MANAGEMENT SCIENCE
Vol 28, No. 5, May 1982
Printed in U.S.A

DYNAMIC ANALYSIS OF CONSUMER RESPONSE TO MARKETING STRATEGIES*

JOHN R. HAUSER† AND KENNETH J. WISNIEWSKI‡

This paper develops a methodology for modeling consumer response that integrates previous research in stochastic brand selection, diffusion of innovation, test market analysis, and new product design. The methodology makes it practical to extend brand selection models to include diffusion phenomena such as awareness, trial, and information flow. Purchase timing and brand selection are interdependent and both phenomena depend jointly on managerial controls such as advertising, coupons, price-off promotion, product positioning, and consumer characteristics.

Within this general structure, we provide practical estimation procedures (a least squares approximation to the maximum likelihood estimates) to determine the parameters which link managerial controls to consumer response. Closed form solutions are derived for cumulative awareness, cumulative trial, penetration, expected sales, and purchases due to promotion—all as a function of time. We also provide simplified expressions for equilibrium ($t \to \infty$) market share. Tradeoffs among complexity of the diffusion process, number of managerial variables, nonstationarity, complexity of purchase timing, consumer segmentation, and sample size are made explicit so that the marketing scientist can customize his analyses to the managerial problems that he faces.

The effects of sample size, data interval frequency, and collinearity in the explanatory variables are investigated with simulations based on a five-state consumer response process which depends on 8–10 marketing variables.

The paper closes with a brief description of the application and predictive test of a consumer response model based on the methodology.

(MARKETING; CONSUMER BEHAVIOR; MARKOV ANALYSIS)

1. Perspective

An important goal in modeling consumer response is to understand and forecast the impact of marketing strategies. For example, if a marketing manager mails out free samples during a new product launch, he wants to know how many consumers will try the sample and how this will affect long run market share. But, particularly during a test market or a new product launch, the dynamics of consumer response are important. For example, the marketing manager also wants to know how quickly consumers will try the sample and what impact this has on the speed with which the new product achieves its long run share. Furthermore, today's manager wants to improve strategy as he gains more experience with his market and adapt strategy as market conditions change. To address these issues a consumer response model should be dynamic and dependent on managerial controls (marketing strategy) and the market environment.

For example, suppose we are in a market where the only brands of analgesics are Bayer, Excedrin, and Tylenol. Suppose we recruit a panel of consumers and periodically observe what brand was last purchased. The observations are simple, but the stochastic process generating the observations is quite complex. The interpurchase time is a random variable which may depend on the brand last purchased (Excedrin may be more effective and hence used up at a slower rate) and on marketing strategy (a

*Accepted by Gary L. Lilien, former Departmental Editor; received November 1980. This paper has been with the authors 2 months for 2 revisions.

†Massachusetts Institute of Technology.
‡University of Chicago.

0025-1909/82/2805/0455$01.25

consumer decides to buy earlier because of a 'limited time offer'). The switching among brands may depend on marketing strategy (Bayer users may be more likely to switch to Excedrin which contains some aspirin than to Tylenol which contains none), and on diffusion phenomena (a consumer may be more likely to adopt Tylenol if other consumers are using it). Finally, there are often multiple purchases. If on July 1 the consumer was using Bayer and on July 15 he says his last purchase was Tylenol, he may have purchased Tylenol in the two week period, but alternatively he may have purchased Bayer, Bayer, Excedrin, and then Tylenol in the two week period. Thus a "flow" from Bayer to Tylenol may be a series of "flows" which depend jointly on the strategies of all the brands. This complex consumer response process is the process we seek to model.

We seek a mathematical structure which allows switching and interpurchase times to depend upon marketing strategies and other explanatory variables such as word of mouth. We seek a practical estimation procedure that allows us to estimate the parameters of this continuous time process (which relate marketing strategies to consumer response) by observing the process at discrete points in time (e.g., weekly). Finally, once the process parameters are estimated, we seek closed form expressions which dynamically forecast sales, market share, and penetration for new marketing strategies.

We seek a flexible evolutionary structure. For example, suppose the manager believes that consumer response depends upon whether or not the last purchase was made "on deal". Such a consumer response can be described by expanding the three "behavioral states" (Bayer, Excedrin, Tylenol) to six "behavioral states" (Bayer-regular price, Bayer-deal, Excedrin-regular price, Excedrin-deal, Tylenol-regular price, Tylenol-deal). Similar state expansions can handle diffusion phenomena such as awareness, trial, and repeat. Thus we seek a general model that can handle any number of "behavioral states" and explanatory variables—subject of course to data limitations.

Fortunately there has been a wealth of research in marketing, sociology and operations research that has investigated consumer response. This work is diverse, but when examined within a common framework it provides the basis for a practical, theory-based, evolutionary model of consumer response. This paper provides an analysis methodology that integrates many of the important models in these diverse literatures. Although the methodology is designed to handle complex phenomena, we provide estimation procedures and managerially relevant statistics that are implementable with standard regression and eigenstructure computer packages. Tradeoffs among the complexity of the consumer process, the number of behavioral states, the number of control variables, non-stationarity, and consumer segmentation are based on data requirements (sample size), not analytic or computational limitations.

The basic model is a dynamic, semi-Markov model of consumer response. Transition probabilities and purchase timing jointly depend upon consumer characteristics and marketing strategy and change as the market environment or marketing strategy changes. Diffusion phenomena such as awareness, trial and brand switching are modeled explicitly and depend upon marketing strategy. We derive methods to estimate the parameters of the system from panel data (mature products) or questionnaire data (new products). Once the parameters are estimated, we provide closed form expressions for statistics of managerial interest such as the expected sales over time, the penetration of the market over time, cumulative awareness, and cumulative trial. Analytic arguments and simulation provide the modeler with the tradeoffs necessary to customize the analysis to a wide variety of managerial problems.

We begin with a review of the relevant literature.

2. Existing Literatures

Because of the breadth of the research on consumer response, the following brief review can only highlight some of the important issues. We refer the reader to the references for more detail.

Stochastic Models of Brand Selection

Since the early 1950s researchers have used probabilistic models as a modeling strategy to describe consumer response (Brown [18], Cunningham [23], [24], Maffei [76], Harary and Lipstein [40], Howard [52], Styan and Smith [104], Massy [79], Morrison [89], Kuehn and Day [68], Rao [96], and Bass, Pessemier, and Lehmann [13]). These models address two primary issues: brand selection and purchase timing. For example, Herniter [44] uses a homogeneous Markov process (all consumers have the same probabilities) to describe brand selection and a heterogeneous Erlang distribution (the parameters vary across the population) to model the time between purchases. He assumes brand selection is independent of purchase timing. Other researchers extended these models to include learning (Kuehn [67], Carman [19], Jones [58, 59], Aaker and Jones [1]), heterogeneity in the parameters of brand selection (Frank [36], Bass [10], Bass, Jeuland, and Wright [12], and Kalwani and Morrison [63]), non-stationarity (Howard [52], Montgomery [86], Ehrenberg [31], Massy and Morrison [80]), more complex models or purchase timing (Zufryden [111]), heterogeneity in models (Jones [58], Blattberg and Sen [17], Givon and Horsky [37]), and risk aversion (Hauser [41]). But all models assume independence of brand selection and purchase timing (or do not address purchase timing). Many models limit their analytic solutions to a two-state process consisting of "purchase our brand" and "purchase all other brands" although most do suggest extensions to the multibrand case. See Blattberg and Sen [17] for a discussion of the dangers of two-state processes. Recently, Jeuland, Bass, and Wright [57] have extended their multinomial multibrand brand selection model to include an independent model of heterogeneous Erlang purchase times. Other notable exceptions to two-state models are Bass [10], Herniter [44, 45], and Kalwani and Morrison [63] who are primarily concerned with using observations on interbrand switching to identify hierarchies within the market. Except for identifying hierarchies, all of the above models *describe* consumer response rather than *prescribe* managerial action. More recently, researchers have used stochastic models to diagnose market research data such as purchase intentions (Morrison [91]) and taste tests (Morrison and Brockway [92]).

Early attempts to model the effect of marketing mix variables were moderately successful, although limited to one or a few marketing mix variables. For example, Herniter and Magee [48] suggest that different transition matrices should be used for different levels of marketing expenditure. Albright and Winston [2] extended their model and provided theorems to indicate the directionality of advertising and price strategies. Telser [105] modeled transition probabilities in a two-state homogeneous Markov process as a linear function of price and then modified this model to include advertising rather than price (Telser [106]). Horsky [50] added an optimization component to Telser's model while Lilien [72], [73] extended the concept to a two-state, heterogeneous, linear learning model. Haines [39] also used a linear learning model to represent consumer response. After estimating the parameters of the model, he let equilibrium market share and "approach-to-equilibrium" depend on a linear combination of marketing variables. (Advertising and availability were significant.) MacLachlin [75] modeled brand switching with a Markov transition matrix in which the probabilities were linear functions of marketing variables. Finally, Massy, Montgomery, and Morrison [81, p. 429] suggested that parameters of their stochastic evolutionary model be modeled as linear functions of the marketing mix variables. However, they were careful to point out that estimation is potentially very difficult.

In economics, McFadden [82], [83] has developed a multinomial model (probabilities independent of last purchase) that models purchase probabilities as a function of any number of marketing variables. McFadden's model, called the logit model, has been utilized by Jones and Zufryden [61] in a two-state, heterogeneous multinomial model and by Zufryden [110] in a two-state homogeneous Markov model. Both models include purchase timing but model it as independent of brand selection. Zufryden's model is a practical application of a homogeneous semi-Markov model proposed by Howard [52]. While theoretically quite general, Howard's model has not been widely applied because it is not dependent on marketing variables and because, as Montgomery and Urban [87, p. 78] point out, "there is a need for development of statistical methods to render this model more empirically viable . . . and the model may place excessive burdens on a data base when used in its most general form." Recently Lerman [71] has applied a four-state version of Howard's model to the choice of transportation mode and destination. What is interesting about his application is that he uses the logit model for brand selection and allows purchase timing to be dependent on whether or not the trip was the first trip of the day. However, because Lerman relied heavily on Laplace transforms he found it difficult to extend his model to more complex processes and did not model purchase timing as dependent on managerial controls.

In other applications Ezzati [34], Kao [65], and Shachtman and Hogue [99] have used homogeneous Markov models to successfully project consumer response to services.

Diffusion of Innovation

Sociologists have long recognized the diffusion of information through the population as important phenomena affecting consumer response. (Rogers and Shoemaker [98] review approximately 1500 articles and books on the communication of innovations.) These phenomena have formed the basis for some important models of consumer information acquisition (Lavidge and Steiner [70], Rogers [97]) and behavior (Midgley [84], [85]), and have led to the definition and identification of those consumers, called innovators, who have a major effect on the communication process. In marketing, diffusion phenomena have been particularly effective in modeling the growth, and possibly decline, of sales in new product categories (Bass [9], Nevers [94], Dodds [25], Mahajan and Peterson [78], Mahajan and Muller [77]). While these models are primarily descriptive, recent extensions have added the influence of marketing mix variables on the growth of a product category (Bass [11], Dodson and Muller [27], Horsky and Simon [51], Dolan and Jeuland [28]). Balachandran and Deshmukh [7] provide a theoretical five-state semi-Markov diffusion model for arbitrary communications processes. However, their model becomes intractable for practical problems, does not address heterogeneity, and is limited to one or a few marketing mix variables.

A major contribution of this literature is the recognition and explicit modeling of the dynamics of consumer response. To date all applications have dealt with product categories and have not modeled the dynamics of brand selection within a product class. Nonetheless, the empirical evidence and theory suggest that the *concept* of communication through advertising, trial incentives, and word of mouth is an important phenomena for marketing strategy affecting brand selection for new and mature products. A model that is sensitive to these decision variables should include consumer information acquisition states such as awareness, trial, and repeat purchase.

Test Market Models

One model that uses concepts from the diffusion of innovation to model consumer response is Urban's SPRINTER [108]. Urban models consumer response during new product test markets as a homogeneous, discrete time Markov process. He allows the parameters of the system to change over time (nonstationarity) and achieves sensitivity to marketing mix variables by having the transition probabilities be functionally dependent on managerial actions. Although he combines data from panels, questionnaires, and store audits, Urban's model requires a large number of nonstationary parameters which put a strain on data resources. Furthermore the flows are estimated independently and period by period and thus do not make full use of the stochastic properties of the Markov model. Other researchers (Assmus [5], Wachsler et al., [109]) have developed similar analyses based on Urban's model.

Blattberg and Golanty [15] present an alternative approach to test market modeling. They use recursive regression equations to operationalize an awareness, trial, and repeat process. While much simpler than Urban's model, their model is quite successful in projecting test market trends. In other models, Eskin [33] and Fourt and Woodlock [35] recognize the importance of consumer dynamics in models of awareness, trial, and repeat.

New Product Design Models

While stochastic models postulate a distribution of purchase probabilities across the population and test market models use homogeneous probabilities to incorporate consumer dynamics, new product design models have concentrated on estimating purchase probabilities based on product characteristics and managerial actions. See Shocker and Srinivasan [100] for a recent review of these models. For example, Hauser and Urban [42] provide a means to estimate heterogeneous multinomial probabilities based on product positioning strategy and advertising and distribution strategy. While these models have been successful in predicting equilibrium market share they have not addressed purchase timing or consumer dynamics.

Discussion and Integration

Although consumer response has been studied extensively, it is clear that the research is quite diverse. Each literature has approached an aspect of consumer response with different methods, interests and outputs. Stochastic modelers have explicitly recognized the probabilistic nature of consumer response. They have addressed the relationship between brand selection and purchase timing, the potentially Markovian nature of the process, and consumer heterogeneity. But all analyses deal with markets in equilibrium and assume independence of purchase timing and brand selection. But most managerial actions, by their very nature, are designed to upset the equilibrium and gain advantage for the firm that is acting. For example, if General Mills uses a 50%-off promotion, consumers stock up (thus affecting next purchase timing) and there is a diffusion process as consumers become aware of the deal and perhaps try the product for the first time. Most importantly, the stochastic models are either not dependent on managerial controls, limited by estimation difficulty to one or a few managerial controls, or limited to only a few behavioral states. In all cases, extensions of the models are theoretically possible, but limited by practical considerations in either estimation or analytic tractability.

Diffusion modelers have explicitly recognized consumer dynamics, but their models have not been modified for brand strategies (rather than product categories) and have not incorporated the complexities (purchase timing, Markovian nature, consumer heterogeneity) that have been addressed with stochastic models. Furthermore the modeling of managerial actions has been restricted by analytic necessity to one or a few managerial variables. On the other hand, test market modelers have incorporated diffusion phenomena for individual brands but have not included some of the phenomena identified by stochastic modelers and, in some cases, require more parameters to be estimated than are needed to control the process. Finally, new product design models do not address consumer dynamics or purchase timing, but provide a theory-based link from product positioning to (multinomial) purchase probabilities.

While each literature has successfully modeled a specific aspect of consumer response, we believe that each application can be improved and new issues addressed with a methodology that integrates these diverse, yet complementary, literatures. We provide a modeling approach that incorporates consumer dynamics, allows interdependence among purchase timing and brand selection, is theoretically and practically dependent on managerial actions, and is non-stationary in the sense that transition probabilities depend on the state of the system and managerial actions.

3. Mathematical Model

In this section we derive the mathematical model. The theory applies to both transient processes (test markets and new product launches) and equilibrium processes (mature products).

Semi-Markov Process

We model consumer response as a semi-Markov process. In particular, there is some set of consumer states, S_i, that consumers occupy and flow between over time. In a general semi-Markov process there are transition probabilities, q_{ij}, which determine the probability of a transition to state S_j given that the consumer is now in state S_i, but the time, t, at which the consumer makes a transition is a random variable with density function dependent upon state S_i.

For example, Figure 1 is one simplified model (including diffusion) for a new product entering a market dominated by two existing products. Note that a purchase is a flow (including self-loops) into a purchase state. (We use this simple model to illustrate the mathematics. In theory, the analysis can handle any consumer response process representable by behavioral states. Flows can include communication, forgetting, inventory effects, or brand purchase. We discuss tradeoffs among simplicity vs. complexity and sample size in §4.)

Although the evidence is mixed on whether the equilibrium consumer response process is first-order Markov, we model it as such because: (1) simpler models such as multinomial probabilities can be modeled as a zero-order Markov process, (2) higher-

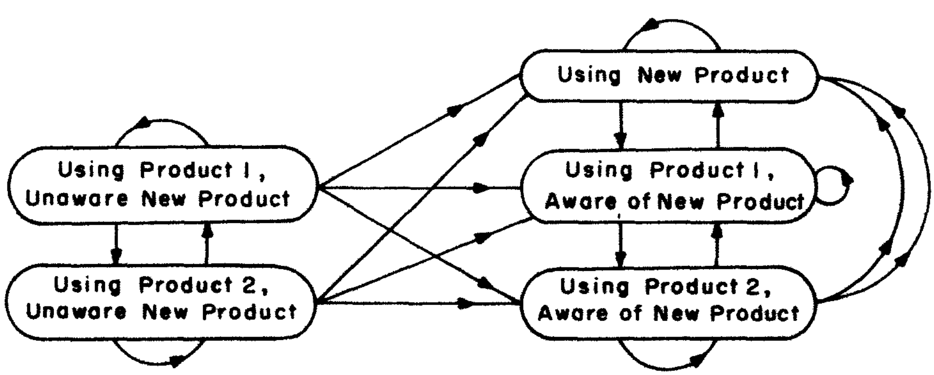

FIGURE 1. Example Semi-Markov Model for Simplified Market.

order processes can be represented as first-order processes by properly defining the consumer response states, and (3) even if the equilibrium process is zero-order we expect the dynamic process to be at least first order. (E.g., purchase may depend upon whether or not the consumer is aware of the product.)

In order to achieve tractibility and estimation capabilities for large problems, we model the probability of making a transition as either a Poisson random variable or as a compound Poisson random variable. For example, we might expect flow rates into awareness and trial to be Poisson[1] (the distribution of the time until the next transition being negative exponential) while interpurchase intervals such as flow rates from product 1 to product 2 might be compound Poisson (the distribution of time until the next purchase being Erlang and thus having its maximum value at some time other than $t = 0$). While this may seem restrictive, there is empirical and theoretical evidence that either a negative exponential or an Erlang distribution is a good model of purchase timing (Ehrenberg [30], [33], Massy, Montgomery and Morrison [81], Herniter [44], Jeuland, Bass, and Wright [57], Chatfield and Goodhardt [20], and Zufryden [111]). Furthermore, the Poisson and compound Poisson processes can approximate a wide range of phenomena.

Mathematically, we model a ρth order Erlang transition time as ρ sequential Poisson processes with the same flow rates, μ_i, where we are concerned only with the last transition. (The probability of a transition from state S_i in time Δt is $\mu_i \Delta t$.) Thus, to handle Erlang transition times we expand our state space to include what we will call "shadow states" such as in Figure 2. *If we assume ρ_i is known for each state*, the "shadow states" do not increase the number of unknown parameters since the flow rates are equal for "shadow states" and the extra transition probabilities are uniquely defined ($q_{i,i+1} = 1$, $q_{ij} = 0$ otherwise). With this construction the semi-Markov process of consumer response is transformed to a continuous time Markov process consisting of a set of real and "shadow" behavioral states with transition probabilities, q_{ij}, and state dependent Poisson flow rates, μ_i. The continuous time process is easier to handle mathematically but can be transformed back to the semi-Markov process for interpretation.

We provide analytic procedures for estimating the dependence of q_{ij} and μ_i on explanatory variables. The integers, ρ_i, can be determined by (1) a priori theory as illustrated in our empirical example, (2) statistical examination of interpurchase times (Lerman [71], Dodson [26]), or (3) iterative search. For example, one can estimate the process assuming $\rho_i = 1, 2, 3, \ldots, h$ and select the best fit if $h \times$ (number of states) is not too large. For small h this is reasonable computationally since each estimation for fixed ρ_i costs only a few dollars.

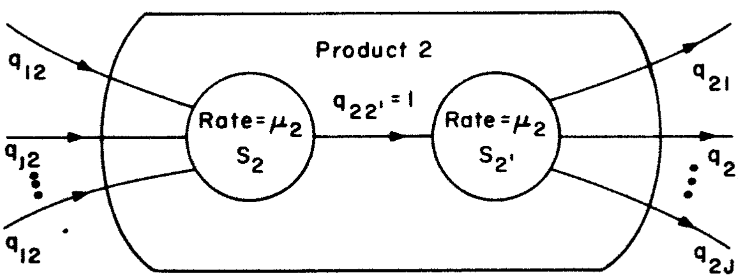

FIGURE 2. Shadow States to Model Erlang (Order 2) Interpurchase Times.

[1] For example, the well-known Lanchester advertising response model is a two-state semi-Markov process with Poisson flow rates proportional to advertising. See discussion in Little [74].

We consider collecting data in at least two ways (1) periodic questionnaires and (2) panel data. In both cases we observe the state the consumer was in at the beginning of the nth measurement, $S_i(T_{n-1})$, and the state he is in at the end of that measurement, $S_j(T_n)$. (Define $S_i(T_n) = 1$ if the consumer is in state i at time T_n and $S_i(T_n) = 0$ otherwise.) Thus the statistic we use to describe the process is the probability, $p_{ij}(t_n)$, that the consumer is in state S_j at time T_n given that he started in state S_i at time T_{n-1}. I.e.,

$$p_{ij}(t_n) = \text{Prob}\{S_j(T_n) = 1 | S_i(T_{n-1}) = 1\} \tag{1}$$

where $t_n = T_n - T_{n-1}$. Any data collection procedure that provides observations on $S_j(T_n)$ *and* $S_i(T_{n-1})$ can be used to implement the model.

To derive an expression for $p_{ij}(t_n)$ for each consumer we consider small time intervals, Δt_n (Cox and Miller [21], Drake [27]). Using the Poisson nature of the process, the only two ways for the consumer to be in state S_j at time $t_n + \Delta t_n$ are (1) to be in state S_j at time t_n and remain there (probability $= 1 - \sum_{k \neq j} \mu_j q_{jk} \Delta t$) and (2) to be in a state S_k at time t_n and flow to state S_j in Δt_n (probability $= \mu_k q_{kj} \Delta t_n$). Thus, algebraically we write these conditions as:

$$p_{ij}(t_n + \Delta t_n) = p_{ij}(t_n)\left[1 - \sum_{k \neq j} \mu_j q_{jk} \Delta t_n\right] + \sum_{k \neq j} p_{ik}(t_n) \mu_k q_{kj} \Delta t_n. \tag{2}$$

By redefining notation we can simplify (2). Define $a_{kj} = \mu_k q_{kj}$ for $k \neq j$ and $a_{jj} = -\sum_{k \neq j} \mu_j q_{jk}$ and if we let $\Delta t_n \to 0$ we can write (2) as:

$$\frac{d}{dt_n} p_{ij}(t_n) = \sum_k p_{ik}(t_n) a_{kj} \tag{3}$$

or in matrix form:

$$\frac{d}{dt_n}[P_n(t_n)] = P_n(t_n) A_n \tag{4}$$

where we have added a subscript, n, to the A_n matrix to indicate possible non-stationarity because of different managerial actions (which will be shown to affect A_n) in different time periods of observation. ((4) is known as the Chapman-Kolmogorov forward equation.) Note that the length of observation periods, t_n, can vary freely. This allows data collection strategies such as weekly sampling at the beginning of a new product launch and monthly sampling after the first few months.

The solution to (4) with initial conditions, $p_{iin}(t_n = 0) = 1$, $p_{ijn}(t_n = 0) = 0$ for $j \neq i$, is given by (Cox and Miller [21], Athans and Falb [6]):

$$P_n(t_n) = \exp(A_n t_n) \equiv \sum_{r=0}^{\infty} A_n^r t_n^r / r!. \tag{5}$$

While simple in appearance, (5) is complex since it requires the exponentiation of the matrix $A_n t_n$. The elements of $P_n(t_n)$ are *not*[2] exp $(a_{ij} t_n)$ but rather given by the infinite

[2]For example, if

$$A_n t_n = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

then

$$\exp(A_n t_n) = \begin{bmatrix} \cosh(t_n) & \sinh(t_n) & 0 \\ \sinh(t_n) & \cosh(t_n) & 0 \\ 0 & 0 & \exp(t_n) \end{bmatrix}.$$

convergent series in equation 5. Fortunately, optimal control theory provides a computational procedure for $P_n(t_n)$. Athans and Falb [6] show that if $P_n(t_n)$ has independent eigenvectors,[3] $P_n(t_n)$ can be computed analytically as $E_n\,[\exp(\Lambda_n t_n)]E_n^{-1}$ where Λ_n is the diagonal matrix of eigenvalues, λ_{nj}, of A_n, $[\exp(\Lambda_n t_n)]$ is a diagonal matrix with elements $\exp(\lambda_{nj} t_n)$, and E_n is the matrix of eigenvectors. Since A_n is a differential matrix by definition (rows sum to zero), it has at least one eigenvalue equal to zero. The other eigenvalues can be shown to have negative real components. Thus each element of $\exp(A_n t_n)$ is equal to a constant (the equilibrium probability) plus a series of exponentially decaying terms. Furthermore, by using the series definition of $P_n(t_n)$ one can show that $P_n(t_n)$ is a stochastic matrix (positive values, rows sum to 1.0) for any A_n satisfying the definitions above. Together these results are computationally convenient since (1) $P_n(t_n)$ can be readily obtained from A_n using widely available eigenstructure computer packages and (2) $P_n(t_n)$ is automatically a matrix of probabilities for any A_n with nonnegative off-diagonal elements and with a_{jj} defined such that the rows sum to zero.

To this point we have dealt with one consumer. To achieve market response we have to aggregate consumers in some manner. We address this issue in the next section. For ease of exposition only, the present development assumes complete heterogeneity (each consumer has a different A_n matrix).

Explanatory Variables

The above analysis is useful for predicting behavior once the parameters of the process (a_{jkn}'s) are known. However, for modeling consumer response to managerial actions we must allow the parameters of the process to depend on variables, such as advertising or price, that managers can control.

We first recognize that both the transition probabilities, q_{ijn}, and the flow rate parameters, μ_{in}, can be recovered from A_n if we know q_{iin}. I.e., $q_{ijn} = a_{ijn}(1 - q_{iin})/\sum_{k\neq i} a_{ikn}$ and $\mu_{in} = \sum_{k\neq i} a_{ikn}/(1 - q_{iin})$. Thus we let the explanatory variables directly affect the elements of A_n. Note that the elements, a_{ikn}, of A_n have an intuitive interpretation as the flow rates from state S_i to state S_k. I.e., the probability that the consumer (who is in state S_i) flows to state S_k in time Δt_n is simply $a_{ikn}\Delta t_n$. We discuss estimation of q_{ii} below.

Let x_{ijln} be the value that the lth explanatory variable takes on during the nth observation period for transitions from state S_i to state S_j. The dependence on the transition $(i \to j)$ is important since it lets some variables, such as trial incentives, affect only certain states, such as flows into trial states while other variables, such as relative preference, take on different values depending on the two states that are involved.

For example, if, for ease of exposition, we model the process in Figure 1 as having Poisson flows ($\rho_i = 1$ for all i), we have five states, $(1U)$ using product 1 unaware of new product, $(2U)$ using product 2 unaware of new product, $(1A)$ using product 1 aware of new product, $(2A)$ using product 2 aware of new product, and (NA) using new product. Suppose we hypothesize that flows among purchase states depend on the relative advertising for each product. This hypothesis is shown mathematically in Table 1a. The minus sign $(-)$ indicates the diagonal values are chosen so the rows sum to zero. We might also hypothesize that the rate at which consumers become aware of the new product is independent of whether they are using product 1 or product 2 but dependent upon the advertising of the new product. This hypothesis is shown in Table 1b. Other hypotheses would imply different structures for the X_{ln} matrices. Table 1

[3] If the eigenvalues are not independent, one can use a generalized diagonalization, known as the Jordan canonical form (Noble [95, Theorem 11.12]), which is only slightly more complex. Alternatively, one can perturb the eigenvectors to make them independent (Noble [95, p. 374]).

TABLE 1

Example X_{ln} Matrices for Relative Advertising Affecting Purchase and Actual Advertising Spending Affecting Awareness (α_j = advertising index for product j)

$$\begin{array}{c} \\ 1U \\ 2U \\ 1A \\ 2A \\ NA \end{array}\left[\begin{array}{ccccc} 1U & 2U & 1A & 2A & NA \\ - & \alpha_2/\alpha_1 & & & \\ \alpha_1/\alpha_2 & - & & & \\ & & - & \alpha_2/\alpha_1 & \alpha_N/\alpha_1 \\ & & \alpha_1/\alpha_2 & - & \alpha_N/\alpha_2 \\ & & \alpha_1/\alpha_N & \alpha_2/\alpha_N & - \end{array}\right]\cdot W_1 \qquad \left[\begin{array}{cccccc} & 1U & 2U & 1A & 2A & NA \\ 1U & - & & \alpha_N & & \\ 2U & & - & & \alpha_N & \\ 1A & & & & & \\ 2A & & & & & \\ NA & & & & & \end{array}\right]\cdot W_2$$

might represent only two of many explanatory variables in the model. Table 2 lists a few of the explanatory variables that we have used in applications of the methodology.

As Table 2 illustrates, the explanatory variables can be defined directly, e.g., advertising, or can be the outputs of submodels, e.g., relative preference can be based on product characteristics or product positioning. For example, in the empirical test, preference was the output of a logit model which was based on perceived product features. Similarly, one can use retail price as an explanatory variable but model retail price as a function of other variables such as wholesale price.

We allow each of the explanatory variables to have a differential impact on the flow rates and we assume that the causal relationship can be modeled as linear-in-the parameters.[4] I.e.,

$$a_{ijn} = \sum_l w_l x_{ijln} \tag{6}$$

where the w_l are parameters to be estimated. To assure that $a_{ii} = -\sum_{j \neq i} a_{ij}$ we define $x_{iiln} = -\sum_{j \neq i} x_{ijln}$. We also define variables x^0_{iiln} to carry the information about flow from S_i to S_i. For example, in Table 1a we would define $x^0_{iiln} = \alpha_i/\alpha_i$ and in Table 1b, $x^0_{iiln} = 0$. If we define a^0_{iin} by $\sum_l w_l x^0_{iiln}$ then q_{iin} can be determined by the equation,

$$q_{iin} = a^0_{iin} \Big/ \Big(a^0_{iin} + \sum_{j \neq i} a_{ijn} \Big),$$

once the w_l are estimated. By having the explanatory variables carry information on self-flows we avoid the usual continuous time analysis problem of unidentified self-flows. In the empirical tests of the methodology (section 8) the definition of x^0_{iiln} was obvious, e.g., advertising for i, and the models appear to predict well. However, we

TABLE 2

Example Explanatory Variables

	Variable	*Flows Affected*
1.	Innovator index	to awareness, trial
2.	Introductory price-off deal	to trial
3.	Advertising (absolute)	to awareness
	(relative)	among products
4.	Social and personal norms	to new product
5.	Relative preference including	among products
	a) product improvements	
	b) product positioning	
6.	Word of mouth (cumulative trial)	to awareness
7.	Consumer search	to awareness

[4]Nonlinearities and interactions are handled as in regression by properly redefining the explanatory variables.

caution the reader that great care should be exercised in the selection of x_{iiln}^0. Alternatively one can double the number of behavioral states to avoid self-flows. This formulation avoids the issue of defining x_{iiln}^0, but comes at a cost of increased sample size requirements.

(6) is a generalization of the relative attractiveness measures axiomatically developed by Bell, Keeney, and Little [14] and Barnett [8]. To see this recognize that for independent Poisson processes the underlying transition probabilities, q_{ijn}, are given by $q_{ijn} = a_{ijn}/\sum_k^0 a_{ikn}$ where the summation notation, $\sum_k^0$, indicates a_{iin}^0 is substituted for a_{iin}. Thus the event-wise transition probabilities are given by

$$q_{ijn} = \Big(\sum_l w_l x_{ijln}\Big) \Big/ \sum_k^0 \Big(\sum_l w_l x_{ikln}\Big) \tag{7}$$

which is an (US)/(US + THEM) model in the tradition of Luce's model [83] and the logit model (McFadden [82]). Viewed in this way, our procedure is a generalization of this class of models to continuous time and with the added potential of a wide range of behavior states to include diffusion and experience phenomena.

4. Estimation

In order for our model to be practical we must be able to estimate the w_l's. In general $P_n(t_n)$ varies across the population. If the model is well specified then the variation in the explanatory variables (x_{ijln}'s) should explain the variation in consumer response. Alternatively one can group consumers into homogeneous segments and analyze each segment separately. We introduce the subscript, c, for *c*onsumer to indicate variation in the explanatory variables and resulting consumer response. (5) is rewritten as:

$$P_{nc}(t_n) = \exp(A_{nc}t_n) = \exp\Big(\sum_l w_l X_{nlc} t_n\Big), \tag{8}$$

where X_{nlc} is the matrix ($i \to j$ transitions) for the lth explanatory variable for the cth consumer in the nth observation period. (In a segmented analysis w_l would vary by segment.)

This method of handling variation across consumers is somewhat different than the technique normally used in the stochastic brand selection literature. For example, Bass [10], Blattberg and Sen [16], [17], Givon and Horsky [37], Jones and Zufryden [61], Kalwani and Morrison [64], and Lilien [72], [73] characterize heterogeneity as a probability distribution over brand selection probabilities where the distribution depends on aggregate or segmented statistics. As in the econometric literature (e.g., McFadden [83]), (8) assumes that the explanatory variables vary across consumers and affect $P_{nc}(t_n)$ through the w_l's. The result is a derived distribution over $P_{nc}(t_n)$. To avoid confusion, we use the econometric term, 'disaggregate', to describe variation in explanatory variables and the marketing term, 'heterogeneous' to describe probability distributions over brand selection probabilities.

General Maximum Likelihood

Since our model predicts probabilities and we observe events, $S_i(T_{n-1})$ and $S_n(T_n)$ we can use maximum likelihood estimates to determine the w_l. (Maximum likelihood estimators are consistent, asymptotically efficient, and a function of minimal sufficient statistics.)Let δ_{ijnc} be a state indicator variable such that $\delta_{ijnc} = 1$ if consumer c is in state S_i at time T_{n-1} and in state S_j at time T_n and $\delta_{ijnc} = 0$ otherwise. Let $\{M\}_{ij}$ be the $i-j$th element of matrix M. Then the general log-likelihood function, L_0, for our model is:

$$L_0 = \log\left[\prod_n \prod_c \prod_{(i,j)} \left\{\exp\left(\sum_l w_l X_{ncl} t_n\right)\right\}_{ij}^{\delta_{ijnc}}\right]. \tag{9}$$

If we define $\mathcal{C}_{ijn} = \{c | \delta_{ijnc} = 1\}$ then L_0 can be simplified to be:

$$L_0 = \sum_n \sum_{(i,j)} \sum_{c \in \mathcal{C}_{ijn}} \log\left\{\exp\left(\sum_l w_l X_{ncl} t_n\right)\right\}_{ij}. \tag{10}$$

The maximum likelihood estimators, w_l^*, are then the w_l that maximize L_0. The maximization is unconstrained[5] since the definition of x_{ulc} assures that A_{nc} is a differential matrix (rows sum to zero) hence exp $(A_{nc}t_n)$ is a stochastic matrix (rows sum to one). If X_{ncl} were a scalar rather than a matrix we could simplify (10) further, but $\exp(M_1)^*\exp(M_2) \neq \exp(M_1 + M_2)$ for matrices unless M_1 and M_2 commute and $\prod_c \{M\}_{ij} \neq \{\prod_c M\}_{ij}$ for matrices (Aleksandrov, et al., [3, p. 93]).

Maximization of L_0 is not yet computationally feasible without further simplification. For example, a gradient search algorithm would require $N \cdot C$ exponentiations of a large matrix for every iteration. [$N \cdot C =$ (number of time periods)*(number of consumers in sample)]. Thus, while general maximum likelihood estimators exist in theory, we require simplification for practical solutions. What follows are three potentially practical maximum likelihood solutions, each successively less restrictive.

Case 1: Homogeneous process. As indicated in §2, many of the stochastic models have described consumer behavior quite well with homogeneous processes, which in our case is $X_{ncl} = X_{nl}$ for all c. Furthermore, Givon and Horsky [35] present theoretical arguments that in many cases a homogeneous process is a good approximation to a heterogeneous process. Similar results may hold for a disaggregate process. If X_{ncl} is homogeneous then equation 10 can be simplified and each iteration of the gradient search algorithm requires at most $N \cdot S^2$ matrix exponentiations where S is the number of states. If the Givon-Horsky conditions apply to disaggregate processes, we will show such processes can be approximated with a least squares solution.

Case 2: Macro-flow analysis. In 1970, Urban [108] introduced the concept of macro-flow to deal with the infeasibility of implementing the micro-analytic models of Amstutz [4] and Herniter and Cook [47]. We face a similar problem. In macro-flow analysis we represent the population as a composite process. (This is not an assumption of homogeneity but rather a representative approximation of consumer response as an aggregation of individuals with varying explanatory variables.) Urban shows empirically it is indeed a good approximation for many test market analyses.

For sufficiently small Δt_n, the probability that consumer c flows from S_i to S_j is $a_{ijnc}\Delta t_n$. If we have a variety of consumers in state S_i, then for very small Δt_n the probability that one consumer flows from S_i to S_j is $\sum_c a_{ijnc}\Delta t_n$. But we are interested in the number of consumers flowing from S_i to S_j, thus we represent the process as C_{in} consumers with flow rates, $r_{ijn} = (1/C_{in})\sum_c a_{ijnc}\Delta t_n$. ($C_{in}$ = number of consumers in state S_i at the beginning of observation period n.) Following the same derivation for each consumer we get a representative process, R_n, with state equation $P_n(t_n) = \exp(R_n t_n)$. Since each independent consumer is now characterized by $P_n(t_n)$, the log-likelihood function, L_2, for macro-flow analysis is given by

$$L_2 = \sum_n \sum_{(i,j)} C_{ijn} \log\left\{\exp\left(\sum_l w_l R_{nl} t_n\right)\right\}_{ij} \tag{11}$$

[5]Technically, problems could occur if enough w_l were sufficiently negative to force the $a_{ij} < 0$ for $j \neq i$. In such cases we would constrain the w_l's or a subset of the w_l's to be non-negative. However, we have yet to see these problems in practice and thus leave this issue to future research.

where C_{ijn} is the number of consumers flowing from s_i to s_j in t_n,

$$R_{nl} = (1/C_{in}) \sum_{c \in \mathcal{C}_{in}} X_{ncl},$$

and $\mathcal{C}_{in} = \{c \mid S_i(T_{n-1}) = 1\}$. Note that L_2 reduces to the homogeneous case if $X_{ncl} = X_{nl}$ for all c. Again L_2 requires at most $N \cdot S^2$ matrix exponentiations for each iteration and we will show it can be approximated with a least squares solution. See appendix for gradient search procedure and equations. The empirical examples in §8 use the macro-flow analysis.

Case 3: Intermediate analysis. Once we have defined the macro-flow analysis, it is straightforward to extend the analysis to mutually exclusive and collectively exhaustive segments. Thus, depending on computer resources and data resources, the analyst can choose any level of intermediate analysis. If each consumer is a segment by himself then we have (10), if we group all consumers into one segment then we have (11). The consumer-segment likelihood function, L_3, is the same as (11) except that we subscript C_{ijns} and R_{nls} by s to indicate segment s.

$$L_3 = \sum_n \sum_s \sum_{(i,j)} C_{ijns} \log\left\{ \exp\left(\sum_l w_l R_{nls} t_n \right) \right\}_{ij}. \tag{12}$$

Thus, in principle, we can obtain estimates of w_l for homogeneous, macro-flow, and disaggregate processes. But, practically, the maximum-likelihood estimates are difficult to obtain without specialized computer software.

Least Squares Approximation

Because case 1 and case 2 are specializations of case 3 we deal with the segmented analysis. Since the optimization is unconstrained we maximize equation 12 by setting $\partial L_3 / \partial w_l = 0$ for all w_l where:

$$\frac{\partial L_3}{\partial w_l} = \sum_n \sum_s \sum_{(i,j)} C_{ijns} \sum_k r_{iknls} \frac{p_{kjns}(t_n)}{p_{ijns}(t_n)}. \tag{13}$$

See appendix for algebra. We show in the appendix that if $p_{ijns}(t_n) = \tilde{p}_{ijns} \equiv C_{ijns}/C_{ins}$ then the expression in (13) vanishes and we have achieved at least a local maximum. (The second order conditions, negative definite Hessian, are given in the appendix.) Thus a good approximation to the maximum likelihood solution is:

$$\exp\left(\sum_l w_l R_{nls} t_n \right) \approx \tilde{P}_{ns} \quad \text{for all } n \text{ and } s \tag{14}$$

where $\tilde{P}_{ns}$ is the matrix with elements equal to the observed frequencies, C_{ijns}/C_{ins}.

However, (14) is still a highly nonlinear set of equations. To simplify (14) we use the special eigenstructure properties of exp(A). Let $\tilde{\Lambda}_{ns}$ be the diagonal matrix of eigenvalues for $\tilde{P}_{ns}$ and let $\tilde{E}_{ns}$ be the corresponding eigenvectors. Let $[\log \tilde{\Lambda}_{ns}]$ be a diagonal matrix with elements equal to the logarithms of the eigenvalues.[6]

If the elements of $\tilde{\Lambda}_{ns}$ are not positive then $[\log \tilde{\Lambda}_{ns}]$ has imaginary components (Hildebrand [49, pp. 514-516]). A sufficient condition for the elements of $\tilde{\Lambda}_{ns}$ to be positive is that $\tilde{P}_{ns}$ be positive definite, but in general a stochastic matrix, $\tilde{P}_{ns}$, is not necessarily positive definite. A necessary condition for $\tilde{P}_{ns}$ to be positive definite is that $\tilde{p}_{iins} + \tilde{p}_{jjns} > \tilde{p}_{ijns} + \tilde{p}_{jins}$ for all i and $j, i \neq j$. Empirically this will occur when the

[6] As before, we either use the Jordan canonical form or the perturbation theory for $\tilde{\Lambda}_{ns}$ when the eigenvectors of $\tilde{P}_{ns}$ are not independent. It is easy to show that a Jordan canonical form also exists for $[\log \tilde{\Lambda}_{ns}]$. Fortunately, the need for the Jordan form "tends to occur with a frequency near zero" (Singer and Spilerman [102, p. 28]).

diagonal elements are relatively large (e.g., $p_{11} + p_{22} > 1$ for the two states). Thus if we select the interval of measurement, t_n, to be sufficiently short, the matrix, $[\log \tilde{\Lambda}_{ns}]$, will have only real components. This condition makes intuitive sense since we expect that it would be easier to determine the dynamic effects of the control variables if we make observations of short enough duration to observe those effects. For example, if, on average, consumers purchase analgesics once per month, t_n should not be significantly larger than a month. Yearly measurements would miss the purchase dynamics. For alternative necessary conditions and for perturbation techniques to deal with violations in the necessary conditions see Singer and Spilerman [102, p. 11].

We further investigate this phenomena with the simulations in §7. The simulation results show that the estimation is still feasible when t_n is large enough to produce imaginary elements in $[\log \Lambda_{ns}]$, however, the mean absolute errors of the estimates of w_l increase for large t_n. Thus the positive definite conditions provide guidelines for selecting the periods of observation. However, the model still applies (but with greater estimation error) when $\tilde{P}_{ns}$ is not positive definite.

To simplify (14), let U_{ns} be a matrix defined such that $\exp(U_{ns}) = \tilde{P}_{ns}$. Note that $\exp[\log \Lambda_{ns}] = \tilde{\Lambda}_{ns}$ because both are diagonal matrices. Reduce $\tilde{P}_{ns}$ to its diagonal form (Noble [95, Theorem 11.3]), i.e., $\tilde{P}_{ns} = \tilde{E}_{ns}\tilde{\Lambda}_{ns}\tilde{E}_{ns}^{-1}$. Combining these results we get $\exp(U_{ns}) = \tilde{E}_{ns} \exp[\log \tilde{\Lambda}_{ns}]\tilde{E}_{ns}^{-1} = \tilde{P}_{ns}$. Athans and Falb [6, pp. 132-138] show that e^M and M have the same eigenvectors and functionally related eigenvalues, thus $U_{ns} = \tilde{E}_{ns}[\log \tilde{\Lambda}_{ns}]\tilde{E}_{ns}^{-1}$. Finally recognizing that $U_{ns} = \sum_l w_l R_{nls} t_n$ we substitute in (14) to obtain:

$$\textstyle\sum_l w_l R_{nls} t_n \approx \tilde{E}_{ns}\big[\log \tilde{\Lambda}_{ns}\big]\tilde{E}_{ns}^{-1} \quad \text{for all } n, s. \tag{15}$$

(15) is the computational result we seek. The right hand side of equation 15 is simply a transformation of the frequency data. Intuitively, it is useful to think of the right hand side as a numerical estimate of the flow rates, $\tilde{A}_{ns}$. Efficient computer algorithms exist for such eigenstructure transformations (e.g., IMSL [56]). The left hand side is linear in the unknown parameters. Thus a least squares regression solution to (15) will approximate the maximum likelihood estimates, w_l^*. As illustrated in Table 3, for one time period and one segment, (15) is a matrix equation which represents $\mathbb{S} \cdot N \cdot S^2$ scalar equations where $\mathbb{S}$ is the number of segments. Since the shadow state and diagonal flows are functions of the other (independent) flows, the regression is run across the $\mathbb{S} \cdot N \cdot S^0(S^0 - 1)$ equations that are independent where S^0 is the number of real (not shadow) states. Of course, if a flow is defined as zero (e.g., purchase new product to unaware of new product) it is not included in the regression. §7 provides simulation

TABLE 3

Schematic Representation of the Regression Equations Implied by (15) for One Time Period and One Consumer Segment

$$W_1 \cdot \begin{bmatrix} -5 & 2 & 3 & \cdots \\ 0 & -1 & 1 & \cdots \\ & \vdots & & \\ 4 & 2 & \cdots & -6 \end{bmatrix} + W_2 \begin{bmatrix} \\ X_{2NS} \\ \\ \end{bmatrix} + \cdots + W_L \begin{bmatrix} \\ X_{LNS} \\ \\ \end{bmatrix}$$

$$\approx (1/T) \begin{bmatrix} \\ \tilde{E}_{NS} \\ \\ \end{bmatrix} \begin{bmatrix} \mathrm{LOG}\,\lambda_{1NS} & & 0 \\ & \mathrm{LOG}\,\lambda_{2NS} & \\ 0 & & \ddots \end{bmatrix} \begin{bmatrix} \\ \tilde{E}_{NS}^{-1} \\ \\ \end{bmatrix}$$

$$= \begin{bmatrix} -0.84 & 0.32 & 0.36 & \cdots & 0.15 \\ 0 & -2.1 & 1.0 & \cdots & 1.1 \\ & & \vdots & & \\ 0.6 & 1.2 & 0.9 & \cdots & -2.7 \end{bmatrix} \qquad \begin{array}{l} \tilde{E}_{NS} = \text{EIGENVECTORS OF } \tilde{P}_{NS} \\ \lambda_{JNS} = J\text{TH EIGENVALUE OF } \tilde{P}_{NS} \end{array}$$

results testing the accuracy of the least squares approximation for one consumer model.

At this point we digress to better understand (15) by considering a two-state process. For a one segment, two-state process, (15) becomes

$$\sum_l w_l r_{12ln} t_n = -(\tilde{p}_{12n}/\tilde{p}_{12n} + \tilde{p}_{21n})\log(1 - \tilde{p}_{12n} - \tilde{p}_{21n}),$$

$$\sum_l w_l r_{21ln} t_n = -(\tilde{p}_{21n}/\tilde{p}_{12n} + \tilde{p}_{21n})\log(1 - \tilde{p}_{12n} - \tilde{p}_{21n})$$

which gives us $2N$ equations to estimate L parameters. Note also that the right hand side is real iff $\tilde{p}_{12n} + \tilde{p}_{21n} < 1$. With (5) we can show that $p_{12n} + p_{21n} = 1 - \exp[-(a_{12n} + a_{21n})t_n]$. Thus as t_n gets large, $p_{12n} + p_{21n} \rightarrow 1$ and $\log(1 - p_{12n} - p_{21n}) \rightarrow -\infty$.

Interpreting these equations intuitively we see that estimates can be obtained as long as the process has not reached equilibrium ($t_n \rightarrow \infty$). Indeed, $E_{ns}[\log \Lambda_{ns}]E_{ns}^{-1}$ (as opposed to $\tilde{E}_{ns}[\log \tilde{\Lambda}_{ns}]\tilde{E}_{ns}^{-1}$) is unique for finite t_n. Since the $\tilde{p}_{ijn}$ are estimates of the p_{ijn}, the above equations indicate that as the process nears equilibrium ($t_n \rightarrow \infty$) and $(1 - p_{12n} - p_{21n}) \rightarrow 0$, the right hand side, $\log(1 - \tilde{p}_{12n} - \tilde{p}_{21n})$, of the two-state process becomes extremely sensitive to errors in $\tilde{p}_{ijn}$. We extend these results to S^0 states, to show that (15) will give reasonable estimates of the w_l if:

(1) there are enough consumers in each segment so that $\tilde{p}_{ijn}$ is a good estimate of p_{ijn}, and

(2) the observation period, t_n, is short relative to the time it takes the process to reach equilibrium.

These conditions provide useful guidelines to the model builder. While there is no conceptual limit on the number of explanatory variables, the complexity of the diffusion process, nonstationarity, and consumer segmentation, the regression format shows explicitly what sample sizes are required to identify these effects. First recognize that the dependent measure is a transformation of $\tilde{p}_{ijns}$. Thus, if C^0 is the average number of observations per cell to get a "reasonable" estimate of a probability,[7] then the sample must be at least as large as C^0 times the number of independent $\tilde{p}_{ijns}$'s. This implies the number of observations *per behavioral state per time period* must be greater than $C^0 \cdot \mathbb{S} \cdot S^0$. Finally, the number of independent variables, L, must be less than the number of independent equations, $\mathbb{S} \cdot N \cdot S^0 \cdot (S^0 - 1)$.

Summary of Estimation

This section has derived a battery of estimation procedures for the general continuous time Markov process which, in our case, represents semi-Markov consumer models with negative exponential or Erlang holding times. The three maximum-likelihood procedures, which vary in their degree of consumer aggregation, provide the theoretical structure for estimation. While they are thoeretically feasible, we know of no existing computer software that can readily implement these procedures. (A gradient search maximum-likelihood algorithm is sketched in the appendix.)

The key result of this section, (15), shows that there is a simple approximation to the general maximum likelihood estimators. (15) provides a number of practical advantages. Besides making estimations feasible for large problems (many behavioral states, explanatory variables, time periods, and segments) it enables the model builder to

[7]For example, if we have C^0 observations per cell then the standard deviation of the probability in the i-jth cell is $\sqrt{p_{ijns}(1 - p_{ijns})/C^0}$. Thus a larger C^0 gives less sampling error. §7 investigates how large C^0 should be.

make explicit tradeoffs among (1) the complexity of the behavioral process (S^0), (2) the number of managerial variables (L), (3) the number of observation periods (N), (4) the length of the observation periods (t_n), (5) the number of consumer segments ($\mathbb{S}$), and (6) sample size.

5. Statistics of Managerial Interest

Once the w_l's are estimated, the stochastic consumer response process is fully specified. A marketing scientist can use the w_l's to determine A_{ns}, and hence $P_{ns}(t_n)$, for any set of levels of the marketing variables (X_{nsl}'s). For example, Table 4 provides an abridged process specification for 10 explanatory variable matrices (X_{lns}'s) and an example set of levels for the managerial control variables (the new product's advertising, preference, and promotional deals) for a five-state process consisting of aware of new product (A), intent to purchase new product (I), purchase of new product (NP), purchase product 1 ($P1$), and purchase product 2 ($P2$). Assume Erlang ($\rho_t = 2$) flows out of NP, $P1$, and $P2$ and Poisson ($\rho_t = 1$) flows our of A and I.

The manager is likely to be interested in further descriptions of the process such as expected sales, market share, cumulative trial, cumulative awareness, and purchases on deal. While such statistics can be obtained from a stochastic process by Monte Carlo simulation, e.g., see Urban [108], it is possible by exploiting the matrix properties of the continuous time Markov model to provide closed form solutions for these statistics.

TABLE 4

Example Process Specification and Levels of the Marketing Variables over 10 Periods

PROCESS SPECIFICATION		
Advertising	*Flows*	*Elements in X_{ln} matrix*
$w_1 = 0.1$	Product → Product	Relative Adv. Index
$w_2 = 0.4$	Into Awareness	$(1 - \beta_{n-1})\alpha_0$
$w_3 = 0.2$	Awareness → Intent	
	Intent → Purchase	α_0
Preference		
$w_4 = 0.4$	Product to Product	Relative Preference Index
$w_5 = 0.3$	Out of Awareness	p_j/p_0'
$w_6 = 0.3$	Out of Intent	p_j/p_0'
Promotion (e.g., deals)		
$w_7 = 0.2$	Out of Product 1	d_0
$w_8 = 0.2$	Out of Product 2	d_0
Word of Mouth		
$w_9 = 0.1$	Into Awareness	β_{n-1}
Active Search by Consumers		
$w_{10} = 0.4$	Into Awareness	$(1 - \beta_{n-1})p_0'/p_j$

LEVELS OF THE MARKETING MIX VARIABLES*	$n =$	1	2	3	4	5	6	7	8	9	10
New Product Advertising	(α_0)	0.05	0.10	0.20	0.30	0.40	0.50	0.40	0.30	0.30	0.30
Product 1 Advertising	(α_1)	0.72	0.68	0.60	0.53	0.45	0.38	0.45	0.53	0.53	0.53
Product 2 Advertising	(α_2)	0.23	0.22	0.20	0.17	0.15	0.12	0.15	0.17	0.17	0.17
Pre-trial Preference	(p_0')	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
Post-trial Preference	(p_0)	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Product 1 Preference	(p_1)	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49
Product 2 Preference	(p_2)	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21
Deal Level	(d_0)	1	1	1	0	0	0	0.5	0.5	0	0

*Units are illustrative and hence arbitrary β_{n-1} = cumulative trial at time T_{n-1}.

$$\begin{array}{r} \\ \text{JTH ROW} \\ \\ \end{array} \left[\begin{array}{c} a_{ikn} \\ \hline 0\,0\cdots\qquad\qquad 0 \\ \hline a_{ikn} \end{array}\right] = {}_{1}A_N$$

FIGURE 3. Flow Matrix for Calculating Penetration.

The advantage of closed form solutions is that they simplify the computational demands of the model.

We derive formulas for three types of statistics: (1) cumulative statistics for one time only events such as trial, (2) rate statistics for repeated events such as sales, and (3) equilibrium statistics. For simplicity of exposition we drop the subscript s. Aggregation of these statistics over segments requires summation (averaging) over segments (s).

Penetration and Other Cumulative Statistics

Penetration, the percent of consumers purchasing a given product at least once in time t_n, is an important output from stochastic models. Because the process is stochastic, a greater percentage of consumers actually purchase a product at least once than would be suggested by market share. To compute penetration we redefine state S_j as an absorbing state (trapping state) such that ${}_1q_{jj} = 1$ and ${}_1q_{ji} = 0$ for $j \neq i$. In other words, every consumer who flows into state S_j stays there. If S_j corresponds to purchasing product j then, for the trapping process, penetration is simply the weighted sum of the probabilities that the consumer flows from S_i to S_j.

If we define ${}_1A_n$ such that ${}_1a_{ik} = a_{ik}$ for $i \neq k$ and ${}_1a_{jk} = 0$ (see Figure 3) then penetration is given by:

$$\text{penetration (into state } S_j) = \sum_i \pi_i(T_{n-1})\,{}_1p_{ijn}(t_n) \qquad (16)$$

where

$${}_1P_n(t_n) = \exp({}_1A_n t_n)$$

and $\pi_i(T_{n-1})$ is the probability that the consumer is in state S_i at time T_{n-1}. (16) is used recursively when calculating penetration over more than one observation period. I.e., $\pi_i(T_n) = \sum_i p_{ijn}(t_n)\pi_j(T_{n-1})$, recursively back to $n - 1 = 0$. The starting states, $\pi_j(T_0)$, must be based on direct measurement or assumption.

Cumulative trial, cumulative awareness, cumulative penetration of a deal, cumulative use of a free sample, and other cumulative statistics are all computed by suitably defining the absorbing states. For example, if we want to compute cumulative awareness for the process in Figure 1, we make "awareness" an absorbing state. If we want to separate awareness due to advertising and promotion from awareness due to diffusion phenomena (word of mouth and active search) we define two (or more) absorbing states such that flows due to advertising go into one awareness state and flows due to other variables go into the other awareness state. Figure 4 displays cumulative awareness computed in this manner for the example process in Table 4.

Sales (Mean and Variance) and Other Rate Statistics

To determine the number of purchases, M_j, of product j we determine the number of times the consumer flows into state S_j in time t_n. Let $f_{iln}(m_j, t_n)$ be the probability

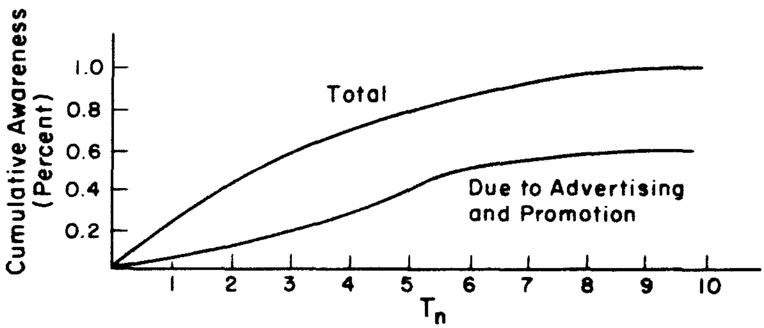

FIGURE 4. Cumulative Awareness (for the Process in Table 4).

that the consumer has made m_j purchases in time t_n *and* is in state S_l given he started in state S_i at the beginning of the observation period. I.e.:

$$f_{iln}(m_j, t_n) = \text{Prob}\{m_j \text{ purchases}, S_l(T_n) = 1 | S_i(T_{n-1}) = 1\}. \tag{17}$$

Following the same type of reasoning used to derive (2), we get:

$$f_{iln}(m_j, t_n + \Delta t_n) = \Big(1 - \sum_{k \neq l} a_{lkn}\Delta t_n\Big) f_{iln}(m_j, t_n) + \sum_{k \neq l} f_{ikn}(m_j, t_n) a_{kln} \Delta t_n, \qquad l \neq j,$$

$$f_{ijn}(m_j, t_n + \Delta t_n) = \Big(1 - \sum_{k \neq j} a_{jkn}\Delta t_n - a^0_{jjn}\Delta t_n\Big) f_{ijn}(m_j, t_n)$$

$$+ \sum_{k \neq j} f_{ikn}(m_j - 1, t_n) a_{kjn} \Delta t_n + f_{ijn}(m_j - 1, t_n) a^0_{jjn} \Delta t_n. \tag{18}$$

We redefine matrices to simplify (18). Let ${}_2A_n$ be defined such that ${}_2a_{kln} = a_{kln}$ for $l \neq j$, ${}_2a_{kjn} = 0$ for $k \neq j$, ${}_2a_{jjn} = a_{jjn} - a^0_{jjn}$. See Figure 5. Let ${}_3A_n$ be defined such that ${}_3a_{kl} = 0$ for $l \neq j$, ${}_3a_{kjn} = a_{kjn}$ for $k \neq j$, ${}_3a_{jjn} = a^0_{jjn}$. See Figure 5. Note that $A_n = {}_2A_n + {}_3A_n$.

With these definitions we can rewrite (18) in matrix form:

$$\frac{d}{dt} F_n(m_j, t_n) = F_n(m_j, t_n)[{}_2A_n] + F_n(m_j - 1, t_n)[{}_3A_n]. \tag{19}$$

To calculate the mean and variance of sales we first determine the moment generating function, $\check{F}_n(z, t_n)$, for $F_n(m_j, t_n)$ by taking the discrete geometric transform of (19).

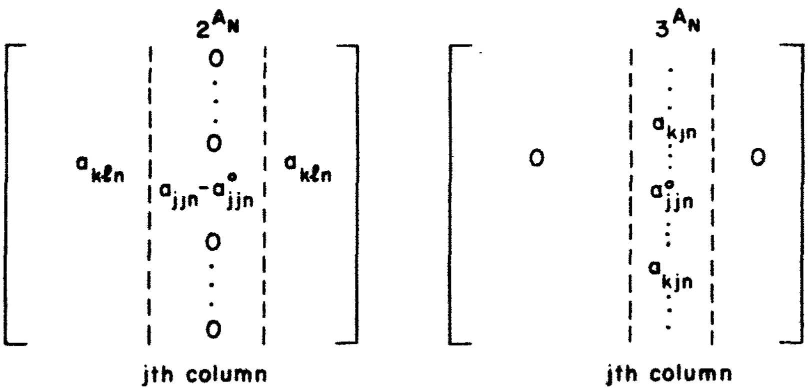

FIGURE 5. Partition of the Flow Matrix to Calculate Sales.

This gives us:

$$\frac{d}{dt} F_n(z, t_n) = \check{F}_n(z, t_n)\left[{}_2A_n + z({}_3A_n)\right]. \tag{20}$$

The solution to (20) with boundary conditions, $F_n(z,0) = I$, is:

$$\check{F}_n(z, t_n) = \exp\left[{}_2A_n + z({}_3A_n)\right]. \tag{21}$$

Those readers who are familiar with Howard's [53], [55, pp. 604-611] transform analysis for counting transitions will recognize Figure 5 as a matrix formulation of Howard's flow graph procedure for "tagging". However the closed form solution in (21) provides computational simplifications over the flow graph analysis since we can proceed from (21) to determine direct formulae for the mean and variance of sales.

We use the properties of the moment generating function to determine the mean and variance of conditioned sales. Let $e_{ikn}(m_j, t_n)$ be the expected value of sales given that the consumer began in state S_i and is in state S_k at time t_n, let $v_{ikn}(m_j, t_n)$ be the corresponding variance, then in matrix notation:

$$E_n(m_j, t_n) = \frac{d}{dz} \check{F}_n(z, t_n)\big|_{z=1} = \left[\exp(A_n t_n)\right]\left[{}_3A_n t_n\right]$$

or

$$E_n(m_j, t_n) = P_n(t_n)\left[{}_3A_n t_n\right]. \tag{22}$$

A similar derivation gives:

$$V_n(m_j, t_n) = P_n(t_n)\left\{{}_3A_n t_n + ({}_3A_n t_n)^2\left[I - P_n(t_n)\right]'\right\}. \tag{23}$$

Finally we recognize that expected sales is simply the weighted sum of conditional sales, thus:

$$\text{Expected sales} = \sum_i \pi_i(T_{n-1})\left\{\sum_k e_{ikn}(m_j, t_n)\right\}$$

or using (22):

$$\text{Expected sales} = \sum_i \sum_k{}^0 \pi_i(T_{n-1}) p_{ikn}(t_n) a_{kjn} t_n \tag{24}$$

where $\sum_k^0$ means we use a_{jjn}^0 in the sum. The variance in sales can be computed by expanding (23) in a similar manner.

(24) is intuitive and particularly simple to use. Given A_n, (24) is a function of known values which lends itself readily to computation via an electronic computer. Furthermore since it is explicitly dependent on t_n, we can plot the growth (or decline) in expected sales over time. Finally we use (24) recursively if we want expected sales over a number of observation periods.

For example, the solid curve in Figure 6 is a plot of expected sales for the example process in Table 4. The downturn in periods 4 and 9 result from the retraction of a deal in those periods. To gain further insight into the effect of marketing variables we used (24) to compute what expected sales would be without any dealing (dotted curve) and without any advertising (dashed curve). The marginal impact of dealing (or advertising) can then be computed by subtracting the dotted curve (or dashed curve) from total sales.

Other statistics such as the number of purchases of a new product (including trial) and the number of purchases of a product being sold with a reduced price can be

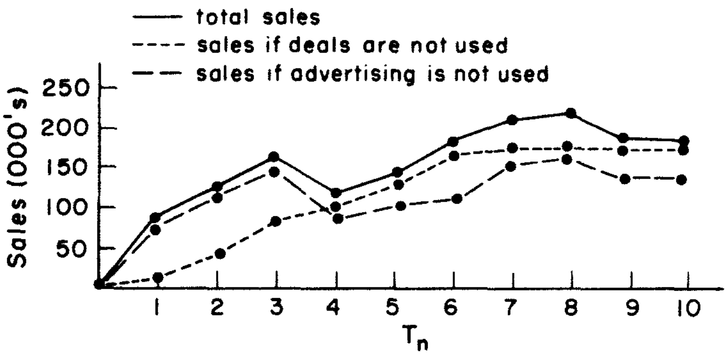

FIGURE 6. Expected Sales per Million Customers (for the Process in Table 4).

computed by suitably partitioning the A_n matrix. The dynamics of market share are obtained by computing the expected sales for each product in the market.

Equilibrium Statistics

In test markets and new product launches we are concerned with the transient nature of consumer response, but in the long run we are concerned with market share and penetration under equilibrium conditions. We can calculate equilibrium conditions by allowing $t_n \to \infty$ in the appropriate equations. For simplicity let A be the long run managerial actions. (Drop the subscript n.)

Let $t = 0$ be the time at which the managerial actions stablize. Let π_j be the equilibrium probability that the consumer is in state S_j. Then if the process is ergodic, $\pi_k = \lim_{t\to\infty} p_{ik}(t)$ for all i (Howard, [55]). In other words, the equilibrium probabilities are equal to any row of $\lim_{t\to\infty}\exp(At)$. (If, in equilibrium, the transient states such as awareness and trial are empty, we deal with the ergodic subchain.) Thus one way to determine π_k is to use the non-decaying portion of $p_{ik}(t)$ corresponding to the zero eigenvalue of A. These probabilities are the same for all i. A simpler method, derived by Howard [55], is to solve the matrix equation $\prod' A = 0$ subject to $\sum_k \pi_k = 1$. Using these results in (24) yields:

$$\text{Expected equilibrium sales rate} = \sum_k {}^0\pi_k a_{kj} = \pi_j\left(a_{jj}^0 - a_{jj}\right) \tag{25}$$

where the right hand side comes from using $\prod' A = 0$. Thus the equilibrium market share of product j is given by

$$\text{Equilibrium market share} = \pi_j\left(a_{jj}^0 - a_{jj}\right) / \sum_k \pi_k\left(a_{kk}^0 - a_{kk}\right) \tag{26}$$

which is both intuitive and particularly simple to implement since the π_k comes from an algebraic equation rather than matrix exponentiation.

The penetration rate under equilibrium conditions can be calculated in a similar manner using the limiting values for ${}_1A$.

6. Summary of Theoretical Development

Many innovative marketing science models have been developed over the past 20 years to model consumer response. Stochastic models, which have their roots in empirical data, have been quite successful in describing consumer response and more recently have been modified to include managerial control variables. Diffusion models have a good record of predicting consumer response and recently, they too have been extended to include control variables. Test market and new product models have proven valuable in diagnosing consumer response and suggesting managerial actions. The published evidence in these four research streams (as reviewed in §2) is extensive.

The semi-Markov theory developed in §3 provides an integrative framework for essential elements from these four research streams. For example, the Jeuland, Bass and Wright [57] model is a semi-Markov process where each behavioral state represents 'last purchase was product j.' Transitions are multinomial ($q_{ij} = q_j$ for all j) and interpurchase times are Erlang. §§3, 4, and 5 provide a practical framework to extend the Jeuland-Bass-Wright model to include diffusion states and explanatory variables. The Jones and Zufryden [61] model is a homogeneous two-state semi-Markov process where q_{ij} is determined by a formula similar to (7). Transitions are multinomial and interpurchase times Erlang. §§3, 4, and 5 provide a practical framework to extend the Jones-Zufryden model to include multiple products, diffusion states, and interpurchase times dependent on explanatory variables. Similarly §§3, 4, and 5 provide a theoretical framework and closed form solutions for Urban's [108] model and practical estimation procedures for the Midgley [84] and Balachandran and Deshmukh [7] models.

But generalization alone is not sufficient. A theory must be practical in order to be used. While many of our derivations are algebraically tedious, the end results (estimation and managerial statistics) are easy to implement with standard computer packages. The estimation derived in §4 is feasible because it relies solely on eigenstructure routines to transform the data and on regression to obtain the unknown parameters. Regression is well studied, easy to implement, and readily lends itself to generalizations. The diagnostic use of the model is practical because all cumulative, rate, and equilibrium statistics (such as penetration, sales, and equilibrium share) are given by closed-form formulae that computationally require, at most, matrix manipulation and eigenstructure transforms.

Finally the semi-Markov model is a flexible methodology for representing consumer response. While, *in theory*, the model can handle highly complex diffusion processes, non-stationarity, disaggregate data, and any number of managerial controls, the estimation regression equation makes explicit the tradeoffs among these effects and sample size so that the marketing scientist can customize the model to his application. We envision the following typical modeling scenario: (1) The modeler first selects the consumer states appropriate for modeling the phenomena he needs to model. More states mean more detail but larger samples. Fewer states mean less detail but greater simplicity for evolutionary development. (2) The modeler next determines what variables the manager can control, how these variables impact flows, and formulates the X_{ln} matrices. (3) The needed sample size and observation periods are determined from the theoretical guidelines. Data is collected and regression determines the w_l's. (4) Using the formulae in section 5, the modeler computes the cumulative, rate, and equilibrium statistics for various levels of the managerial control variables. And (5) together the manager and modeler search alternative combinations of the control variables until a "best" set of levels are found. At present step 5 is a manual search aided by the model's diagnostic information. Future research may provide optimization algorithms to automate step 5 and value of information guidelines to allow iteration through steps 1, 2, and 3. Steps 1 through 4 are illustrated with Tables 1–4 and Figures 1, 2 and 4. Other application scenarios might begin with data limitations, the control variables or the consumer process. For example, one might use F-tests based on the regression approximation to investigate whether certain consumer states or control variables add significant explanatory power to the model.

We feel the semi-Markov analysis addresses many important problems. Limitations, unsolved problems and future research are discussed in §9, but first we address the questions: "How well does the least squares approximation recover known data? And, what are the effects of sample size, observation period length, and collinearity?

7. Sample Size, Observation Period, and Collinearity—Simulation Analysis

The literature provides good evidence for the basic modelling assumption—a negative exponential or Erlang semi-Markov stochastic process. However, any specific model (set of consumer states and structure for the X_{ln} matrices) must be empirically validated before it is used to describe managerial actions.

An empirical test confounds a number of errors including specification error,[8] sampling error and estimation error. Before proceeding with an empirical test, we use Monte Carlo simulation to investigate sampling and estimation error without confounding it with specification error. We do this by investigating whether or not the least squares approximation can recover a known model.

We use the illustrative model and example parameters in Table 4 as the known model. If the least squares approximation cannot recover this model, it is unlikely to recover more complex models. If the approximation does recover this model, then our faith in the estimation is increased and the simulation results suggest a first set of guidelines for more complex models. Generalizations beyond the specific results should be viewed with caution and taken as propositions not "theorems".

The accuracy of the approximation will depend upon sample size, observation period and collinearity. The theory suggests the required sample per behavioral state per time period should be larger than $C^0 \cdot \mathbb{S} \cdot S^0$, that the number of control variables should be less than $\mathbb{S} \cdot N \cdot S^0(S^0 - 1)$, and that the estimation works best if the observation period is short compared to the time it takes the process to reach equilibrium. Simulation investigates (1) how large is C^0, (2) how short is short enough, and (3) how badly does the estimation degrade if "long" observation periods are used. Finally all linear systems are subject to problems of multicollinearity among the explanatory variables. Simulation provides a rough guideline to compare errors due to collinearity to errors due to sampling and long observation periods.

Simulation Design

For our analyses we chose a $5 \times 3 \times 2$ design. The 5 levels of sample size include small (50), moderate (100), large (500), very large (1000), and asymptotic (∞). All numbers are per period per behavioral state. The 3 levels of observation period include short relative to purchase cycle ($\tilde{p}_{njj} + \tilde{p}_{nkk} \gg \tilde{p}_{njk} + \tilde{p}_{nkj}$ for $k \neq j$), near the purchase cycle ($\tilde{p}_{njj} + \tilde{p}_{nkk} \gtrsim \tilde{p}_{njk} + \tilde{p}_{nkj}$), and long relative to the purchase cycle ($\tilde{p}_{njj} + \tilde{p}_{nkk} < \tilde{p}_{njk} + \tilde{p}_{nkj}$). With the parameters in Table 4, these levels correspond to $t_n = 0.25$, 1.0, and 3.0 respectively. Finally, the 2 levels of collinearity are (1) the full process of 10 variables (high collinear case) and (2) the same process with two collinear variables deleted (low collinear case).[9]

In empirical applications there are at least four types of error: specification error, measurement error, sampling error, and error due to the fact that the least squares solution approximates the maximum likelihood solution. We call the last error "estimation error". We chose simulation to avoid confounds with specification error and because specification error for regression is well studied. (Theil [107], Green [38], and Morrison [88]). Measurement error for regression is also well studied and there are a variety of analytic results available. See Theil [107, pp. 607-613]. Here we make the

[8] By specification error we mean errors introduced because the selection of consumer states and control variables is only an approximation to the "real world". For a discussion of error types in probabilistic modeling see Koppelman [66].

[9] In the collinear case $\text{cor}(X_{1n}, X_{4n}) = 0.676$, $\text{cor}(X_{2n}, X_{10n}) = 0.604$. In the independent case X_{4n} and X_{10n} are deleted. Similar results were obtained when other pairs were deleted.

standard assumption of general linear models (Green [38, p. 61] that the X_{ln}'s are random variables but "once a particular realization of X_{ln} occurs, for some observation, that realization is held constant and the dependent measure is observed conditionally on that fixed realization." Thus we are most interested in sampling error and estimation error. We introduce sampling error by using a random number generator (SPURT [21]) with the known probabilities, P_n, to create a stochastic process from which a sample is drawn. Estimation error is the error inherent in the approximation. Note that sampling error disappears for asymptotic (∞) sample sizes. I.e., $\tilde{P}_n \equiv P_n$.

For each level of sample size, observation period, and collinearity we (1) use the X_{ln}'s and the known w_l's to create a "true" P_n, (2) use a random number generator to create C_{in} observations from P_n for each starting state S_i and time period n, (3) use the regression across states and time periods to estimate the $\hat{w}_l$'s and (4) compare the process resulting from the $\hat{w}_l$'s to that resulting from the known w_l's. For ease of exposition we report mean absolute error (MAE) in the w_l's and adjusted R^2 for the regression. Other statistics such as the correlation between the estimated $\tilde{P}_n$ (or $\tilde{A}_n$) and the "true" P_n (or A_n) follow the same pattern as MAE and R^2.

Note that for our process, $\mathbb{S} = 1$, $S^0 = 5$, $N = 10$, and $L = 8$ or 10. Thus the degrees of freedom in the regression are more than sufficient to identify L parameters. C^0 will be approximately one-fifth $(1/\mathbb{S} \cdot S^0 = 1/5)$ of the sample size required to obtain reasonable estimates.

Simulation Results

The basic simulation results are shown in Table 5. To help visualize these results we have plotted the marginal values (main effects plus grand mean) of R^2 and MAE for each of the treatment levels. See Figure 7. All ANOVA main effects are significant at the 0.01 level except observation time for MAE which has a significance level of 0.12. (Very short ($t_n = 0.25$) observation periods cause a qualitative upswing in MAE [downswing in R^2] for the low collinear case, but the ANOVA suggests this may be random error. It does not occur for the high collinear case.)

It appears from Table 5, that estimation error is manageable, especially for large samples ($C_{in} \geqslant 500$), short observation periods ($t_n \leqslant 1.0$), and low collinearity, i.e., $R^2 \gtrsim 0.95$ and MAE $\lesssim 0.20$. Furthermore, it appears that samples should be at least 100 consumers to get reasonable goodness of fit measures for a 5-state process. If these

TABLE 5
Simulation Results

ADJUSTED R^2

t_n \ C_{in}	50	100	500	1000	∞	t_n \ C_{in}	50	100	500	1000	∞
0.25	0.59	0.80	0.94	0.96	1.00	0.25	0.68	0.80	0.95	0.97	1.00
1.0	0.77	0.87	0.98	0.98	1.00	1.0	0.29	0.53	0.88	0.95	1.00
3.0	0.38	0.61	0.88	0.77	1.00	3.0	0.01	0.09	0.33	0.09	1.00
	(low collinearity)						(high collinearity)				

MEAN ABSOLUTE ERROR

t_n \ C_{in}	50	100	500	1000	∞	t_n \ C_{in}	50	100	500	1000	∞
0.25	1.11	0.52	0.23	0.24	0.00	0.25	1.81	0.51	0.94	0.14	0.00
1.0	1.03	0.22	0.19	0.14	0.00	1.0	2.49	1.56	1.03	0.75	0.01
3.0	0.78	0.95	0.56	0.22	0.00	3.0	6.37	0.94	1.47	1.68	0.50
	(low collinearity)						(high collinearity)				

Key: c_{in} = sample size (∞ means no sampling error)
t_n = observation period

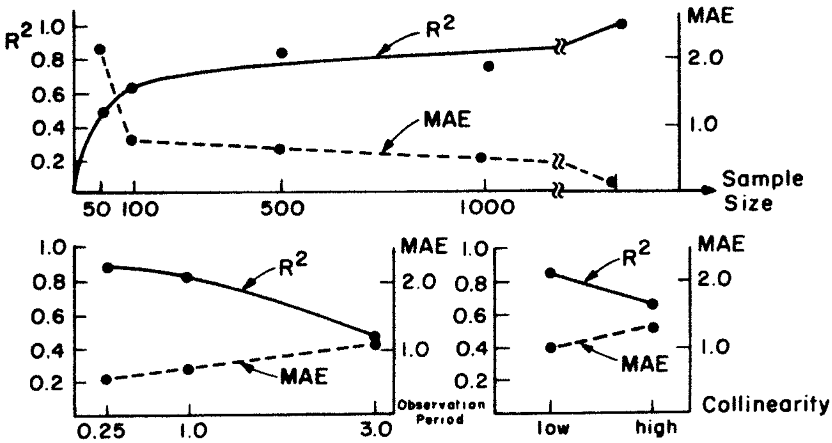

FIGURE 7. Marginal Effects of Sample Size, Observation Period, and Collinearity (R^2 is Adjusted R^2, MAE is the Mean Absolute Error of the Estimates).

results are generalizable then the theoretical sample size guidelines could be used with $C^0 \gtrsim 20$.

Overall, the simulation results are encouraging. They suggest that under the right conditions, the least squares estimation can recover known parameters. However, the simulation results also suggest caution for small samples, long observation periods, and conditions of high collinearity. Finally, it is interesting to note that each simulation run (including data generation, eigenstructure manipulation and regression) for the 5-state, 10-time period model costs less than $5 on a CDC-6600 at $900 per cpu hour.

8. Empirical Applications

The methodology has now been applied twice, once to the launch of a new transportation service (Hauser and Wisniewski [43]) and once to the analysis of purchases of ground coffee (Lange [69]). We describe briefly the former application.

In October 1979 the Village of Schaumburg, Ill. launched a new demand-actuated transportation service, DART, in a community that was previously served only by automobile and conventional bus. To monitor this service we mailed out randomly 16 waves of periodic surveys averaging 625 surveys per wave (response rate 30.4%) This provided the dependent measures and explanatory variables to test the Markov methodology. In addition, we monitored dispatch records for the survey periods to obtain actual ridership. This data is used to test the predictive and external validity of the model based on the methodology.

The basic model consisted of 5 behavioral states: (1) unaware of DART and aware of DART but last mode used was (2) car as a driver, (3) car as a passenger, (4) conventional bus, and (5) DART. The explanatory variables were direct mail (# of pieces), advertising (probabilistic sum of readership times insertions), word of mouth (cumulative trial by the beginning of the observation period), availability (self-reported), budget allocation (self-reported), inertia (a dummy variable for DART), preference, and a constant. Preference was derived from a submodel, a logit model, that computed the probability of a mode being preferred. The explanatory variables in the submodel were measured perceptions of "convenience", "ease of use", "safety" and "normative beliefs". All estimation data were obtained from the surveys after corrections for recall bias.

The model was estimated using data from the first 8 time periods with the regression approximation yielding an adjusted R^2 of 0.82. The significant variables were direct mail/advertising, preference, availability, and the constant.

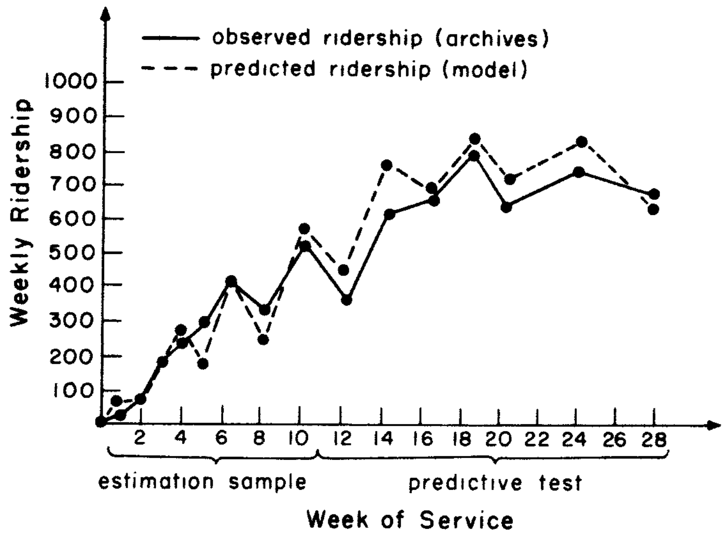

FIGURE 8. Predictive Test for Dynamic Consumer Model Based on Seven Months of DART Service in Schaumburg, Ill.

To test the predictive validity of the model we forecasted for the estimation period (first 8 time periods totalling 11 weeks) *and* for the next 7 time periods totalling 19 weeks. The explanatory variables were obtained from the survey but the model, i.e., the w_l's, were based only on the estimation period. The results are shown in Figure 8.

Note that the dependent variables are transition probabilities in the first 11 weeks, (15), but the predictions are structural outputs of the model in the next 19 weeks, (24). Furthermore, the explanatory measures are taken from surveys, with all their potential biases, but the predictions are compared to actual ridership as measured by dispatch records.

Based on Figure 8 we conclude that the empirical application of a model based on the methodology is encouraging. The correlation between predicted and actual ridership is 0.94 in the estimation period and 0.83 in the prediction period although there is a slight upward bias in the prediction period. Many tradeoffs were made in the empirical application and much was learned about the measurements used to implement the methodology. The interested reader is referred to Hauser and Wisniewski [43].

In another application Lange [69] used automated supermarket checkout data (UPC data) to observe last brand and size (of coffee) purchased. His explanatory variables were price, price-off deal, promotion, inventory, and whether the brand was in or out of stock. Adjusted R^2 was lower, 0.32, but he obtained predictions of transition probabilities which appear promising.

Together these applications indicate the potential empirical utility of the methodology.

9. Conclusions, Limitations and Future Work

The emphasis of this paper is to provide a practical, flexible and integrative structure for modeling dynamic consumer response to marketing strategies. Our main analytic contribution is (1) a practical estimation procedure for a general continuous time Markov process with flow rates linearly dependent on explanatory variables, (2) an ability to estimate the parameters of the continuous time process by observing the process only at discrete intervals, (3) closed form expressions for statistics of managerial interest, including the dynamics of the system, and (4) guidelines for tradeoffs among process complexity, managerial controls, observation period, segmentation, and sample size. In addition, our simulation results suggest quantitatively how estimation error depends on sample size, length of observation period and collinearity.

Limitations

We feel that the semi-Markov interpretation holds many advantages for implementing marketing science consumer models. But to achieve these advantages we have had to make a number of tradeoffs in deriving our equations.

We have already discussed the limitations that the complexity of the process is limited by sample size and that the observation periods must be short enough to capture the dynamics of the system. In addition, we limit the process to either negative exponential or Erlang interpurchase distributions and require a priori specification of the order of the Erlang process. These are not major restrictions since the Erlang family is quite flexible and the order of the Erlang often can be identified empirically through a priori testing or iterative estimation.

A practical limitation is the segmented macro-flow approximation to the fully disaggregate process. Segmented macro-flow is not an assumption of homogeneity, but neither is it a fully disaggregate process. Instead it is an approximation which is made necessary by the need to group consumers together to obtain $\tilde{P}_{ns}$ for the regression approximation. This assumption can be relaxed when full maximum-likelihood estimation software is available.

Another limitation is the stationary Markov assumption between periods. In particular, between T_{n-1} and T_n, the explanatory variable matrix, X_{nlc} or R_{nls}, does not depend on the evolution of the process since T_{n-1}. For most variables this does not matter. However for accumulated variables, such as word of mouth, this requires that we use variables such as "cumulative trial at T_{n-1}" throughout the period T_{n-1} to T_n. Thus for accumulated variables we make the same "discrete" analog approximation that Bass [9] and his colleagues make in their diffusion models. To date, empirical evidence seems to support the discrete analog, but this assumption does represent an area of caution in using our methodology to model diffusion processes.

Future Work

As we develop empirical experience with the semi-Markov methodology, we will be able to identify priority areas for future research. Among these might be (1) relaxation of the macro-flow assumption by developing practical computer software for full maximum likelihood estimators of the w_l, (2) incorporation of nonlinear regression in (15), (3) simultaneous estimation of the w_l and the ρ_l, (4) investigation of the discrete analog for diffusion processes, (5) generalizations beyond negative exponential or Erlang purchase timing, and (6) modeling of heterogeneity with mixing distributions on the parameters of the continuous time Markov process. See discussions by Morrison [90], Morrison and Schmittlein [93], Singer and Spilerman [103] and Zufryden [111]. Some sample size considerations for heterogeneous processes are given in Kalwani and Morrison [64]. One might investigate the need for mixing distributions by simulating a heterogeneous process and analyzing it with segmented w_l's.

Appendix: Necessary and Sufficient Conditions for the Maximum Likelihood Estimators

In this appendix we deal with intermediate level likelihood function, L_3, since the other likelihood functions can be derived from L_3. From the text we have:

$$L_3 = \sum_n \sum_s \sum_{(i,j)} C_{ijns} \log\left\{\exp\left(\sum_l w_l R_{nls} t_n\right)\right\}_{ij}.$$

Since the R_{nls} are differential matrices, the maximum likelihood estimates, w_l^*, are those (unconstrained) values that maximize L_3.

Necessary Conditions (First Order Conditions)

For w_l^* to maximize L_3 it is necessary that $\partial L_3/\partial w_l = 0$ for $w_l = w_l^*$ for all l. Let W be the vector of the w_l's. Then since $p_{ijns}(t_n, W) = \{\exp(\sum_l w_l R_{nls} t_n)\}_{ij}$ we have:

$$\frac{\partial L_3}{\partial w_l} = \sum_n \sum_j \sum_{(i,j)} C_{ijns}\left(\frac{1}{p_{ijns}(t_n, W)}\right)\left\{\frac{d}{dw_l}\exp\left(\sum_l w_l R_{nls} t_n\right)\right\}_{ij}$$

$$= \sum_n \sum_s \sum_{(i,j)} C_{ijns}\left(\frac{1}{p_{ijns}(t_n, W)}\right)\{R_{nls} t_n \cdot P_{ns}(t_n, W)\}_{ij}$$

$$= \sum_n \sum_s \sum_{(i,j)} C_{ijns}\left(\frac{\sum_k r_{iknls} t_n p_{kjns}(t_n, W)}{p_{ijns}(t_n, W)}\right) = 0 \tag{27}$$

for all l as necessary conditions.

While (27) (for all l) can be solved by gradient search procedures, it is clear that the solution is difficult unless we have a good starting solution.

Approximate Solution (Starting Solution)

Suppose that $p_{ijns}(t_n, W) = C_{ijns}/C_{ins}$. Substituting in (27) we get:

$$\frac{\partial L_3}{\partial w_l} = \sum_n \sum_s \sum_{(i,j)} C_{ijns}\left(\frac{\sum_k r_{iknls} t_n C_{kjns}/C_{kns}}{C_{ijns}/C_{ins}}\right)$$

$$= \sum_n \sum_s \sum_i C_{ins}\left[\sum_n \left(r_{iknls} t_n \cdot \sum_j C_{kjns}/C_{kns}\right)\right]$$

$$= \sum_n \sum_s \sum_i C_{ins}\left[\sum_k r_{iknls}\right] t_n = 0 \quad \text{for any } l \tag{28}$$

where we have used $\sum_j C_{kjns} = C_{kns}$ to obtain the third expression and $\sum_k r_{iknls} = 0$ to obtain the result.

Thus if we can select W such that $p_{ijns}(t_n, W) = \tilde{p}_{ijns} \equiv C_{ijns}/C_{ins}$ we satisfy the necessary condition for a maximum. Since in general this requires $\mathbb{S} \cdot (S^0)^2 \cdot N$ equations for L unknowns we use the least squares solution derived in the text as a starting solution.

Sufficient Conditions (Second Order Conditions)

The sufficient conditions for W^* to be a maximum are that the matrix of second partial derivatives be negative definite. (Negative definite Hessian.) Taking the partial derivative of (27) with respect to w_m we obtain after suitable manipulation:

$$\frac{\partial^2 L_3}{\partial w_l \delta w_m} = \sum_n \sum_s \sum_{(i,j)} C_{ijns}\left(\frac{1}{p_{ijns}(t_n, W)}\right)^2$$

$$\cdot \left\{\sum_k \sum_h r_{iklns} p_{hjns}(t_n, W)\left[r_{khmns} p_{ijns}(t_n, W) - r_{ikmns} p_{kjns}(t_n, W)\right]\right\}. \tag{29}$$

The matrix of $\{\partial^2 L_3/\partial w_l \delta w_m\}$ must then be negative definite at W^* for W^* to be a maximum.

For example, for two behavioral states, two control variables, one segment and one time period, the Hessian, evaluated at $p_{ij}(t_n, W) = \tilde{p}_{ij}$, is given by (suppress the n and the s subscripts and the arguments, t_n, W):

$$\text{Hessian} = K\begin{bmatrix} \dfrac{C_1 r_{121}^2}{k_{12}} + \dfrac{C_2 r_{211}^2}{k_{21}} & \dfrac{C_1 r_{121} r_{122}}{k_{12}} + \dfrac{C_2 r_{211} r_{212}}{k_{21}} \\ \dfrac{C_1 r_{121} r_{122}}{k_{12}} + \dfrac{C_2 r_{211} r_{212}}{k_{21}} & \dfrac{C_1 r_{122}^2}{k_{12}} + \dfrac{C_2 r_{212}^2}{k_{21}} \end{bmatrix}$$

where $K = 2(p_{12} + p_{21}) - (p_{12} + p_{21})^2 - 1$ and $k_{ij} = p_{ij}(1 - p_{ij})$. The proof is by expansion of (29). Calculus shows that $K \leqslant 0$ for all $p_{12}, p_{21} \in [0, 1]$ and $K = 0$ iff P is singular. Furthermore it can be shown that the matrix multiplying K is positive definite for all $R_1 \not\equiv R_2$. (The proof is by completing the square to show that the diagonal elements and the determinant are positive.) Thus, for two behavioral states and two controls the Hessian is negative semi-definite at $p_{ij} = \tilde{p}_{ij}$ for all P, R_1, R_2 and negative definite for all nonsingular P. The extension to multiple segments, multiple time periods, multiple behavioral states, and multiple control variables represents future research. However, we have found that the Hessians in our simulations (§7) are negative definite at the regression solutions.

Gradient Search Procedure

Let W^h be the best solution after the hth iteration of the gradient search. Let W^0 be the starting solution obtained from the least squares approximation. Let G^h be the vector of $\partial L_3/\partial w_l$ evaluated at W^h. Then a gradient search procedure of step size Δ and stopping rule δ is: (1) Compute G^h for W^h. (2) Is $|G^h| < \delta$? If yes, go to step 3. If no $W^{h+1} = W^h + G^h\Delta$ and return to step 1. (3) Are the second order conditions approximately satisfied? If yes, stop. If no, select another starting solution and return to step 1.

Since W^0 approximates the necessary conditions and since the condition $p_{ijns}(t_n, W) \approx \tilde{p}_{ijns}$ is intuitive, we expect a gradient search procedure to converge rapidly. Even without the true maximum likelihood estimates, w_l^*, we expect the least squares approximation to be sufficient for many managerial needs, especially for large samples and relatively short observation periods.[10]

[10]The authors wish to thank Steve Graves, John Little, Don Morrison, Al Silk and Glen Urban for their suggestions and provocative questions that have improved the derivation and exposition of the methodology and the empirical test. In addition, Moshe Ben-Akiva, Phil Kotler, and Alice Tybout provided useful criticisms and suggestions on the empirical application.

The empirical work was supported in part by a grant from the U.S. Department of Transportation, Urban Mass Transportation Authority, Office of Policy and Research. Mary Martha Churchill, Nat Jasper and Bruce Spear at the U.S. D.O.T. have provided valuable input. Representatives of the Village of Schaumberg were involved in the application and support of the empirical test of this model. In particular, we wish to thank Ken Dallmeyer (DART manager), Ray Kessell (mayor), the Board of the Village of Schaumberg, and the staff at the Regional Transportation Authority in Chicago.

This work has benefited from presentations at Northwestern University, M.I.T., University of Chicago, and the ORSA/TIMS Joint National Meetings in Milwaukee, Wisc.

References

1. Aaker, D. A. and Jones, J. M., "Modeling Store Choice Behavior," *J. Marketing Res.*, Vol. 8 (February 1971), pp. 38–42.
2. Albright, S. C. and Winston, W., "Markov Models of Advertising and Pricing Decisions," *Operations Res.*, Vol. 27, No. 4 (July-August 1979), pp. 668–681.
3. Aleksandrov, A. D., Kolmogorov, A. N., and Larrent'ev, M. A., *Mathematics: Its Content, Methods, and Meaning*, Vol. 3, Translated by S. H. Gould, K. A. Hirsch, and T. Bartha, Amer. Math. Soc. and MIT Press, Cambridge, Mass., 1965.

4. AMSTUTZ, A. E., *Computer Simulation of Competitive Market Response*, MIT Press, Cambridge, Mass., 1967.
5. ASSMUS, G., "Newprod: The Design and Implementation of a New Product Model," *J. Marketing*, Vol. 39, (No. 1 January 1975), pp. 16–23.
6. ATHANS, M., AND FALB, P. L., *Optimal Control: An Introduction to the Theory and its Applications*, McGraw-Hill, New York, 1966.
7. BALACHANDRAN, V. AND DESHMUKH, S. D., "A Stochastic Model of Persuasive Communication," *Management Sci.*, Vol. 22, No. 8 (April 1976), pp. 829–840.
8. BARNETT, A. I., "More on a Market Share Theorem," *J. Marketing Res.*, Vol. 13, No. 1 (February 1976), pp. 104–109.
9. BASS, F. M., "A New Product Growth Model for Consumer Durables," *Management Sci.*, Vol. 15, No. 5 (January 1969), pp. 215–227.
10. ———, "The Theory of Stochastic Preference and Brand Switching," *J. Marketing Res.*, Vol. 11 (February 1974), pp. 1–20.
11. ———, "The Relationships Between Diffusion Rates, Experience Curves, and Demand Elasticities for Consumer Durable Technological Innovations," *J. Business*, Vol. 53, No. 3, Part 2 (July 1980), pp. 551–578.
12. ———, JEULAND, A. AND WRIGHT, G. P., "Equilibrium Stochastic Choice and Marketing Penetration Theories: Derivations and Comparisons," *Management Sci.*, Vol. 22, No. 10 (June 1976), pp. 1051–1063.
13. ———, PESSEMIER, E. A., AND LEHMANN, D. R., "An Experimental Study of Relationships Between Attitudes, Brand Preference, and Choice," *Behavioral Sci.*, Vol. 17, No. 6 (November 1972), pp. 532–541.
14. BELL, D. E., KEENEY, R. L. AND LITTLE, J. D. C., "A Market Share Theorem," *J. Marketing Res.*, Vol. 12 (May 1975), pp. 136–141.
15. BLATTBERG, R. AND GOLANTY, J., "Tracker: An Early Test Market Forecasting and Diagnostic Model for New Product Planning," *J. Marketing Res.*, Vol. 15, No. 2 (May 1978), pp. 192–202.
16. ——— AND SEN, S. K., "A Bayesian Technique to Discriminate Between Stochastic Models of Brand Choice," *Management Sci.*, Vol. 21 (February 1975), pp. 682–696.
17. ——— AND ———, "Market Segments and Stochastic Brand Choice Models," *J. Marketing Res.*, Vol. 13 (February 1976), pp. 34–45.
18. BROWN, G., "Brand Loyalty—Fact or Fiction," *Advertising Age*, Vol. 23 (1952), June 9, pp. 53–65, June 30, pp. 45–47, October 6, pp. 82–86, December 1, pp. 76–79.
19. CARMAN, J. M., "Brand Switching and Linear Learning Models," *J. Advertising Res.*, Vol. 6 (June 1966), pp. 23–31.
20. CHATFIELD, C. AND GOODHARDT, F. "Results Concerning Brand Choice," *J. Marketing Res.*, Vol 12 (February 1975), pp. 110–113.
21. COHEN, C. AND ROBINSON, B., "Simulation Marketing and Programming with SPURT/76, A Simulation Package for University Research and Teaching," Vogelback Computing Center, Northwestern University, September, 1976.
22. COX, D. R. AND MILLER, H. D., *The Theory of Stochastic Processes*, Wiley, New York, 1965.
23. CUNNINGHAM, R. M., "Brand Loyalty—What, Where, How Much?," *Harvard Business Rev.*, Vol 24 (January-February 1956), pp. 116–128.
24. ———, "Customer Loyalty to Store and Brand," *Harvard Business Rev.*, Vol. 39 (November-December 1961), pp. 127–137.
25. DODDS, W., "Application of the Bass Model in Long-Term New Product Forecasting," *J. Marketing Res.*, Vol. 10 (August 1973), pp. 308–311.
26. DODSON, J. A., "An Empirical Examination of Buyer Behavior: Individual and Brand Analyses," unpublished doctoral dissertation, Purdue University, West Lafayette, Ind., 1975.
27. ——— AND MULLER, E., "Models of New Product Diffusion Through Advertising and Word of Mouth," *Management Sci.*, Vol. 15 (November 1978), pp. 1568–1578.
28. DOLAN, R. J. AND JEULAND, A. P., "The Experience Curve Concept: Implications for Optimal Pricing Strategies," working paper, Graduate School of Business, Unviersity of Chicago (revised February 1979).
29. DRAKE, A., *Fundamentals of Applied Probability Theory*, McGraw-Hill, New York, 1967.
30. EHRENBERG, A. S. C., "The Pattern of Consumer Purchases," *Appl. Statist.*, Vol. 8 (March 1959), pp. 26–41.
31. ———, "An Appraisal of Markov Brand Switching Models," *J. Marketing Res.*, Vol. 2 (November 1965), pp. 347–362.
32. ———, *Repeat Buying*, North-Holland, Amsterdam, 1972, p. 266.
33. ESKIN, G. J., "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model," *J. Marketing Res.*, Vol 10, No. 2 (May 1973), pp. 115–129.

34. Ezzati, A., "Forecasting Market Shares of Alternative Home-Heating Units by Markov Processes Using Transition Probabilities Estimated from Aggregate Time Series Data," *Management Sci.*, Vol. 21, No. 4 (December 1974), pp. 462–473.
35. Fourt, L. A. and Woodlock, J. W., "Early Prediction of Market Success for Grocery Products," *J. Marketing*, Vol. 25, No. 2 (October 1960), pp. 31–38.
36. Frank, R. E., "Brand Choice as a Probability Process," *J. Business*, Vol. 35 (January 1962), pp. 43–56.
37. Givon, M. and Horsky, D., "Market Share Models as Approximators of Aggregated Heterogeneous Brand Choice Behavior," *Management Sci.*, Vol. 24, No. 13 (September 1978), pp. 1404–1416.
38. Green, P. E., *Analyzing Multivariate Data*, Dryden Press, Hinsdale, Ill., 1978.
39. Haines, C., "A Theory of Market Behavior after Innovation," *Management Sci.*, Vol. 10 (July 1964), pp. 634–658.
40. Harary, F. and Lipstein, G., "The Dynamics of Brand Loyalty: A Markov Approach," *Operations Res.*, Vol 10 (January-February 1962), pp. 19–46.
41. Hauser, J. R., "Consumer Preference Axioms: Behavioral Postulates for Describing and Predicting Stochastic Choice," *Management Sci.*, Vol. 23, No. 13 (September 1978), pp. 1331–1341.
42. ——— and Urban, G. L., "A Normative Methodology for Modeling Consumer Response to Innovation," *Operations Res.*, Vol. 25, No. 4 (July-August 1977), pp. 579–619.
43. ——— and Wisniewski, K. J., "Application, Predictive Test, and Strategy Implications of a Dynamic Model of Consumer Response to Marketing Strategies," M.I.T. Working Paper, Cambridge, Mass., June 1981. Forthcoming in *Marketing Science*.
44. Herniter, J., "A Probabilistic Market Model of Purchase Timing and Brand Selection," *Management Sci.*, Vol. 18, No. 4, Part II (December 1971), pp. P-102–P-113.
45. ———, "An Entropy Model of Brand Purchase Behavior," *J. Marketing Res.*, Vol. 10 (November 1973), pp. 361–375.
46. ———, "A Comparison of the Entropy Model and the Hendry Model," *J. Marketing Res.*, Vol. 11 (February 1974), pp. 21–29.
47. ——— and Cook, V. J., "Microsimulation Evaluation of Advertising Budget Strategies," working paper, Boston University, 1969.
48. ——— and Magee, J., "Customer Behavior as a Markov Process," *Operations Res.*, Vol. 9 (January-February 1961), pp. 105–122.
49. Hildebrand, F. B., *Advanced Calculus for Applications*, Prentice-Hall, Englewood Cliffs, N.J., 1962, pp. 514–516.
50. Horsky, D., "An Empirical Analysis of Optimal Advertising Policy," *Management Sci.*, Vol. 23, No. 10 (June 1977).
51. ——— and Simon, L. S., "A New Product Growth Model Incorporating the Effects of Advertising and Cross-City Differences," presented at the Joint National ORSA/TIMS Meeting, New York City, May 1–3, 1978.
52. Howard, R. A., "Stochastic Process Models of Consumer Behavior," *J. Advertising Res.*, Vol. 3 (September 1963), pp. 35–42.
53. ———, "System Analysis of Semi-Markov Processes," *IEEE Trans. Military Electronics*, Vol. MIL-8, No. 2 (April 1964), pp. 114–124.
54. ———, "Dynamic Inference," *Operations Res.*, (September-October 1965), pp. 712–33.
55. ———, *Dynamic Probabilistic Systems: Vol. 1, Markov Models, Vol. 2, Semi-Markov and Decision Processes*, Wiley, New York, 1971.
56. *International Mathematics and Statistics Library*, Vol. 1, 2, 3, IMSL LIB-0007, IMSL, Inc., Houston, Tex., January, 1979.
57. Jeuland, A. P., Bass, F. M., and Wright, G. P., "A Multibrand Stochastic Model Compounding Heterogeneous Erlang Timing and Multinomial Choice Processes," *Operations Res.*, Vol. 28, No. 2 (March-April 1980), pp. 255–277.
58. Jones, J. M., "A Comparison of Three Models of Brand Choice," *J. Marketing Res.*, Vol. 7 (November 1970), pp. 466–473.
59. ———, "A Dual-Effects Model of Brand Choice," *J. Marketing Res.*, Vol. 7 (November 1979), pp. 458–464.
60. ———, "A Composite Heterogeneous Model for Brand Choice Behavior," *Management Sci.*, Vol. 19, No. 5 (January 1973), pp. 499–509.
61. ——— and Zufryden, F. S., "Adding Explanator Variables to a Consumer Purchase Behavior Model: An Exploratory Study," *J. Marketing Res.*, Vol. 17, No. 3 (August 1980), pp. 323–334.
62. Jöreskog, K. G., "A General Method for Estimating a Linear Structural Equation System," in *Structural Equation Models in the Social Sciences*, A. S. Goldberger and O. P. Duncan, eds., Seminar Press, New York, 1973, pp. 85–112.

63. KALWANI, M. U. AND MORRISON, D. G., "A Parsimonious Description of the Hendry System," *Management Sci.*, Vol. 23, No. 5 (January 1977), pp. 467–477.
64. ——— AND ———, "Sample Size Requirements for Zero-Order Models," *J. Marketing Res.*, Vol. 17, No. 3 (May 1980), pp. 221–7.
65. KAO, E. P., "Modeling the Movement of Coronary Patients within a Hospital by Semi-Markov Processes," *Operations Res.*, Vol. 22, No. 4 (July-August 1974), pp. 683–699.
66. KOPPELMAN, F. S., "Methodology for Analysis of Errors in Prediction with Disaggregate Choice Models," *Transportation Res. Record*, No. 610, Transportation Research Board, Washington, D.C., 1976.
67. KUEHN, A. A., "Consumer Brand Choice as a Learning Process," *J. Advertising Res.*, Vol. 2 (December 1962), pp. 10–17.
68. ——— AND DAY, R. L., "Strategy of Product Quality," *Harvard Business Rev.*, (November-December 1962), pp. 100–110.
69. LANGE, E., "Measuring Market Response to Marketing Mix Variables Using Dynamic Modeling and its Implications for Brand Stretegy," unpublished Master of Science Thesis, M.I.T., Cambridge, Mass., June 1981.
70. LAVIDGE, R. J. AND STEINER, G. A. "A Model for Predictive Measurements of Advertising Effectiveness," *J. Marketing*, Vol. 25, No. 6 (October 1961), pp. 59–62.
71. LERMAN, S. R., "The Use of Disaggregate Choice Models in Semi-Markov Models of Trip Chaining Behavior," *Transportation Sci.*, Vol. 13, No. 4 (November 1979), pp. 273–291.
72. LILIEN, G. L., "A Modified Linear Learning Model of Buyer Behavior," *Management Sci.*, Vol. 20 (March 1974), pp. 1027–1036.
73. ———, "Application of a Modified Linear Learning Model of Buyer Behavior," *J. Marketing Res.*, Vol. 11 (August 1974), pp. 279–285.
74. LITTLE, J. D. C., "Aggregate Advertising Models: The State of the Art," *Operations Res.*, Vol. 27, No. 4 (July-August 1979), pp. 629–667.
75. MACLACHLIN, D. L., "A Model of Intermediate Market Response" *J. Marketing Res.*, Vol. 9, No. 4 (November 1972), pp. 378–384.
76. MAFFEI, R. B., "Brand Preferences and Simple Markov Processes," *Operations Res.*, Vol. 8 (March-April 1960), pp. 210–218.
77. MAHAJAN, V. AND MULLER, E., "Innovation, Diffusion, and New Product Growth Models in Marketing," *J. Marketing*, Vol. 43, No. 4 (Fall 1979), pp. 55–68.
78. ——— AND PETERSON, R. A., "Innovation Diffusion in a Dynamic Potential Adopter Population," *Management Sci.*, Vol. 24, No. 5 (November 1978), pp. 1589–1597.
79. MASSY, W. F., "Order and Homogeneity of Family Specific Brand-Switching Process," *J. Marketing Res.*, Vol. 3 (February 1966), pp. 48–54.
80. ——— AND MORRISON, D., "Comments on Ehrenberg: Appraisal of Brand Switching Models," *J. Marketing Res.*, Vol. 5 (May 1968), pp. 225–227.
81. ———, MONTGOMERY, D. B. AND MORRISON, D. G., *Stochastic Models of Buying Behavior*, MIT Press, Cambridge, Mass., 1970.
82. MCFADDEN, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, P. Zarembka, ed., Academic Press, New York, 1973, pp. 105–142.
83. ———, "Econometric Models for Probabilistic Choice Among Products," *J. Business*. Vol. 53, No. 3, Part 2 (July 1980), pp. 513–536.
84. MIDGLEY, D. F., "A Simple Mathematical Theory of Innovative Behavior," *J. Consumer Res.*, Vol. 3 (June 1976), pp. 31–41.
85. ———, *Innovation and New Product Marketing*, Croom Helm, London, 1977.
86. MONTGOMERY, D. B., "A Stochastic Response Model with Application to Brand Choice," *Management Sci.*, Vol. 15, No. 7 (March 1969), pp. 323–337.
87. ——— AND URBAN, G. L., *Management Science in Marketing*, Prentice-Hall, Englewood Cliffs, N.J., 1969.
88. MORRISON, D. G., *Multivariate Statistical Methods*, McGraw-Hill, New York, 1976.
89. ———, "Testing Brand-Switching Models," *J. Marketing Res.*, Vol. 3 (November 1966), pp. 401–409.
90. ———, "On Linearly Increasing Mean Residual Lifetimes," *J. Appl. Probability*, Vol. 15 (1978), pp. 617–620.
91. ———, "Purchase Intentions and Purchase Behavior," *J. Marketing*, Vol. 43, No. 2 (Spring 1979), pp. 65–74.
92. ——— AND BROCKWAY, G., "A Modified Beta Binomial Model with Applications to Multiple Choice and Taste Tests," *Psychometrika*, Vol. 44, No. 4 (December 1979), pp. 427–442.
93. ——— AND SCHMITTLEIN, D. C., "Jobs, Strikes, and Wars: Probability Models for Duration," *Organizational Behavior and Human Performance*, Vol. 25 (1980), pp. 224–251.

94. NEVERS, J. V., "Extensions of a New Product Growth Model," *Sloan Management Rev.*, Vol. 13, No. 2 (Winter 1972), pp. 77–90.
95. NOBLE, B., *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, N.J., 1969.
96. RAO, T. R., "Consumer's Purchase Decision Process: Stochastic Models," *J. Marketing Res.*, Vol. 6 (August 1969), pp. 321–329.
97. ROGERS, E. M., *The Diffusion of Innovation*, The Free Press, New York, 1962.
98. ——— AND SHOEMAKER, F. F., *Communications of Innovations: A Cross-Cultural Approach*, The Free Press, New York, 1971.
99. SHACHTMAN, R. H. AND HOGUE, C. J., "Markov Chain Model for Events following Induced Abortion," *Operations Res.*, Vol. 24, No. 5 (September-October 1976), pp. 916–932.
100. SHOCKER, A. D. AND SRINIVASAN, V., "Multiattribute Approaches to Product Concept Evaluation and Generation: A Critical Review," *J. Marketing Res.*, Vol. 16 (May 1979), pp. 159–180.
101. SILK, A. J. AND URBAN, G. L., "Pre-Test Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," *J. Marketing Res.*, Vol. 15, No. 2 (May 1978), pp. 171–191.
102. SINGER, B. AND SPILERMAN, S., "The Representation of Social Processes by Markov Models," *Amer. J. Sociology*, Vol. 82 (July 1976), pp. 1–54.
103. ——— AND ———, "Trace Inequalities for Mixtures of Markov Chains," *Advances in Appl. Probability*, Vol. 9, No. 4 (December 1977), pp. 747–764.
104. STYAN, G. P. H. AND SMITH, H. JR., "Markov Chains Applied to Marketing," *J. Marketing Res.*, Vol. 1 (February 1964), pp. 50–55.
105. TELSER, L. G., "The Demand for Brand Goods as Estimated from Consumer Panel Data," *Rev. Econom. Statist.*, Vol. 44 (August 1962), pp. 300–324.
106. ———, "Advertising and Cigarettes," *J. Political Economy*, (October 1962), pp. 471–499.
107. THEIL, H., *Principals of Econometrics*, Wiley, New York, 1971.
108. URBAN, G. L., "SPRINTER mod III: A Model for the Analysis of New Frequency Purchased Consumer Products," *Operations Res.*, Vol. 18, No. 5 (September-October 1970), pp. 805–853.
109. WACHSLER, R. A., PRINGLE, L. G. AND BRODY, E. I., "NEWS: A Systematic Method for Diagnosing New Product Marketing Plans and Developing Actionable Recommendations," Batten, Barton, Durstine, and Osburn, Inc., working paper, New York, February, 1972.
110. ZUFRYDEN, F. S., "A Dynamic Model of Market Response to Influences of Explanatory Variables," working paper, University of Southern California, in *Research in Marketing*, J. Sheth, ed. (to appear).
111. ———, "A Composite Heterogeneous Model of Brand Choice and Purchase Timing Behavior," *Management Sci.*, Vol. 24, No. 2 (October 1977), pp. 121–136.