmar.2011.08.005



Legend: This figure presents stylized examples of "own" impulse response functions (IRFs) for sales and marketing efforts. A sales IRF represents the impact of a sales shock in period zero on future sales. The marketing IRF represents the impact of a marketing shock in period zero on future marketing effort. When the IRFs converge to zero, the series are stationary; IRFs converging to a non-zero value mean that the series are non-stationary (i.e., a unit root is present). We note that although there might be a relation between the two series (e.g., the evolving behavior of sales could be due to the evolving behavior of marketing), the first-stage classification into business scenarios does not require such cause-effect relationships. This classification is important for the second-stage modeling phase because the specification of the VAR models depends on the stationary or evolving nature of the series.

Fig. 1. The four basic business scenarios—stylized examples of impulse response functions.

separate PVAR models to investigate potential differences in response dynamics across groups.

To test our approach, we use individual physician-level time-series panel data on prescribing and direct-to-physician promotion (DTP) from a pharmaceutical prescription drug market. Differences across doctors in age, experience, practice size and type, risk aversion and adoption timing could lead to differences in the level of time-series integration for individual prescribing (i.e., prescription data series might be evolving or stationary depending on these factors). Different levels of marketing activity targeted by pharmaceutical companies across physicians can also give rise to evolution versus stationarity of DTP activity at the individual level. Thus, different dynamic business scenarios might be present across physicians.

Empirically, we find that multiple dynamic scenarios coexist within a single drug market, each with markedly different response magnitudes and patterns. We assess the performance of the proposed method relative to alternative segmentation approaches and present evidence that common segmentation variables used by the industry fail to detect meaningful response differences across groups. Tests using cross-sectional and longitudinal holdout samples support the superiority of the proposed dynamic business scenario-based segmentation, which produces not only better in-sample fit but also better out-of-sample fit. Hence, we believe that the proposed segmentation approach provides a useful new tool for enhancing the productivity of marketing resources through better targeting at the individual level.

2. Motivation

In markets where sales and/or marketing activity may be evolving, econometricians have emphasized the importance of handling potential nonstationarity in time-series data. These "persistence modeling" methods (1) employ unit-root tests to ascertain the stationary versus evolving nature of the data and (2) estimate appropriate vector autoregressive (VAR) models given the integration level of the data to assess the market dynamics and long-run effects. The application of VAR-based persistence models to marketing data has yielded key insights about dynamic response in a series of studies conducted on aggregate-level data (e.g., Dekimpe & Hanssens, 1995, 1999; Bronnenberg, Mahajan, & Vanhonacker, 2000; Pauwels, Hanssens, & Siddarth, 2002; Horváth & Franses, 2003). Dekimpe and Hanssens (1999), for example, have demonstrated how time-series methods applied to aggregate historical data can identify different dynamic business scenarios: business as usual, escalation, hysteresis, and evolving business practice (or coevolution). Implications for return on marketing spending, marketing strategy, and profitability can differ dramatically across these scenarios. The possibility that different dynamic business scenarios may coexist within a single market, however, has not been examined.

Another branch of marketing science has extensively studied disaggregate-level data in consumer panels, based largely on UPC scanner data for packaged goods. This stream has emphasized the study of short-run response to marketing activity across individual

decision makers and the potential for segmentation in individual-level response (e.g., Kamakura & Russell, 1989; Bucklin, Gupta, & Siddarth, 1998; Kamakura & Wedel, 2000). Despite the emphasis on short-run response in individual-level studies, a number of authors working with individual scanner panel data have also attempted to examine dynamic effects or long-run changes in response parameters (e.g., Papatla & Krishnamurthi, 1996; Mela, Gupta, & Lehmann, 1997). These studies have found empirical evidence for long-run changes in the nature of consumer response, but they have not investigated the time-series properties of the underlying series.

Pauwels et al. (2002) applied VAR-based persistence modeling to UPC scanner data and found that evolution (i.e., the presence of unit roots) was uncommon in aggregate-level scanner data, a phenomenon they attributed to the maturity of most products studied. The link between product maturity and a business scenario may also explain the results of Bronnenberg et al. (2000). In their study of new product sales and distribution, the authors analyzed aggregate-level data for the launch phase of the ready-to-drink tea category and found empirical patterns consistent with coevolution for sales and distribution.

We contend that just as different products might be at different stages of the life cycle, so might individual consumers or markets. Indeed, most theories of new product diffusion and new product adoption rely on the presence of consumer segments that learn about and adopt the product at different times (Bass, 1969; Rogers, 2003). Dekimpe, Parker, and Sarvary (2000), for example, showed that technological products may be at different points of the life cycle across countries depending both on the time of the initial adoption and the level of current adoption. Thus, product sales can have different dynamics across countries. Parsons (1975), Lieberman (1987), and Osinga, Leeflang, and Wieringa (2010) reported that marketing-mix responsiveness can vary across the stages of the life cycle. However, when using aggregate-level time-series data for dynamic scenario and response analysis, the investigator is required to treat all the firm's customers within a single scenario classification, which is a significant limitation. If analysis is restricted to aggregate-level data, there is no opportunity to observe heterogeneity in the dynamics of firm and customer behavior.

Pauwels et al. (2005) explicitly called for more research into the problems associated with cross-sectional heterogeneity in VAR-based models. Lim, Currim, and Andrews (2005) presented a VAR-based analysis of scanner panel data that took a step in this direction. They first assigned panelists to segments determined a priori based on panelists' brand loyalty (loyals versus switchers) and category usage rates (heavy versus light users). The authors then aggregated data to the segment level. The estimated segment-specific VARs revealed differences in response and adjustment period across segments. Perhaps due to the mature product categories involved, unit roots showed the aggregate time series to be almost universally stationary (similar to the results of Pauwels et al., 2002). This constrained the analysis of response dynamics to differences within the "business as usual" scenario.

In addition to preventing the identification of coexisting business scenarios, estimating aggregate-level VAR models presents another limitation: the difficulty of handling the typically large number of endogenous variables and/or lag effects. A lack of degrees of freedom has led researchers to impose a variety of restrictions on the structure of VAR models to limit the number of parameters. Panel data can be used to increase the degrees of freedom under appropriate pooling assumptions (e.g., fixed effects, Horváth & Wieringa, 2008). Thus, PVAR models offer the advantage of imposing fewer restrictions on lag structure and explanatory variables.

3. Modeling approach

Granger and Newbold (1974) first highlighted the "spurious regression" problem, in which unrelated unit-root series appear related with a very high probability if conventional estimation methods are employed. This finding stimulated interest in appropriately modeling time-series dynamics, and it led to a multitude of research studies testing dynamic patterns and addressing the advantages and disadvantages of various modeling approaches to control for the dynamic properties of the series (e.g., Plosser & Schwert, 1977; Stock & Watson, 1988; Nelson & Kang, 1984). More recently, researchers in econometrics have begun scrutinizing the methods involved in the analysis of panel data and examining how these data can be used to improve the study of time-series dynamics (e.g., Baltagi & Kao, 2000). New methodological advancements in panel data analysis provide a foundation for researchers in marketing to integrate the two streams of research: (1) the study of segmentation with panel data and (2) the study of long-run dynamic effects with (potentially) evolving time-series data.

Our approach builds on these advancements. Methodologically, our approach is based on unit-root tests and PVAR modeling techniques. The proposed method is applicable to data in which different dynamic business scenarios are thought to coexist. It is not limited to individual-level data but can be applied where cross-sectional units are, for example, markets or geographical regions.

3.1. The proposed procedure

Our approach proceeds as follows. First, we conduct unit-root tests for the outcome and marketing variables for each panelist at the individual level.[1] Based on the unit-root test results, we classify each panelist into one of the four dynamic business scenarios: evolving business practice (coevolution), hysteresis, escalation, or business as usual. We then specify the PVAR models for each group such that variables enter either in levels (if they were deemed stationary) or in differences (if evolving), depending on the unit-root test results (e.g., Campbell & Perron, 1991). Fig. 1 depicts the four groups and the corresponding stylized representations of the impulse response functions (IRFs) of marketing activity and sales response.

In Fig. 1, Group 1 is the "evolving business practice" scenario. Here, all variables are evolving, and all enter the PVAR model in first differences. Group 4 is the "business as usual" scenario: all variables are stationary and enter the PVAR model in levels. Group 2 is the "hysteresis" scenario (evolving in sales, stationary in marketing activity), and Group 3 is the "escalation" scenario (stationary in sales, evolving in marketing activity). Both Group 2 and Group 3 are modeled with mixed PVARs. The outcome variables (e.g., sales) enter the model in levels for Group 3 and in differences for Group 2. Marketing activity variables enter in levels for Group 2 and in differences for Group 3. To address potential endogeneity in marketing activity and response, we treat all variables for all groups as endogenous.

Next, we ascertain the presence of cross-sectional heterogeneity in the data. We conduct Hausman (1978) specification tests for the presence of individual-specific fixed effects. When necessary, the PVAR model is specified to incorporate fixed effects. All PVAR models include time-specific indicator variables to control for unobserved time-specific effects and to ensure that the models are robust to structural changes. The time-specific intercepts also serve as equilibrium-correction terms (e.g., Theil, 1961, Clements & Hendry, 1999).

We test each PVAR model for the appropriate number of lags using the Schwarz criterion (SC). After estimation of the appropriate PVAR models for each scenario, we obtain IRFs and gauge the effect of marketing effort on the performance variable. To provide a benchmark for our results, we estimate a pooled PVAR model across all panelists, in levels and in differences. (Later, we discuss other benchmark alternatives, including a pooled Bayesian random-effects formulation.)

[1] A number of different unit-root tests has been advanced in the time series literature (Maddala, 1992). Our segmentation approach relies on the outcomes of a unit-root test but does not depend on the specific unit-root test employed.

Finally, we assess the performance of the dynamic scenario-based segmentation in cross-sectional and longitudinal holdout samples.

4. Research on pharmaceutical marketing dynamics

Recent studies on pharmaceutical marketing have used different data as well as different methods to investigate the effects of marketing activity on physicians' prescribing decisions (see Kremer, Bijmolt, Leeflang, & Wieringa, 2008, Leeflang & Wieringa, 2010, Manchanda et al., 2005 for comprehensive reviews). Although some of these studies incorporate dynamics, to some extent, in their modeling, we seldom see the simultaneous investigation of heterogeneity and the time-series properties of the data. For example, using aggregate data, Dekimpe and Hanssens (1999) employed VAR models to investigate the long-run effects of changes in marketing activity and to draw implications for profitability. Narayanan, Manchanda, and Chintagunta (2005) also used aggregate-level data on new prescriptions to investigate temporal differences in the role of detailing and other marketing expenditures. These studies did not deal with potential heterogeneity in the data.

Studies conducted at the disaggregate level typically investigate the role of heterogeneity in response across physicians. For example, Manchanda and Chintagunta (2004) used a Poisson model to study the effect of marketing activity on the number of prescriptions written by physicians on a quarterly basis. Manchanda, Rossi, and Chintagunta (2004) extended the analysis to incorporate potential endogeneity in prescribing and marketing activity but did not investigate the time-series properties of the data. Mizik and Jacobson (2004) employed fixed-effects instrumental variable estimation to address both heterogeneity and endogeneity in physician response to direct-to-physician marketing activities. Finally, Narayanan and Manchanda (2005), Janakiraman, Dutta, Sismeiro, and Stern (2008), and Janakiraman, Sismeiro, and Dutta (2009) used physician panel data to estimate an individual-level model of prescription choice within a therapeutic class. Although these studies did not investigate the evolving nature of the corresponding prescription and marketing mix series, these authors find significant differences across physicians in response to pharmaceutical detailing, in choice state-dependence, and in learning rates.

In sum, existing disaggregate-level pharmaceutical models have accounted for some dynamic effects (e.g., carryover effects, Bayesian learning, and lagged dependent variables), but they have not provided a full analysis of the time-series properties of the data. In addition, to the best of our knowledge, previous research has not investigated the time-series properties of the data at the individual physician level and has not yet attempted to model the coexistence of multiple dynamic business scenarios.

5. Empirical application

5.1. Data

Our data come from an anonymous panel of 5000 U.S. physicians tracked monthly over a period of approximately 2 years (October 2001–August 2003) for a single drug. Due to confidentiality requirements, the identity of the drug and the company are masked. For each physician in the panel, the data set includes general demographic information (age, gender, and year of graduation), the number of new prescriptions written, and the number of sales calls (details) and samples received each month.

To focus on decision makers who were at least minimally involved with the product, we have excluded physicians with fewer than two prescriptions and with less than one sales call per year. To avoid the undue influence of outliers, we also removed physicians with extremely high levels of prescribing activity (top one-half of 1%). According to the collaborating firm, the activity assigned to these physicians might reflect large group practices or hospitals rather than representing individual doctors. Our final study sample consisted of 3942 physicians (79% of the original sample). We randomly assigned two-thirds (2628) of physicians to the estimation sample and the remaining one-third (1314) to a holdout sample.

Table 1
Summary of descriptive statistics.

	Mean	Standard error of the mean	Min	Lower 10%	Median	Upper 90%	Max
New prescriptions	3.33	.05	.15	.79	2.55	6.77	21.02
Details	3.09	.03	.09	.91	2.91	5.52	9.91
Samples	15.71	.22	.22	3.09	11.65	33.76	93.87
Physician national deciles	6.17	.03	4.00	4.00	6.00	9.00	10.00
Years since graduation	20.09	.15	3.00	8.00	20.00	32.00	58.00

Legend: Descriptive statistics for prescriptions, details, and samples reflect monthly values per physician and are computed across the 3942 physicians in the study sample. Physicians were assigned a national decile by the company based on their prescribing volume. A decile assignment of 7, for example, indicates that the doctor is in the 70th percentile for new prescription volume. The variable "years since graduation" uses 2001, the first year of our sample, as the reference year.

Table 1 presents a summary of descriptive statistics for the final sample of 3942 physicians, including both estimation and holdout samples. The mean number of new prescriptions written per month was 3.33,[2] the mean number of details was 3.09, and the mean number of samples received was 15.71. The range statistics show that our data sample includes a wide cross-section of physicians, with panelists varying significantly in experience, level of prescribing, and attention from pharmaceutical sales representatives. The variable of 'years since graduation' is based on the number of years between the physician's graduation year and the first year of our sample period. Physicians were also assigned a national decile value by the company based on their past prescribing volume. A decile assignment of 7, for example, indicates that the doctor is in the 70th percentile for new prescription volume. The average decile value was 6.17. According to the collaborating firm, the characteristics of our sample are representative of the general population of physicians prescribing this drug.

Fig. 2 graphs the aggregate time series of key variables over the study period. An examination of the time-series patterns in Fig. 2 suggests possible evolution in the aggregate series for prescriptions, detailing, and sampling. Augmented Dickey–Fuller (ADF) tests performed on the aggregate data shown in Fig. 2 are all close to the 95% critical value, with prescriptions marginally evolving and stationary marketing activity.

5.2. Unit-root tests

As the first step of our approach, we conducted individual-level ADF unit-root tests on all 3942 physicians in the sample. We tested prescription and marketing activity variables (details and samples).[3] We relied on the ADF test to classify our panelists into the evolving or stationary conditions for several reasons. First, previous research has shown that the data-generating process for drug prescriptions has a high autoregressive order (e.g., Mizik & Jacobson, 2004 report six significant lags for the first-differenced prescriptions series for all three drugs in their study), and our data exhibit significant lag structure. The ADF test has been shown to perform well in Monte Carlo studies under these

[2] It is important to note that the prescription data at the individual level are not count data. Because prescription size (i.e., the number of pills) varies significantly across prescriptions, prescriptions are tracked and recorded as the number of "standardized" prescription units issued in a given month. In addition, caution is needed in the interpretation and use of averages for evolving variables because population means are not defined for I(1) processes (evolving series).

[3] As a sensitivity test, we replicated our analysis using the KPSS test (Kwiatkowski, Phillips, Schmidt, & Shin, 1992). We report the results in the Sensitivity analyses section.

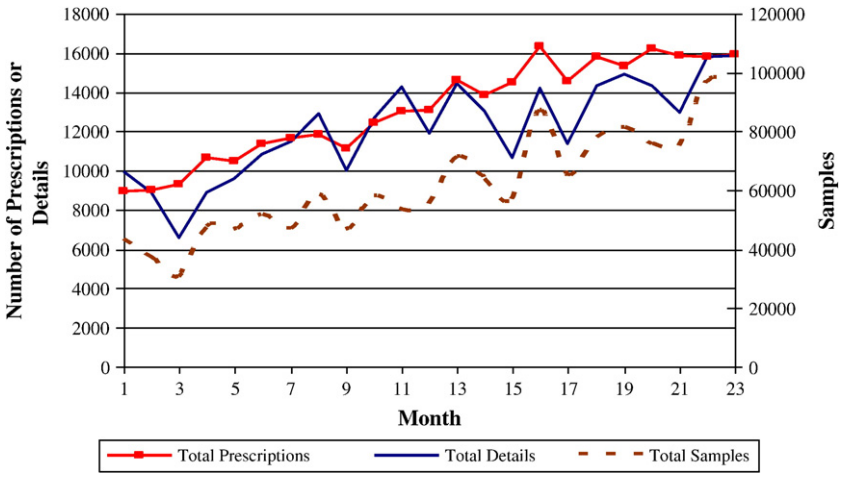


Fig. 2. Number of new prescriptions, details, and samples by month.

conditions because it increases in power as the number of lags increases (Harris, 1992; Haug, 1996). Second, the null hypothesis in ADF tests is the presence of a unit root, and the consequences of modeling a stationary process as having a unit root are less severe than the opposite. The negative consequences of ignoring unit roots involve false inferences and incorrect conclusions, whereas treating a stationary process as a unit root (and potentially overdifferencing the data) only leads to a decrease in the efficiency of the estimates (Plosser & Schwert, 1977). Thus, the negative consequences of underdifferencing significantly outweigh the consequences of overdifferencing.

Finally, we chose the ADF because it has been shown in simulation studies to have very close, albeit slightly more conservative, test statistics for count data (our detailing and sampling data are integer counts) and to perform well in short time-series data (i.e., small sample sizes similar to our data). Hellström (2001), for example, conducted a large-scale simulation study and reported small sample distributions for the ADF test statistic for count data. He specifically examined short time series ($T=25$) and derived approximation equations to enable calculation of critical values for any values of time-series length (T) and drift. Given our data characteristics, the Hellström (2001) corrections suggest very small deviations from the classic ADF statistic.

The ADF tests we performed for each physician included an intercept and a trend. The number of lags used for each physician was selected based on the SC. Given the test results, we classified each physician into one of four groups. Panelists were classified as evolving in the number of new prescriptions if we failed to reject the null hypothesis of a unit root at the .05 significance level. Otherwise, they were classified as stationary. Evolving prescription behavior was found in the physician-level data for 22% of the doctors, whereas 78% was stationary. For the marketing activity variables, we classified panelists as evolving when, for at least one of the marketing variables (detailing or sampling), we failed to reject the null hypothesis of a unit root (in such cases, all marketing variables are modeled in differences). Marketing activity was classified as evolving for 52% of the doctors and as stationary for 48%.[4] In addition, we found that each of the four business scenario groups was populated by a large, albeit unequal, number of physicians. Group 1 (coevolution) had 12%, Group 2 (hysteresis) had 10%, Group 3 (escalation) had 40%, and Group 4 (business as usual) had 38% of the physicians.[5]

Table 2 presents the sample means for the characteristics of the resulting four business scenario groups. We see no significant differences in the number of monthly new prescriptions or experience (as measured by years since graduation). Table 2 also presents the national decile rankings for physicians assigned to each of the four groups. Pharmaceutical firms use national decile rankings of prescription volume as a key component in targeting marketing efforts to physicians (Manchanda et al., 2004). Table 3 shows the percentage of each group's membership corresponding to the seven different decile ratings available (numbered 4 through 10). For example, in Group 1, about 21.2% of the doctors fall into decile 4, whereas 2.7% fall into decile 10. We see few differences in decile rankings across the four groups.

The results from Tables 2 and 3 suggest that readily available descriptive information does not predict the classifications produced by the unit-root tests. If the segmentation based on the unit-root classifications carries through to meaningful differences in response dynamics, it could open new opportunities for segmentation and targeting. For example, at first glance, Group 2 (hysteresis) would seem to be the most interesting and profitable for marketers to identify and target. Indeed, it may take only a temporary increase in marketing activity to produce a permanent increase in new prescription activity. We note, however, that the unit-root classifications (evolving outcomes, stationary marketing) simply make it possible for hysteresis to occur but do not ensure that it occurs (evolution in performance variables might not be the result of specific marketing actions but might be derived from other phenomena, such as learning from usage).

In contrast, Group 3 panelists may be the least attractive because the scenario involves permanent increases in marketing effort with only temporary increases in prescribing. Incidentally, Group 3 is the largest group in our data sample. Its size might partly reflect the recent intensely competitive environment in the pharmaceutical industry. This phenomenon has been referred to as the "PSR arms race," in which firms field large numbers of new pharmaceutical sales representatives (PSRs).

5.3. PVAR specifications

The specifications of the PVAR models for the four groups are given below (to keep the model presentation tractable, we have

[4] This approach is parsimonious in that it holds the number of resulting segments to four. If desired, this approach can be extended to all possible combinations of evolving versus nonevolving behavior for all variables. In our sample, there are approximately 2047 physicians with evolving marketing mix (those in Group 1 and Group 3). For 1174 of these physicians (i.e., 57%), only detailing is evolving, for 427 physicians (21%) both marketing variables are evolving, and for the remaining 446 physician (22%), the only evolving marketing variable is samples. A very similar pattern was observed within Group 1 and Group 3.

[5] In previous studies using VARX specifications in the presence of evolving series, researchers excluded the cross-sections with non-stationary series from their analyses. For example, Horváth, Leeflang, Wieringa, and Wittink (2005) excluded two stores showing evolving series for some of the variables and used the 24 remaining stores with stationary series in their fixed-effects VARX model.

Table 2
Summary of descriptive statistics by group.

	Group description		Averages					Number of physicians
	New prescriptions	Marketing	Number of new prescriptions	Details	Samples	National deciles	Years since graduation	
Group 1 Evolving business practice	Evolving	Evolving	3.62	3.02	15.30	6.18	19.15	476
Group 2 Hysteresis	Evolving	Stationary	3.66	2.92	15.69	6.24	19.75	375
Group 3 Escalation	Stationary	Evolving	3.27	3.08	15.33	6.16	20.26	1571
Group 4 Business as usual	Stationary	Stationary	3.22	3.16	16.23	6.17	20.36	1520

dropped the group subscript). In Group 1 (coevolution), all variables are evolving, and therefore all series are specified in differences. In Group 4 (business as usual), all series are stationary, and thus all equations are specified in levels. Groups 2 and 3 are mixed PVAR models with evolving series entering in differences and stationary entering in levels.

Group 1 (coevolution)[6]:

$$\begin{bmatrix}\Delta Q_{i,t}\\ \Delta D_{i,t}\\ \Delta S_{i,t}\end{bmatrix} = \begin{bmatrix}\alpha_{0Q}\\ \alpha_{0D}\\ \alpha_{0S}\end{bmatrix} + \sum_{j=1}^{p}\begin{bmatrix}\beta_{11}^{j} & \beta_{12}^{j} & \beta_{13}^{j}\\ \beta_{21}^{j} & \beta_{22}^{j} & \beta_{23}^{j}\\ \beta_{31}^{j} & \beta_{32}^{j} & \beta_{33}^{j}\end{bmatrix} \times \begin{bmatrix}\Delta Q_{i,t-j}\\ \Delta D_{i,t-j}\\ \Delta S_{i,t-j}\end{bmatrix} + \begin{bmatrix}\sum_{t=1}^{T-1}\delta_{Qt}\\ \sum_{t=1}^{T-1}\delta_{Dt}\\ \sum_{t=1}^{T-1}\delta_{St}\end{bmatrix} \times \Delta Time(t) + \begin{bmatrix}\varepsilon_{Q_{i,t}}\\ \varepsilon_{D_{i,t}}\\ \varepsilon_{S_{i,t}}\end{bmatrix}, \quad (1)$$

Group 2 (hysteresis):

$$\begin{bmatrix}\Delta Q_{i,t}\\ D_{i,t}\\ S_{i,t}\end{bmatrix} = \begin{bmatrix}\alpha_{0Q}\\ \alpha_{iD}\\ \alpha_{iS}\end{bmatrix} + \sum_{j=1}^{p}\begin{bmatrix}\beta_{11}^{j} & \beta_{12}^{j} & \beta_{13}^{j}\\ \beta_{21}^{j} & \beta_{22}^{j} & \beta_{23}^{j}\\ \beta_{31}^{j} & \beta_{32}^{j} & \beta_{33}^{j}\end{bmatrix} \times \begin{bmatrix}\Delta Q_{i,t-j}\\ D_{i,t-j}\\ S_{i,t-j}\end{bmatrix} + \begin{bmatrix}\sum_{t=1}^{T-1}\delta_{Qt}\times\Delta Time(t)\\ \sum_{t=1}^{T-1}\delta_{Dt}\times Time(t)\\ \sum_{t=1}^{T-1}\delta_{St}\times Time(t)\end{bmatrix} + \begin{bmatrix}\varepsilon_{Q_{i,t}}\\ \varepsilon_{D_{i,t}}\\ \varepsilon_{S_{i,t}}\end{bmatrix}, \quad (2)$$

Group 3 (escalation):

$$\begin{bmatrix}Q_{i,t}\\ \Delta D_{i,t}\\ \Delta S_{i,t}\end{bmatrix} = \begin{bmatrix}\alpha_{iQ}\\ \alpha_{0D}\\ \alpha_{0S}\end{bmatrix} + \sum_{j=1}^{p}\begin{bmatrix}\beta_{11}^{j} & \beta_{12}^{j} & \beta_{13}^{j}\\ \beta_{21}^{j} & \beta_{22}^{j} & \beta_{23}^{j}\\ \beta_{31}^{j} & \beta_{32}^{j} & \beta_{33}^{j}\end{bmatrix} \times \begin{bmatrix}Q_{i,t-j}\\ \Delta D_{i,t-j}\\ \Delta S_{i,t-j}\end{bmatrix} + \begin{bmatrix}\sum_{t=1}^{T-1}\delta_{Qt}\times Time(t)\\ \sum_{t=1}^{T-1}\delta_{Dt}\times\Delta Time(t)\\ \sum_{t=1}^{T-1}\delta_{St}\times\Delta Time(t)\end{bmatrix} + \begin{bmatrix}\varepsilon_{Q_{i,t}}\\ \varepsilon_{D_{i,t}}\\ \varepsilon_{S_{i,t}}\end{bmatrix}, \quad (3)$$

[6] We tested for and did not find cointegration in Group 1 in our data.

Group 4 (business as usual):

$$\begin{bmatrix}Q_{i,t}\\ D_{i,t}\\ S_{i,t}\end{bmatrix} = \begin{bmatrix}\alpha_{iQ}\\ \alpha_{iD}\\ \alpha_{iS}\end{bmatrix} + \sum_{j=1}^{p}\begin{bmatrix}\beta_{11}^{j} & \beta_{12}^{j} & \beta_{13}^{j}\\ \beta_{21}^{j} & \beta_{22}^{j} & \beta_{23}^{j}\\ \beta_{31}^{j} & \beta_{32}^{j} & \beta_{33}^{j}\end{bmatrix} \times \begin{bmatrix}Q_{i,t-j}\\ D_{i,t-j}\\ S_{i,t-j}\end{bmatrix} + \begin{bmatrix}\sum_{t=1}^{T-1}\delta_{Qt}\\ \sum_{t=1}^{T-1}\delta_{Dt}\\ \sum_{t=1}^{T-1}\delta_{St}\end{bmatrix} \times Time(t) + \begin{bmatrix}\varepsilon_{Q_{i,t}}\\ \varepsilon_{D_{i,t}}\\ \varepsilon_{S_{i,t}}\end{bmatrix}. \quad (4)$$

In the above equations, $Q_{i,t}$ denotes the number of new prescriptions by physician i in month t, and $D_{i,t}$ and $S_{i,t}$ denote, respectively, the number of details and samples received by physician i in month t. Δ denotes the difference operator. The p matrices of the parameters β depict the effects of past prescriptions, detailing, and sampling. The number of lags, p, in each group is selected to minimize the SC in the PVAR estimation (in each group, SC might lead to the use of a different number of lags). We use Hausman specification tests to assess the presence of fixed effects in our level equations. We denote individual-specific intercepts as α_i and uniform intercepts as α_0. The subscripts Q, D, and S identify the intercepts for prescriptions, detailing, and samples, respectively. $\varepsilon_{Qi,t}$, $\varepsilon_{Di,t}$, $\varepsilon_{Si,t}$ are the error terms.

To model the time-specific effects, we adopt a time indicator specification commonly employed in the time-series models of regime changes. Our indicator variable *Time(t)* is equal to zero before time period t and unity from time t on. In the difference equations, this time indicator variable is differenced ($\Delta Time(t)$) and takes the familiar form of a standard time-period dummy variable equal to 1 if the time period is t and zero otherwise. This specification of time indicators accommodates a very general structure of exogenous shocks with various dynamic effects and allows for robust estimation without prior knowledge of where structural breaks might be occurring (e.g., Clements & Hendry, 1999).

Table 3
Percentage of physicians by group and decile.

Decile	Group 1 (%)	Group 2 (%)	Group 3 (%)	Group 4 (%)
4	21.2	24.0	23.7	24.1
5	18.3	16.0	17.2	18.0
6	18.7	17.1	19.0	18.1
7	18.1	15.5	15.7	13.8
8	13.2	13.1	12.0	12.4
9	7.8	11.2	8.8	9.7
10	2.7	3.2	3.6	3.8
Total	100.0	100.0	100.0	100.0

Table 4
Summary of the estimated models.

	Variable specification		Period effects		Cross-sectional fixed effects		Number of lags
	Prescriptions	Marketing	Prescriptions	Marketing	Prescriptions	Marketing	
Segmented							
Group 1 Evolving business practice	Difference	Difference	ΔTime(t)	ΔTime(t)	No	No	7
Group 2 Hysteresis	Difference	Level	ΔTime(t)	Time(t)	No	Yes	4
Group 3 Escalation	Level	Difference	Time(t)	ΔTime(t)	Yes	No	8
Group 4 Business as usual	Level	Level	Time(t)	Time(t)	Yes	Yes	5
Pooled							
Differences	Difference	Difference	ΔTime(t)	ΔTime(t)	No	No	11
Levels	Level	Level	Time(t)	Time(t)	Yes	Yes	7

Legend: Time(t) represents an indicator variable that is equal to zero before time period t and unity from time t on; ΔTime(t) represents the first difference of Time(t) and is the same as monthly dummy variables. The number of lags for each final model was selected based on the Schwarz criterion. The variables entering the models in differences do not include cross-sectional fixed effects because these are canceled out through first-differencing.

5.4. PVAR model estimation and testing

Testing, estimation, and model selection for the PVARs proceeded as follows. For each group, we first determined the appropriate number of lag terms to include (p) by minimizing the SC.[7] Table 4 provides a summary of the specifications for each of the PVAR models estimated. The last column of the table shows that the selected PVAR models for groups 1, 2, 3, and 4 had lag lengths of 7, 4, 8, and 5, respectively. The final pooled PVAR model in levels had 7 lags, and the pooled PVAR model in differences had 11 lags.

Next, we conducted group-specific Hausman specification tests and, consistent with prior research (Mizik & Jacobson, 2004), detected the presence of significant fixed effects in the levels data.[8] As such, our PVAR models differ across the four groups in that we have an individual-level intercepts specification for all equations in Group 4, for marketing activity equations in Group 2, and for the prescribing equation in Group 3. Due to the differencing of evolving series, any physician-specific fixed effects are eliminated, and the remainder of the equations are specified with uniform intercepts (i.e., without fixed effects).

Table 5 reports SC values for each group model as well as the pooled models. The group-level PVAR models are estimated separately for each group, but we also report the aggregate SC values to facilitate comparison with the pooled PVAR models. Conversely, the pooled PVAR models are estimated across all physicians in the estimation sample. Again, to facilitate comparison, we report the SC values produced by the pooled models specifications within each physician group. The results in Table 5 show that the group-level models are preferred to the pooled models in all cases. The group-level PVAR models have the minimum SC values both overall and for each group taken separately.

We also compared the performance of the group-level models with the pooled models using cross-sectional and longitudinal holdout samples. Table 6 reports results for the cross-sectional holdout. Using the estimates of PVAR parameters from the estimation sample, we predict the number of new prescriptions for physicians in the holdout sample. The table provides two measures of predictive validity, root mean squared error (RMSE) and mean absolute deviation (MAD). The group-level PVAR models produce lower RMSE and MAD values than pooled PVAR models, with one exception. For Group 3, the MAD for the group-level model is equal to the corresponding MAD of the pooled model in differences. The RMSE, however, still favors the group-level model for the physicians in Group 3.

Table 7 presents results for the longitudinal holdout tests. We created a longitudinal holdout sample containing the last 2 months of the data for each physician. We re-estimated our models using the remaining data points (i.e., excluding the holdout time periods for each individual) and used the estimates for prediction. Again, we see clear evidence of superior performance for the dynamic scenario-based segmentation because it generates better quality forecasts.

In sum, the model comparison results consistently support the group-level PVAR models over the pooled PVAR models. For the estimation sample, the SC is minimized by the group-level modeling approach, for all four groups separately and for the estimation sample taken as a whole. For the holdout samples, the forecast errors, as measured by RMSE and MAD, are lower for the group-level models than for the pooled models. Collectively, these results suggest superior fit and forecasting in time-series panel data for a model that incorporates the potential for coexisting business scenarios.

5.5. Nature of dynamic response

To examine whether our group-level approach produces substantive differences in dynamic response patterns, we computed impulse response functions (IRFs) for the effect of changes in detailing and

Table 5
Estimation sample model fit and model selection.

	SC values for group-level PVARs	SC values for pooled models	
		Levels	Differences
Group 1 Evolving business practice	18.378	25.772	18.579
Group 2 Hysteresis	18.404	28.304	18.487
Group 3 Escalation	18.836	25.387	18.922
Group 4 Business as usual	17.956	20.440	17.969
All groups	18.395	21.779	18.476

Legend: Schwarz criterion (SC) values for all models are based on the estimation sample of physicians. SC is defined as $-2(l/T) + k\log(T)/T$, where l is the log likelihood, k is the number of parameters, and T is the number of observations.

[7] The detailed results regarding lag length determinations for the PVAR models are omitted for brevity but are available from the authors upon request. Parameter estimates of the various PVAR models are also omitted because we subsequently report impulse response functions and elasticities.

[8] Although random effects might be appropriate in some contexts, we use a fixed-effects specification because marketing activity is correlated with prescribing at the physician level (Mizik & Jacobson, 2004; Manchanda et al., 2004) and random effects do not accommodate such correlation. Furthermore, Horváth and Wieringa (2008) report that a fixed-effects formulation outperforms a random-effects formulation in a PVAR application.

Table 6
Model comparisons for cross-sectional holdout sample: out-of-sample forecast errors for new prescriptions.

	Group-level models		Pooled model: Levels		Pooled model: Differences	
	RMSE	MAD	RMSE	MAD	RMSE	MAD
Group 1 Evolving business practice	3.398	2.344	3.418	2.400	3.456	2.417
Group 2 Hysteresis	3.769	2.576	3.792	2.593	3.827	2.606
Group 3 Escalation	3.202	2.238	3.217	2.251	3.208	2.238
Group 4 Business as usual	3.083	2.227	3.099	2.250	3.086	2.235
All groups	3.240	2.279	3.257	2.301	3.257	2.288

Legend: RMSE is the root mean square forecast error, and MAD is the mean absolute forecast error.

sampling on prescriptions. We follow previous research (Dekimpe & Hanssens, 1999; Nijs, Dekimpe, Steenkamp, & Hanssens, 2001) to derive generalized impulses to compute IRFs. These generalized impulses do not impose any particular ordering on the effect of the endogenous variables. To compute the standard errors of the IRF estimates, we use a bootstrap procedure repeated 250 times (see Srinivasan, Pauwels, Hanssens, & Dekimpe, 2004 for a discussion). We report the impulse response functions for new prescriptions with respect to a one-standard-deviation change in detailing in Fig. 3 and for new prescriptions with respect to a change of one standard deviation in sampling in Fig. 4. The solid lines give the IRFs, and the dashed lines give their 95% confidence intervals.

The patterns of dynamic response represented by the IRFs are broadly similar for detailing and sampling and are consistent with the nature of the business scenario classifications pertaining to each group (Dekimpe & Hanssens, 1999). For Group 1, the coevolution scenario, the responses over time are positive, significant, and permanent for both detailing and sampling. Group 4, the "business as usual" scenario, shows significant short-run response in both cases but returns to zero after about six periods (we note that only the immediate effect is significant in this group). In Group 3, escalation, a non-significant response to both detailing and sampling is well contained within the error bands except in the case of the immediate effect of sampling.

For Group 2, hysteresis, we also find that response is not significantly different from zero. For detailing, the IRF values are positive for most periods, but the error bands always include zero. This suggests that the sales calls made to physicians in this group, on average, are not producing sufficient short-run response to capitalize on the potential for a permanent effect from evolution in the prescriptions series. In other words, we do not find a marketing-based hysteresis in our data (i.e., the evolving nature of the sales series in this group does not seem to be caused by detailing and sample marketing effort). Indeed, evolution in sales can be due to factors other than marketing stimuli (e.g., physicians learning through usage). Our finding of the absence of marketing-induced hysteresis is consistent with prior research in that hysteresis is a very rarely observed phenomenon.[9]

Our finding that detailing and sampling have a limited impact on prescriptions (albeit significant in some of the smaller groups) is also consistent with prior research documenting modest effectiveness of detailing and sampling. For example, Kremer et al. (2008) surveyed the literature and noted that most recent studies report very modest effectiveness of detailing and sampling (some report no effect, and some studies report negative effects). Some earlier studies that did not control for individual-specific heterogeneity in levels of prescribing (e.g., Wittink, 2002) reported much higher effectiveness for detailing, although such results might have been driven by omitted variable bias (these studies did not account for individual-level heterogeneity via a fixed-effects formulation).

5.6. Comparison with pooled models

The IRFs for the pooled PVAR models show quite different patterns of response (bottom two panels of each figure). For details and samples, the pooled model in levels shows a significant and positive short-run effect that reverts to zero after about eight periods. Most importantly, the pooled model in differences estimates the effects of detailing and sampling to be not only positive and significant but also permanent. This is in stark contrast to the dynamic response implied by the group-level models (we find a significantly smaller group of physicians for whom detailing and sampling have permanent effects). The reason for these differences is that we do not impose the same dynamic patterns on all physicians. The group-level models allow impulse response functions to capture different dynamics at the segment level, revealing heterogeneity in the response dynamics within the same market.

5.7. Elasticities

To better understand the magnitude of the effect sizes for each group and for the sample as a whole, we compute detailing and sampling elasticities at different time points (1, 6, and 12 months). These elasticities are presented in Table 8. We begin by noting that the magnitude of the elasticities differs markedly across the four groups. Consistent with the IRF patterns, both detailing and sampling elasticities are largest for Group 1 and smallest for Group 3, while Groups 2 and 4 fall in between. The detailing elasticities for Group 3 are negatively signed but are very close to zero. These group-level results for the elasticities provide further support for the segmentation and targeting potential of business scenarios.

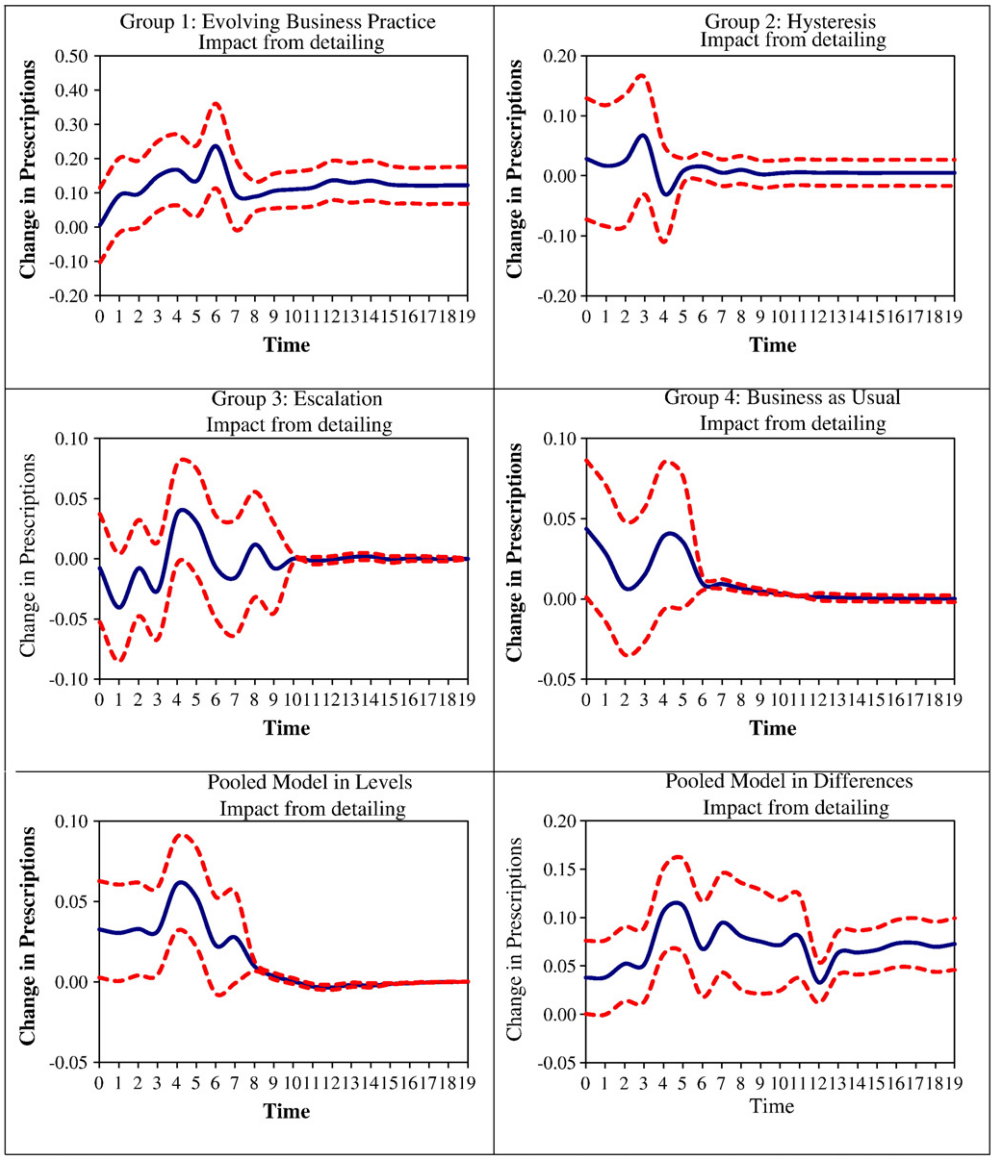
Table 8 also reports the overall elasticity obtained from the group-level PVAR models (computed as a weighted average of group elasticities) as well as the two pooled PVAR models. The two pooled models produce very different elasticity values; those from the pooled model in differences are substantially greater than the pooled model in levels. Indeed, at 12 months, the differenced model gives elasticities for the entire sample about three times larger than those from the group-level model. Because the overall properties of the panel data

Table 7
Model comparisons for longitudinal holdout sample (two-month holdout): out-of-sample forecast errors for new prescriptions.

	Group-level models		Pooled model: Levels		Pooled model: Differences	
	RMSE	MAD	RMSE	MAD	RMSE	MAD
Group 1 Evolving business practice	3.836	2.777	4.201	2.902	3.859	2.891
Group 2 Hysteresis	3.905	2.710	4.143	2.897	3.921	2.821
Group 3 Escalation	3.162	2.176	3.183	2.197	3.166	2.256
Group 4 Business as usual	3.175	2.199	3.194	2.201	3.178	2.259
All groups	3.333	2.308	3.425	2.349	3.339	2.387

Legend: RMSE is the root mean square forecast error, and MAD is the mean absolute forecast error. We present here the two-month holdout results for all physicians (estimation sample and cross-sectional holdout sample combined). We also examine the results using one-month holdout and using the estimation and the holdout sample physicians separately. The results are fully consistent with those reported here. Full details are available upon request.

[9] We thank Dominique Hanssens for very helpful discussions on the nature of the hysteresis scenario and the frequency of its occurrence. We note in the directions for future research that further within-group heterogeneity in response may exist within dynamic scenarios, but this issue is beyond the scope of the present study.



Note: Solid lines correspond to the responses to a generalized impulse of one standard deviation of detailing. Dashed lines represent the 95% confidence intervals.

Fig. 3. New prescription impulse response to change in detailing by group.

suggested possible evolution in the prescribing series, the choice of a pooled model in differences might be justified. However, the elasticity results we present show that a pooled model could suggest very different implications for market response and resource allocation.

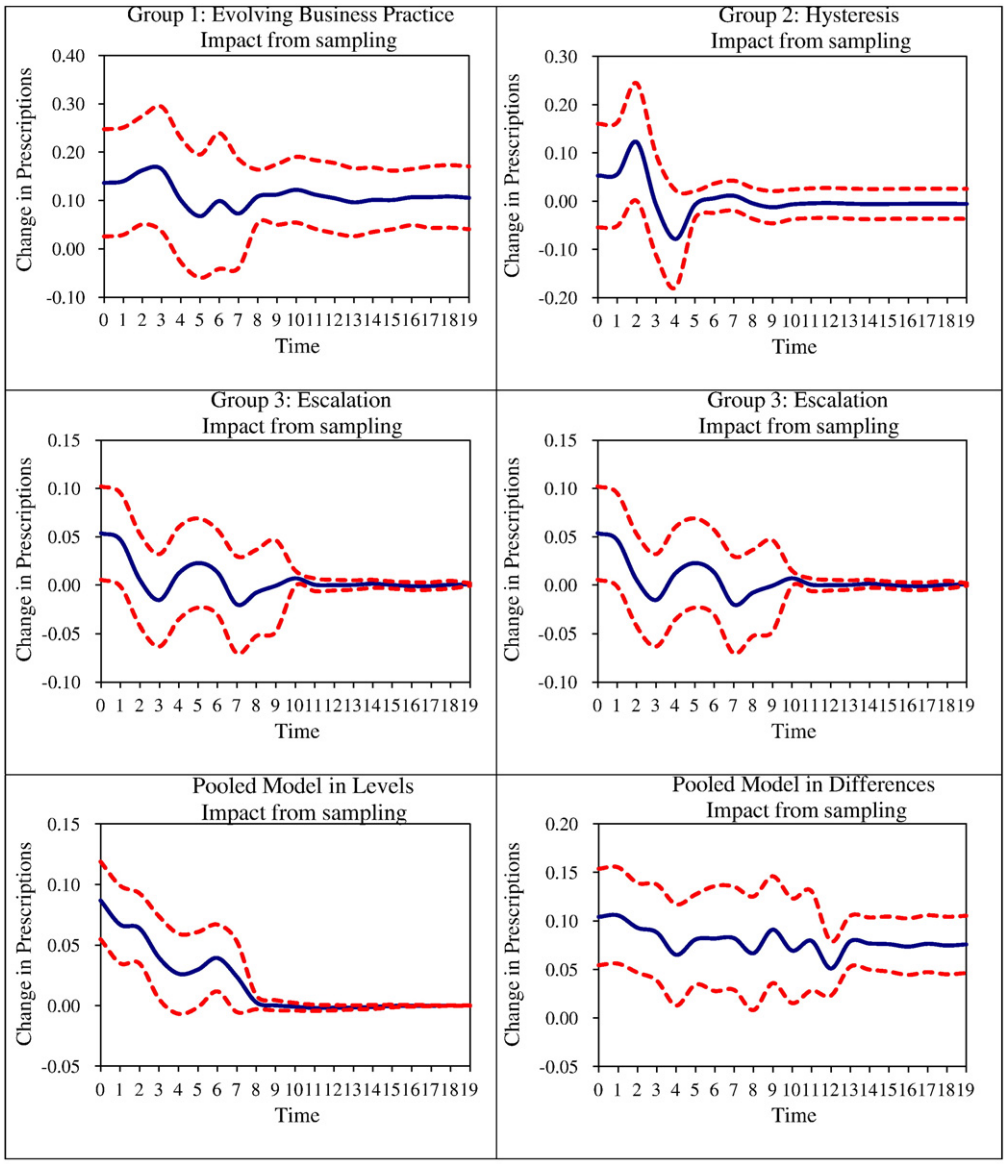
6. Sensitivity analyses

We undertook several sensitivity analyses to ensure the robustness of our results. In this section, we briefly summarize these analyses and our findings. We note that all additional analyses are fully supportive of our proposed approach and suggest the superiority of dynamic business scenario-based segmentation in identifying groupings with meaningful response differences over common alternatives used in the industry. Full results of these additional analyses are available from the authors upon request.

6.1. Alternative classification heuristics

We examined alternative segmentation heuristics and compared their performance to that of the proposed dynamic segmentation. We tested two different a priori groupings, one based on deciles (typically used by the industry) and one based on a combination of deciles, gender and years of experience (i.e., accounting for all physician-specific information available). For the first segmentation benchmark, we performed a median split of physicians based on their deciles (ranging from 4 to 10 and reflecting previous physician prescribing levels; the median decile was 6). We segmented physicians into a group of high deciles (deciles greater than 6) and one of low deciles (deciles less than or equal to 6). For the second segmentation benchmark, we used the k-means clustering algorithm and the Gower distance to cluster physicians using the three physician descriptors.[10] The three-cluster solution was deemed the best based on the silhouette criterion (further details available upon request). This solution segmented doctors into one group with all the female

[10] The Gower metric easily combines discrete and continuous variables in one single measurement, which is important in our specific application. In our case, deciles range from four to ten and reflect previous prescription levels, gender is a dummy variable that takes the value 1 for males and 0 for females, and years of experience ranges from 1 to 55 years for our physician sample.



Note: Solid lines correspond to the responses to a generalized impulse of one standard deviation of sampling.Dashed lines represent the 95% confidence intervals.

Fig. 4. New prescription impulse response to change in sampling by group.

physicians (who are less experienced and are typically in lower deciles) and two groups of male physicians, one with higher levels of past prescribing and the other with lower levels.

For both segmentation alternatives, we find no systematic relation between these alternative segments and the proposed dynamic-based groupings. The four dynamic groups are uniformly distributed across the segments, further strengthening the results in Table 3. In addition, we used panel unit-root tests to test for unit roots in each segment and concluded that no unit roots were present. We then estimated PVAR models in levels for each of the segments, and although unit root tests indicated no unit roots, we estimated a first-difference specification for completeness. As before, we relied on SC to determine the number of lags, allowed for time-specific effects, and included individual-level fixed effects in all models estimated in levels (detailed results available upon request). Finally, we estimated the elasticities for the new segments and computed the in- and out-of-sample performance of these two alternative segmentation benchmarks.

Our cross-sectional and longitudinal holdout tests clearly favor the business scenario-based segmentation (see Table 9 for detailed results). Similar results were found for the in-sample performance comparison. More importantly, we found no meaningful differences in response across segments, contrary to the results of our dynamic-segmentation approach. Thus, in our context, these alternative approaches (and the variables underlying the segmentation) were unable to reveal meaningful diagnostic information to managers, although significant differences were indeed present in the market. These results further highlight the substantive benefits of examining and addressing dynamic response using the proposed segmentation based on dynamic business scenarios.

6.2. Alternative estimation: Bayesian approach

In large samples, Bayesian methods are computationally intensive, and, as such, can be less practical. However, it might be argued that Bayesian methods are feasible and beneficial in moderate-size samples because they can better depict individual-level heterogeneity in response, even in a time-series context. To assess these arguments and to compare the performance of our two-step segmentation procedure against Bayesian estimation, we undertook additional tests.

Table 8
Elasticities for group and pooled models.

	Detailing elasticity			Sampling elasticity		
	1 month	6 months	12 months	1 month	6 months	12 months
Group 1 Evolving business practice	.014	.325	.727	.028	.152	.276
Group 2 Hysteresis	.032	.095	.163	.016	.033	.043
Group 3 Escalation	−.004	−.008	−.019	.012	.024	.025
Group 4 Business as usual	.029	.111	.135	.019	.041	.056
All groups	.014	.088	.147	.017	.047	.069
Pooled in levels	.020	.145	.182	.021	.061	.089
Pooled in differences	.021	.221	.481	.022	.104	.211

Legend: The "All groups" elasticities are computed as a weighted average of the group elasticities. We first compute the generalized IRFs of prescriptions due to a shock of one standard deviation of detailing or sampling (these will depend on the estimated PVAR parameters). We then compute the percentage change variation in prescriptions with respect to the mean prescriptions and the percentage change in detailing or sampling with respect to the detailing and sampling means. We take their ratios to obtain the elasticity. We repeat this for different time horizons. If we wanted to take into account that after a marketing shock the level of the marketing variable might also change over time (alternative discussed in the paper), we need to obtain the generalized IRF of the marketing mix variable with respect to a shock of that same marketing mix variable. We then recompute the percentage change of the marketing mix, considering the means of the marketing mix variables and the subsequent changes at each time period.

We re-estimated the PVAR models using random-effects Bayesian models in which all model parameters were allowed to be physician-specific. Because of the spurious regression problem, we re-estimated the models in differences with all variables—prescriptions, detailing and samples—entering the models in differences. We used conjugate and uninformative priors and estimated the models using standard Markov chain Monte Carlo (MCMC) methods and convergence diagnostics.

Our in- and out-of-sample fit results suggest that a non-segmented random-effects approach (i.e., a single model specified across all panelists allowing for intercept and response heterogeneity) is not appropriate if different dynamic scenarios coexist in the data. Our proposed model out-performed the Bayesian alternative, especially in longitudinal holdout (detailed results available from the authors upon request). One possible explanation for this result is that traditional Bayesian estimation shrinks individual estimates toward a prior and a common mean (or means) without any a priori screening for dynamic patterns (the first step in unit-root testing). This means that the Bayesian procedure may mask or blend dynamic patterns. Thus, for many panelists, the Bayesian estimates yield a dynamic profile that is similar to the largest group, which, in our case is Group 3, the escalation scenario. Given the existing differences in the market, the shrinkage leads to significantly worse out-of-sample predictive performance.

Table 9
Forecasting performance of dynamic business scenario segmentation, decile-based segmentation, and physician characteristics-based segmentation.

	Cross-sectional holdout		Longitudinal holdout	
	RMSE	MAD	RMSE	MAD
Dynamic scenarios segmentation (specification selected based on unit root tests)	3.240	2.279	3.333	2.308
Decile segmentation (in differences)	3.288	2.300	3.375	2.406
Decile segmentation (in levels)	4.103	2.984	3.454	2.368
Characteristics-based segmentation (in differences)	3.283	2.298	3.364	2.404
Characteristics-based segmentation (in levels)	4.083	2.969	3.418	2.346

Legend: RMSE is the root mean square forecast error, and MAD is the mean absolute forecast error. They are both computed for the cross-sectional and longitudinal holdout samples for the proposed dynamic business scenario segmentation, decile-based groupings, and physician characteristics-based groupings. Decile-based segmentation uses median splits to allocate physicians into high and low decile groupings. We used k-means clustering and the Gower distance measure to segment physicians based on their gender, experience, and decile. Longitudinal holdout results are based on all physicians and two-month holdout (results are also available on request for the one-month holdout and estimation and the cross-sectional physician holdout samples). The RMSE and MAD values in the table are aggregated across cluster and are reflective of the total sample.

6.3. Alternative computation of elasticities

The elasticities presented in Table 8 (for both the segmented and pooled models) were computed with respect to the initial exogenous shocks of detailing and sampling following the standard approach. That is, the elasticity numerator takes into account the impact on prescriptions from subsequent changes of the marketing variables (i.e., it is based on the generalized IRF), but the elasticity denominator does not incorporate the accumulated marketing changes. These cumulative changes, however, can contribute significantly to the total cost of the original marketing action, particularly when marketing is evolving.[11] To incorporate the full marketing cost considerations in the elasticity computation, we recomputed the 6-month and 12-month detailing elasticities, including the effect of a detailing shock on future detailing.

Overall, these new elasticities tend to be lower than the ones reported in Table 8, but they follow similar patterns, retaining the significant differences across groups. The recomputed elasticities also confirm that pooled models produce different implications for market response. For example, the recomputed detailing elasticities for the pooled model in differences are .082 and .110 for the 6-month and the 12-month windows, respectively. The corresponding recomputed values obtained for the segmented approach are .046 and .055, which are significantly lower.

6.4. Alternative formulation for time-specific effects

We chose not to use time trends in our models because trends impose the restriction of constant change from one period to the next. Instead, we opted for a more flexible formulation of time-specific effects. The levels of prescribing and marketing effort differ from period to period and are contemporaneously correlated simply due to the different number of work days in a given month. Our sensitivity tests show that failure to explicitly model these time-specific differences results in spurious relationships and biased estimates. We

[11] We would like to thank Professor Els Gijsbrechts for her comments on the alternative elasticity computations. We do not report a similar analysis for sampling because it is typically treated as having zero marginal cost in the industry.

verified the superior performance of our specification by re-estimating all group-level models, replacing time-specific effects with linear and quadratic trends and examining the cross-sectional and longitudinal out-of-sample predictive performance. The results showed that our specification of time-specific effects produced the lowest RMSE and MAD in all cases (e.g., for the cross-sectional holdout, including a linear trend produced a RMSE of 3.470 and a MAD of 2.967; for the longitudinal holdout, the RMSE was 3.374, and the MAD was 2.396; the inclusion of the quadratic trend produced even worse results).

6.5. Alternative unit-root tests

Finally, to assess the robustness of our findings to a particular unit-root test, we re-ran our analyses based on segments formed using the KPSS test. We obtained similar results. Specifically, we found that (1) the proportion of physicians assigned to each business scenario group changed little when using KPSS instead of ADF, (2) the KPSS-based segments also outperformed the pooled models in terms of fit, and (3) the segment elasticities between the ADF and KPSS-based groupings were very similar.

7. Discussion

In cases where marketing outcomes and marketing effort are potentially evolving, researchers have used persistence modeling techniques to identify and study the nature of the business scenario that characterizes time-series data (Dekimpe & Hanssens, 1995, 1999). These scenarios (business as usual, escalation, hysteresis, and evolving business practice) can have radically different implications for management and for the efficient allocation of scarce marketing resources.

We propose and illustrate an approach to modeling potentially coexisting business scenarios within the same product market. In doing so, we combine persistence modeling techniques with the possibility that customers can be segmented using the individual dynamic properties of prescribing and marketing. First, we conduct unit-root tests on the outcome and marketing activity variables at the panelist level and classify these series, for each panelist, as either evolving or stationary. Using this classification, we group panelists into the four business scenario groups and specify and estimate appropriate PVAR models for each group. We use IRFs to study the dynamic properties of the data and compute elasticities.

The proposed approach assesses the response to marketing effort and simultaneously addresses the dynamic properties of marketing effort, performance, and dependencies (feedback) among all series. It allows for a comprehensive assessment of the returns on marketing. For example, in the escalation scenario (evolving marketing effort and stationary outcomes), any (short-run) benefits of marketing will be eventually overshadowed by the escalation of marketing and ever-increasing spending.

We illustrate the approach using physician panel data provided by a pharmaceutical company. The segmentation obtained from the unit-root tests shows that each of the four business scenarios is populated by a sizable proportion of doctors. We also find that the PVAR models estimated at the group level provide a better in-sample and holdout fit to the data than the alternative models. These findings document the ability of a multiple scenario modeling approach to better represent the data than conventional benchmark models.

Most importantly, in addition to producing superior fit, the IRFs and dynamic response elasticities derived from the multiple scenario approach highlight significant differences across the groups. These differences suggest potential benefits of using the proposed approach in practice. Using the approach, firms can draw important new implications for segmentation, targeting, and marketing resource allocation.

7.1. The role of additional factors

The approach we propose can easily be extended to include additional factors that might influence physician prescribing, including other promotional tools or competitive marketing efforts. Unfortunately, we did not have information on other marketing efforts or competitive detailing and sampling activity in our data set. The competitive information is unavailable to the pharmaceutical firm providing the data. Several recent academic studies have also lacked competitive information and have presented a variety of compelling arguments for proceeding without it (e.g., Manchanda & Chintagunta, 2004; Manchanda et al., 2004; Mizik & Jacobson, 2004). Prior research has shown that the level of a physician's prescribing is the major determinant of the frequency of detailing visits. Because we control for physician-specific effects, the major common source of bivariate correlation between own and competitive marketing effort is removed. Given the body of empirical evidence on this issue, the results we report are not likely to be significantly affected by the lack of competitive marketing data.

7.2. Managerial use of the dynamic business scenarios segmentation

Our sensitivity analyses show that business scenario-based groups do not align with traditional targeting used in the industry (physician demographics and decile rankings) and that common segmentation heuristics do not create meaningful groupings with differentiated response or differentiated relevant behavioral dynamics. Many businesses track customer relationship data and transaction data, but they often do not have access to reliable demographic descriptors of the customers. Our findings suggest that meaningful and actionable segments can be constructed using the dynamic business scenario segmentation approach based on activity-only data and can be used for targeting and marketing activity allocation.

The approach is easy to implement and is highly scalable to large databases and continuous testing. Because tests can be run and models re-estimated quickly using standard methods, managers can easily obtain updated groupings and the associated IRFs. The results can be updated every time new data become available or when major events occur in the category, such as new drug entry or loss of patent protection. Gonzalez, Sismeiro, and Dutta (2008) demonstrate that events such as the loss of patent protection and new drug entry can lead to significant changes in the marketing policy of firms.

In the case of the pharmaceutical industry, new information on prescribing and marketing effort is updated every month. Testing and estimation could be performed monthly, and physicians' group assignments could be monitored over time. Indeed, customer membership in a group is unlikely to be stable over a long time period. We would expect customers to migrate from positive evolving consumption, to stationary, and then to negative evolving as the product moves through the adoption, maturity, and decline stages of its life cycle, respectively. That is, in the initial stages of the product life cycle, we would expect consumers to migrate from Groups 1 and 2 into Groups 3 and 4 and then back (assuming that product use is not abruptly discontinued).

Dynamic business scenario segmentation may offer greater benefits if it is implemented as a continuous process rather than as a one-time exercise. Continuously reclassifying customers and tracking changes in classification on a routine basis can enable managers to draw specific implications for when and to whom to accelerate, decelerate, or even stop marketing activity. For example, in our empirical illustration, we find a large group of physicians populating the escalation scenario (Group 3) and generating zero (insignificantly negative) response to marketing effort. Decelerating effort to physicians in this

group can free up resources that could be deployed for physicians in the "evolving business practice" scenario (Group 1) or in the "business as usual" scenario (Group 4), groups that show positive response. Clearly, escalation of marketing spending to non-responsive physicians is not appropriate unless a new and more effective message can be developed. Thus, when the movement of a physician from Group 1 into Group 3 is noticed, it may be advisable to decelerate marketing effort.

We note that such considerations could be subject to the Lucas critique. Because we are using PVAR models, if the acceleration and deceleration of marketing activities are within the typical ranges for a particular group (i.e., within the relevant data range used for the VAR estimation) and do not alter the long-term properties of the series, the problems associated with the Lucas critique are not significant. Indeed, as discussed by van Heerde, Dekimpe, and Putsis (2005), the Lucas critique is relevant in the VAR framework when radical policy shifts are studied.

We further note that our segment-based estimation and cross-sectional holdout tests may provide guidance (although far from a complete analysis) regarding possible outcomes of marketing activity regime shifts. Because we find that the patterns and estimated relationships continue into the cross-sectional holdout period, we can use our estimates to predict physicians' responses to marketing regime shifts.

An important caveat is that our business scenario segmentation is partially driven by consumer response and partially driven by firm marketing efforts. Hence, changes in marketing effort allocation that shift panelists into a different scenario could change the long-term properties of the response series and therefore should be analyzed with care. For example, escalation of marketing efforts to physicians in Group 2 may have no effect on the dynamic properties of the prescription series and thus may move Group 2 physicians to Group 1, resulting in significantly improved response. However, if the regime shift in marketing activity spurs a regime shift in response series for some physicians, it would move them to Group 3 and would result in decreased responsiveness. Similarly, escalating marketing efforts to Group 4 may move these physicians to Group 3 and result in a waste of marketing resources. The hope, of course, would be that a sustained, persistent marketing effort could spur evolution in prescribing and move physicians to Group 1.

Because we did not formally test the process and the consequences of shifting marketing effort regimes (from stationary to evolving and vice versa), further research and careful experimentation (preferably through field experiments) are needed to better understand these processes. For example, future studies can examine physicians' switching patterns across scenarios in response to changing marketing activity regimes. In sum, we view research into regime switching as an important topic for further study.[12]

8. Conclusion

We propose a new approach for segmenting customers based on the coexistence of multiple business scenarios in a single product market. This approach promises advantages in identifying more responsive targets and aiding resource allocation. In our empirical application, using data from a pharmaceutical drug market, the proposed multiple business scenario approach produced superior in- and out-of-sample fit and highlighted large differences across physician business scenario groupings. These differences (and segments) did not align with traditional targeting variables used in the industry (e.g., physician demographics and decile rankings) and could not be captured by other common segmentation tools. Our results demonstrate that the proposed approach can be used by firms to better understand the underlying dynamic market structure and better target their marketing efforts.

There are several limitations to our study that highlight opportunities for future research. First, we rely on unit-root tests to determine the stationary or evolving nature of each panelist's outcome and marketing activity variables. To the extent that the power of these tests is weak, there is a misclassification risk. Development of better diagnostic tests would improve the reliability of our dynamic segmentation approach. Current research on unit-root tests using wavelets shows promise in increasing power, especially when analyzing near unit-root alternatives (e.g., Fan & Gençay, 2010). Although wavelets have been previously used in marketing to study time-series properties in the frequency domain (Lemmens, Croux, & Dekimpe, 2007), their use in unit-root testing is less developed and is a promising area for future research.

Further improvements to this segmentation approach might be developed by incorporating heterogeneity in individual-level response dynamics within each business scenario. Indeed, an optimal allocation of resources requires a careful study of how spending should be allocated across segments and also within a given segment (e.g., based on individual-level response). While our sensitivity tests suggest that initial unit-root testing is always essential for the proper modeling of dynamic data, Bayesian estimation of individual-level response within business scenarios might offer additional insights and refinements to targeting strategy. The choice between classical and Bayesian estimation in the second phase should be determined by context, sample size, and the need for scalability.

Another interesting direction for future research is to address is the applicability of the Dorfman–Steiner theorem for marketing resource allocation in non-stationary environments (our Groups 2, 3 and 4). One way of approaching this may be to consider an extension of the models of Doraszelski and Markovich (2007) and of Dubè, Hitsch, and Manchanda (2005) to incorporate the possibility of dynamic business scenarios. These authors model the demand and supply side of the market and simulate the outcome of a dynamic game using the Markov perfect equilibrium (MPE) concept, allowing for dynamic competitive behavior (in price and advertising), market entry and exit, and time-dependent solutions. Because these papers did not consider dynamic business scenarios, we feel that this presents a very interesting opportunity for future research.

Finally, future research into modeling customer migration from one business scenario to another over time (e.g., from evolving practice to business as usual) and factors explaining the differences in the propensity to switch can generate academically and managerially valuable insights.

[12] We would like to stress the conceptual distinction between our focus on regime shifts (stationary versus evolving nature of the series) and the research on modeling transition processes among hidden states (e.g., inactive, infrequent, frequent), as in Montoya, Netzer, and Jedidi (2010).

References

Baltagi, B. H., & Kao, C. (2000). Nonstationary panels, cointegration in panels and dynamic panels: a survey. In B. H. Baltagi (Ed.), *Advances in econometrics*, *15*. (pp. 7–51). Amsterdam: Elsevier.

Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science*, *15*(5), 215–227.

Bronnenberg, B. J., Mahajan, V., & Vanhonacker, W. R. (2000, February). The emergence of market structure in new repeat-purchase categories: The interplay of market share and retailer distribution. *Journal of Marketing Research*, *37*, 16–31.

Bucklin, R. E., Gupta, S., & Siddarth, S. (1998, May). Determining segmentation in sales response across consumer purchase behaviors. *Journal of Marketing Research*, *35*, 189–197.

Campbell, J. Y., & Perron, P. (1991). Pitfalls and opportunities: What macroeconomists should know about unit roots. *NBER Macroeconomics Annual*, *6*, 141–204.

Clements, M. P., & Hendry, D. F. (1999). *Forecasting non-stationary economic time series*. Cambridge, Mass.: MIT Press.

Dekimpe, M. G., & Hanssens, D. H. (1995). The persistence of marketing effects on sales. *Marketing Science*, *14*(1), 1–21.

Dekimpe, M. G., & Hanssens, D. H. (1999, November). Sustained spending and persistent response: A new look at long-term marketing profitability. *Journal of Marketing Research*, *36*, 397–412.

Dekimpe, M. G., Parker, P. M., & Sarvary, M. (2000). Global diffusion of technological innovations: A coupled-hazard approach. *Journal of Marketing Research*, *37*(1), 47–59.
Doraszelski, U., & Markovich, S. (2007). Advertising dynamics and competitive advantage. *The Rand Journal of Economics*, *38*(3), 557–592.
Dubè, J. -P., Hitsch, G. J., & Manchanda, P. (2005). An empirical model of advertising dynamics. *Quantitative Marketing and Economics*, *3*(2), 107–144.
Fan, Y., & Gençay, R. (2010). Unit root tests with wavelets. *Econometric Theory*, *26*(5), 1305–1331.
Gonzalez, J., Sismeiro, C., & Dutta, S. (2008). Can branded drugs benefit from generic entry? The role of detailing and price in switching to non-bioequivalent molecules. *International Journal of Research in Marketing*, *25*(4), 247–260.
Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, *2*(2), 111–120.
Harris, R. I. D. (1992). Testing for unit roots using the augmented Dickey–Fuller test: Some issues relating to the size, power and lag structure of the test. *Economics Letters*, *38*(4), 381–386.
Haug, A. A. (1996). Tests for cointegration: A Monte Carlo comparison. *Journal of Econometrics*, *71*(1–2), 89–115.
Hausman, J. A. (1978, November). Specification tests in econometrics. *Econometrica*, *46*, 1251–1272.
Hellström, J. (2001). Unit root testing in integer-valued AR(1) models. *Economic Letters*, *70*(1), 9–14.
Horváth, C., & Franses, P. H. (2003). *Deriving dynamic marketing effectiveness from econometric time series models.* Working paper. Rotterdam, Netherlands: Erasmus University.
Horváth, C., Leeflang, P. S. H., Wieringa, J. E., & Wittink, D. R. (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., & Wieringa, J. E. (2008). Pooling data for the analysis of dynamic marketing systems. *Statistica Neerlandica*, *62*(2), 208–229.
Janakiraman, R., Dutta, S., Sismeiro, C., & Stern, P. (2008). Physicians' persistence and its implications for their response to promotion of prescription drugs. *Management Science*, *54*(6), 1080–1093.
Janakiraman, R., Sismeiro, C., & Dutta, S. (2009). Perception spillovers across competing brands: A disaggregate model of how and when. *Journal of Marketing Research*, *46*, 467–481.
Kamakura, W. A., & Russell, G. J. (1989, November). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, *26*, 379–390.
Kamakura, W. A., & Wedel, M. (2000). *Market segmentation: Conceptual and methodological foundations* (2nd Edition). Norwell, MA: Kluwer Academic Publishers.
Kremer, S. T. M., Bijmolt, T. H. A., Leeflang, P. S. H., & Wieringa, J. E. (2008). Generalizations on the effectiveness of pharmaceutical promotional expenditures. *International Journal of Research in Marketing*, *25*(4), 234–246.
Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992, Oct–Dec). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*, 159–178.
Leeflang, P. S. H., & Wieringa, J. E. (2010). Modeling the effects of pharmaceutical marketing. *Marketing Letters*, *21*(2), 121–133.
Lemmens, A., Croux, C., & Dekimpe, M. G. (2007). Consumer confidence in Europe: United in diversity? *International Journal of Research in Marketing*, *24*(2), 113–127.
Lieberman, M. B. (1987). The learning curve, diffusion, and competitive strategy. *Strategic Management Journal*, *8*(5), 441–452.
Lim, J., Currim, I. S., & Andrews, R. L. (2005, December). Consumer heterogeneity in the longer-term effects of price promotions. *International Journal of Research in Marketing*, *22*, 441–458.
Maddala, G. S. (1992). *Introduction to econometrics* (2nd ed.). New York, N.Y.: Macmillan.
Manchanda, P., & Chintagunta, P. (2004). Responsiveness of physician prescription behavior to salesforce effort: An individual level analysis. *Marketing Letters*, *15*(2–3), 129–146.
Manchanda, P., Rossi, P., & Chintagunta, P. (2004, November). Response modeling with non-random marketing mix variables. *Journal of Marketing Research*, *41*, 467–478.
Manchanda, P., Wittink, D. R., Ching, A., Cleanthous, P., Ding, M., Dong, X. J. J., et al. (2005). Understanding firm, physician and consumer choice behavior in the pharmaceutical industry. *Marketing Letters*, *16*(3–4), 293–308.
Mela, C. F., Gupta, S., & Lehmann, D. R. (1997, May). The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research*, *34*, 248–261.
Mizik, N., & Jacobson, R. (2004). Are physicians 'easy marks'? Quantifying the effects of detailing and sampling on new prescriptions. *Management Science*, *50*(12), 1704–1715.
Montoya, R., Netzer, O., & Jedidi, K. (2010). Dynamic allocation of pharmaceutical detailing and sampling. *Marketing Science*, *29*(5), 909–924.
Narayanan, S., & Manchanda, P. (2005). *Heterogeneous learning and the targeting of marketing communication for new products.* Working paper. Palo Alto, Calif.: Stanford University.
Narayanan, S., Manchanda, P., & Chintagunta, P. (2005, August). Temporal differences in the role of marketing communication in new product categories. *Journal of Marketing Research*, *42*, 278–290.
Nelson, C. R., & Kang, H. (1984). Pitfalls in the use of time as an explanatory variable in regressions. *Journal of Business and Economic Statistics*, *2*(1), 73–82.
Nijs, V. R., Dekimpe, M. G., Steenkamp, J. -B. E. M., & Hanssens, D. M. (2001). The category–demand effects of price promotions. *Marketing Science*, *20*(1), 1–22.
Osinga, E. C., Leeflang, P. S. H., & Wieringa, J. E. (2010). Early marketing matters: A time-varying parameter approach to persistence modeling. *Journal of Marketing Research*, *47*(1), 173–185.
Papatla, P., & Krishnamurthi, L. (1996, February). Measuring the dynamic effects of promotions on brand choice. *Journal of Marketing Research*, *33*, 20–35.
Parsons, L. J. (1975, November). The product life cycle and time-varying advertising elasticities. *Journal of Marketing Research*, *12*, 476–480.
Pauwels, K. H., Currim, I., Dekimpe, M. G., Ghysels, E., Hanssens, D. M., Mizik, N., et al. (2005). Modeling marketing dynamics by time series econometrics. *Marketing Letters*, *15*(4), 167–184.
Pauwels, K. H., Hanssens, D. H., & Siddarth, S. (2002, November). The long-term effects of price promotions on category incidence, brand choice and purchase quantity. *Journal of Marketing Research*, *39*, 421–439.
Plosser, C. I., & Schwert, G. W. (1977, September). Estimation of non-invertible moving average processes: The case of overdifferencing. *Journal of Econometrics*, *6*, 199–224.
Rogers, E. (2003). *Diffusion of innovations* (5th ed.). New York, N.Y.: Free Press.
Srinivasan, S., Pauwels, K. H., Hanssens, D. M., & Dekimpe, M. (2004). Do promotions increase profitability and for whom? *Management Science*, *50*(5), 617–629.
Stock, J. H., & Watson, M. W. (1988, December). Testing for common trends. *Journal of the American Statistical Association*, *83*, 1097–1107.
Theil, H. (1961). *Economic forecasts and policy* (2nd ed.). Amsterdam: North Holland Publishing Company.
van Heerde, H. J., Dekimpe, M. G., & Putsis, W. P., Jr. (2005). Marketing models and the Lucas critique. *Journal of Marketing Research*, *42*(1), 15–21.
Wittink, D. R. (2002). *Analysis of ROI for pharmaceutical promotion (ARPP).* Presentation to the association of medical publications, September 18, 2002 available from http://www.vioworks.com/clients/amp (last accessed, March 31, 2011).