INTRODUCTION STRATEGY FOR NEW PRODUCTS WITH POSITIVE AND NEGATIVE WORD-OF-MOUTH*

VIJAY MAHAJAN, EITAN MULLER AND ROGER A. KERIN

Edwin L. Cox School of Business, Southern Methodist University, Dallas, Texas 75275

Business School, Hebrew University of Jerusalem, Jerusalem 91905, Israel

Edwin L. Cox School of Business, Southern Methodist University, Dallas, Texas 75275

Existing innovation diffusion models assume that individual experience with the product is always communicated positively through word-of-mouth. For certain innovations, however, this assumption is tenuous since communicators of the product experience may transfer favorable, unfavorable, or indifferent messages through word-of-mouth. This paper examines a diffusion model for products in which negative information plays a dominant role, discusses its implications for optimal advertising timing policy and presents an application to forecast attendance for the movie Gandhi in the Dallas area.

(MARKETING; NEW PRODUCTS; DIFFUSION MODELS)

Introduction

The rate of new product introductions by industry is expected to double and the profit contribution by new products is anticipated to increase by 40 percent over the next five years (Booz, Allen and Hamilton 1982). The obvious importance of new product successes to the welfare of firms has prompted increased attention to new product forecasting models in marketing (see, for example, Urban and Hauser 1980; Wind, Mahajan and Cardozo 1981).

The theoretical foundation for a number of new product forecasting models lies in diffusion theory (see Bass 1969, Mahajan and Muller 1979). Diffusion theory, in brief, suggests that there is a time lag in the adoption of new products by different members of a social system. A new product is first adopted by some who, in turn, influence others to adopt it. It is the “interaction” or inter-personal communication (word-of-mouth) between early adopters and late adopters that is often posited to account for the rapid growth stage in the diffusion process (Rogers 1983).

Existing new product diffusion models implicitly assume that an individual's experience with a product is communicated positively through word-of-mouth (Mahajan and Muller 1979). This assumption is tenuous since communicators of the product experience may transfer favorable, unfavorable, or indifferent messages to others. Moreover, all adopters do not participate equally in the word-of-mouth process (Midgley 1976; Robertson 1971). Some actively spread favorable word-of-mouth and can be thought of as adding a free flow of information which augments the flow of information generated through advertising. Others might not participate in this social process even though they are aware of a new product or even had a satisfactory or unsatisfactory experience with it. Others still might actively spread unfavorable word-of-mouth due to negative impressions about a product or an unsatisfactory use experience (Richins 1983). Such negative information may consist of:

1. Brand related information such as the rumor that McDonald's hamburgers contained worms to boost the protein content (Wall Street Journal 1978);
2. Product category related information such as the relationship between toxic shock syndrome and tampons (Rotbart 1981);
3. Firm related information such as the rumor that Procter and Gamble was run by Satanists (Yao 1982); and
4. Industry related information such as the health and environmental effects of fluorocarbons in aerosol cans (Margolies 1976).

Empirical studies on the effect of negative word-of-mouth on product evaluation and adoption are limited although some research exists on how to deal with it from an information processing perspective (Tybout, Calder, and Sternthal 1981). Research in consumer behavior indicates that unfavorable information about products tends to carry greater weight with prospective buyers than favorable information (Mizerski 1982; Weinberger and Dillon 1980; Wright 1974). In a study on the diffusion of a new food product, Arndt (1967) reported that persons receiving negative word-of-mouth comments about the product used were 24

*Accepted by John R. Hauser; received March 30, 1982. This paper has been with the authors 12 months for 2 revisions.
percentage points less likely to purchase the product than other individuals. By comparison, persons receiving positive word-of-mouth comments were only 12 percentage points more likely to purchase the product. It is apparent from these studies that negative word-of-mouth can have a harmful effect on the adoption of new products, or at the very least, retard the diffusion process.

This paper is concerned with the diffusion and adoption of new products where negative word-of-mouth might play a significant role. Such products include books, movies, and novelty products. Other examples include reformulated or repackaged brands where the product class itself has encountered considerable negative word-of-mouth such as a new brand of tampons or new packages for nonprescription drugs (e.g., Tylenol).

The objectives of this paper are three-fold. First, we present a model showing how negative word-of-mouth can be incorporated into the diffusion process. Second, we describe a pilot study using the model to predict the viewing behavior of the movie Gandhi when it was introduced to the Dallas, Texas area. Third, based on our model we propose an optimal advertising timing policy for the introduction of products under three conditions: (1) pure negative word-of-mouth, (2) positive and negative word-of-mouth, and (3) pure positive word-of-mouth.

The Model

Since the original work by Bass (1969), a number of models have been suggested to study the dynamics of the diffusion process. As compared to the detailed adoption models (e.g., Urban's SPRINTER 1970) which model the flow of consumers through the various adoption states (e.g., about 500 in SPRINTER Mod III), the Bass model and its extensions (see Mahajan and Muller 1979) have been concerned with modeling the flow of consumers from unaware to trial states. Whereas the adoption models capture richness and reality, the diffusion models embrace simplicity and parsimony. In recent years, however, these two streams of model building seem to be converging. For example, most SPRINTER applications have tended to simplify the number of states (e.g. mod I SPRINTER). On the other hand, some extensions of the Bass model have tended to increase the number of states (e.g., Dodson and Muller 1978; Hauser and Wisniewski 1982; Midgley 1976).

The model presented in this paper is an attempt toward further convergence of these two research streams. In modeling the flow of consumers from unaware to trial states, the proposed model details the adoption process in terms of the nature of the information being communicated thus maintaining a linkage with the diffusion models. The model we propose is an extension of the model proposed by Dodson and Muller (1978), which by itself is an extension of the model proposed by Bass (1969). The market is divided into three main categories—unawares, potential customers and current customers, where in each of the last two categories further distinction is made to allow for a negative or positive word-of-mouth. Consider the following terminology, Figure 1, and Table 1.

\[ N = \text{the number of people in the market (assumed constant for the diffusion period under consideration),} \]

![Customer Flow Diagram for Products in Which Both Positive and Negative Types of Information Are Circulated.](image-url)
TABLE 1

Specification of the Flow Diagram

<table>
<thead>
<tr>
<th>FLOW</th>
<th>THE MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Unaware to Positive Potential</td>
</tr>
<tr>
<td>Advertising + w-o-m</td>
<td>$ux + (k_1 y^+ + k_3 z^+) x$</td>
</tr>
<tr>
<td>B</td>
<td>Unaware to Negative Potential</td>
</tr>
<tr>
<td>- w-o-m</td>
<td>$(k_2 y^- + k_4 z^-) x$</td>
</tr>
<tr>
<td>C</td>
<td>Positive Potential to Negative Potential</td>
</tr>
<tr>
<td>- w-o-m</td>
<td>$(k_2 y^- + k_4 z^-) y^+$</td>
</tr>
<tr>
<td>D</td>
<td>Positive Potential to Positive Triers</td>
</tr>
<tr>
<td>+ trial</td>
<td>$a^+ y^+$</td>
</tr>
<tr>
<td>E</td>
<td>Positive Potential to Negative Triers</td>
</tr>
<tr>
<td>- trial</td>
<td>$a^- y^+$</td>
</tr>
<tr>
<td>F</td>
<td>Negative Potential to Positive Potential</td>
</tr>
<tr>
<td>+ w-o-m</td>
<td>$(k_1 y^+ + k_3 z^+) y^-$</td>
</tr>
<tr>
<td>G</td>
<td>Positive Triers to Positive Triers</td>
</tr>
<tr>
<td>Repurchase</td>
<td>$g z^+$</td>
</tr>
<tr>
<td>H</td>
<td>Aware to Unaware</td>
</tr>
<tr>
<td>Forgetting</td>
<td>$b_1 y^+ + b_2 y^- + b_3 z^+ + b_4 z^-$</td>
</tr>
<tr>
<td>I</td>
<td>Positive Triers to Negative Triers</td>
</tr>
<tr>
<td>- w-o-m</td>
<td>$k_4 z^+ z^-$</td>
</tr>
<tr>
<td>J</td>
<td>Negative Triers to Positive Triers</td>
</tr>
<tr>
<td>+ w-o-m</td>
<td>$k_3 z^+ z^-$</td>
</tr>
</tbody>
</table>

$x(t) =$ the number of people who are unaware of the existence of the product,
$y(t) =$ the number of potential customers who are aware of the product but have not yet purchased it, where $y^+(t)$ are those of $y(t)$ who spread favorable (positive) information about the product and $y^-(t)$ are those who spread negative information,
$z(t) =$ the number of triers or current customers who are aware of the product and have purchased it, where $z^+(t)$ are those of $z(t)$ who spread positive information about the product and $z^-(t)$ are those who spread negative information,
$u(t) =$ advertising reach; i.e., the proportion of the market being informed at time $t$ by advertising,
$a =$ product trial rate where $a^+$ and $a^-$ are the trial rate with positive and negative experiences, respectively,
$b =$ forgetting rate where $b_1$, $b_2$, $b_3$ and $b_4$ are the forgetting rates of $y^+$, $y^-$, $z^+$ and $z^-$, respectively,
$g =$ product repurchase rate,
$k =$ contact rate, where $k_1$, $k_2$, $k_3$ and $k_4$ are the contact rates of $y^+$, $y^-$, $z^+$, and $z^-$, respectively.

Figure 1 depicts the customers flow for the product and Table 1 summarizes the ten different types of flows.

Consider flow A in Table 1. The difference in the number of unaware customers between time $t$ and time $t + \Delta t$, i.e., $x(t + \Delta t) - x(t)$ can be found according to the following considerations: Since $u(t)$ is the advertising reach coefficient, then $\Delta tu(t) N$ people are reached by the advertising during the time interval $\Delta t$. Out of these, only a fraction of $x(t)/N$ are unaware at the time. Thus the number of newly informed people at this period is $x(t)u(t)\Delta t$. During this time period, the positive potential people contact and inform a total of $k^1(y^+)\Delta t$ people out of which only $x(t)/N$ are unaware. Thus defining $k_1 = k^1/N$, the total number of unaware people who are
informed by the positive potentials is \( k_1 y^+ x \Delta t \). In the same manner, the total number of unaware people who are informed by the positive triers is \( k_3 z^+ x \Delta t \). Dividing through by \( \Delta t \) and taking the limit as \( \Delta t \) approaches zero one obtains the first two terms of equation (1), i.e., flow A. This procedure can be applied to yield the rest of equations (1) through (5).

This approach is useful to explain two possible pitfalls in the process. First, a question might be asked as to what happens if an unaware person is contacted in the same period by a positive and a negative trier or by advertising (which is assumed to be positive) and a negative trier. The answer is that these contacts are of second degree importance and can be safely ignored. To see this recall that the total number of unawares who are newly informed by advertising during \( \Delta t \) is \( x(t)u(t)\Delta t \). During this period, the negative triers contact and inform a total of \( k_4 z^-(t)\Delta t \) people out of which only a fraction of \( x(t)u(t)\Delta t / N \) are newly informed by advertising. Thus the total number of people informed both ways is \( k_4 z^-(t)x(t)u(t)(\Delta t)^2 \). When dividing by \( \Delta t \) and taking the limit as \( \Delta t \) approaches zero, this term approaches zero as well.

Second, a question might be raised about the validity of using the same contact coefficient when for example the positive triers contact an unaware and when they contact a negative potential. The answer is simply that the \( z^+ \) group (positive triers) contact and inform a total of \( k_3 z^+ x(t) \Delta t \), out of which \( x(t)/N \) are unaware, and \( y^-(t)/N \) are negative potentials. Thus in flow A the term \( k_3 z^+ x \) is added while in flow F the term \( k_3 z^+ y^- \) is added.

Flows other than word-of-mouth activities in Table I (which are explained much the same way) are D, E, G, H, I and J. In flows D and E, out of the total number of positive potentials \( y^+(t) \), a fraction of \( a^- y^+(t) \Delta t \) try the product during the period \( \Delta t \) and are satisfied, and \( a^- y^+(t) \Delta t \) try the product and are dissatisfied. In the same manner G and H are explained as fractions who repurchase and fractions who forget. Flows I and J represent the effect of negative triers on positive triers and vice versa, respectively.

The equations describing the change in the diffusion variables are given by:

\[
\frac{dx}{dt} = -ux - (k_1 y^+ + k_3 z^+)x - (k_2 y^- + k_4 z^-)x + b_1 y^+ + b_2 y^- + b_3 z^+ + b_4 z^- , \tag{1}
\]

\[
\frac{dy^+}{dt} = u x + (k_1 y^+ + k_3 z^+)x + (k_1 y^+ + k_3 z^+)y^- - (k_2 y^- + k_4 z^-)y^+ - (a^+ + a^- + b_1)y^+ , \tag{2}
\]

\[
\frac{dy^-}{dt} = (k_2 y^- + k_4 z^-)(x + y^+) - (k_1 y^+ + k_3 z^+)y^- - b_2 y^- , \tag{3}
\]

\[
\frac{dz^+}{dt} = a^+ y^+ + (k_3 - k_4)z^+ z^- - b_3 z^+ , \tag{4}
\]

\[
\frac{dz^-}{dt} = a^- y^+ - (k_3 - k_4)z^+ z^- - b_4 z^- . \tag{5}
\]

By definition:

\[
x(t) + y^+(t) + y^-(t) + z^+(t) + z^-(t) = N \tag{6}
\]

and the sales rate \( s(t) \) is given by

\[
s(t) = (a^+ + a^-)y^+ + gz^+ . \tag{7}
\]

The transition equations (1) through (5) are a concise summary of Table 1 where each inflow adds to the respective stock and thus has a positive sign and each outflow has a negative sign. It should be noted that since \( N \) is constant, the summation of the first five equations is zero. Thus only four equations are independent.
There are several assumptions which are implicitly made in the above formalization of the diffusion and adoption process. First, we have ignored the existence of neutral information, which is neither positive or negative. Truly neutral information is hard to convey since conveying a neutral experience with a product will, in most cases, be interpreted as unfavorable. Still, if one is able to convey information which is so impotent as to leave the receiver of the information with exactly the same attitude about the product as he had before, then we can consider these types as conveying no information.

Second, we have ignored the possibility that there is a segment of the population which, though it may purchase the product, does not participate in the word-of-mouth process. If we may borrow a term from epidemiology, this group is called an immuned group. Its size clearly has a negative effect on the size of the relevant word-of-mouth flows. Third, we ignore impulse purchase.

An Application

The applicability of the proposed model was examined in the context of a pilot study on motion picture attendance. Motion pictures were chosen for study because word-of-mouth has been identified as the “most potent and least manageable factor in movie marketing” by industry observers (Lees and Berkowitz 1981) and shown to affect a prospective viewer’s opinion of movies (Handel 1976; Burzynski and Baker 1977; Mizerski 1982).

Subjects in this study consisted of 67 undergraduate students enrolled in an undergraduate Marketing Fundamentals course at Southern Methodist University. Undergraduate college students were considered appropriate as subjects since the Motion Picture Association of America estimates that 76 percent of first-run theatrical film viewers are between the ages of 12 and 24 (Broadcasting 1982). Subjects were pre-screened before being included in the study. Only students who responded yes to the question “Would you typically see two or more movies per month?” were asked to participate. Subjects were rewarded for participation with a drawing for cash prizes.

Subjects agreed to maintain a diary of their movie-related experiences on a weekly basis over a ten-week period beginning the week of January 17 and concluding the week of March 28, 1983. On Tuesday of each week, subjects submitted a self-administered questionnaire. The questionnaire asked participants to report (1) the titles of movies they had become aware of in the preceding week, (2) the primary source of information about each movie mentioned—friends, movie reviews, advertising, other—and whether the information led them to hold a favorable, neutral, or unfavorable opinion about the movie, (3) whether or not the movie was seen during the week, (4) whether or not the movies seen would be recommended to others, and (5) whether or not they were involved in any movie-related conversations during the week.

The movie Gandhi was selected for analysis because advance publicity indicated that this movie would be introduced to the Dallas area during the first week of the study period. Gandhi was shown at a movie theater approximately two miles from the campus.

Table 2 indicates the reported incidence of movie-related word-of-mouth activity, the principal source of information about Gandhi, and whether the information was a positive, negative, or neutral influence on the participant’s opinion of Gandhi. This summary shows that approximately three in five subjects engaged in movie-related conversations during each week of the study. In addition, friends were the major source of movie-related information (except for the first time period reports). It is apparent that Gandhi benefited from favorable information conveyed through friends, movie reviews, and advertising. However, a sizeable number of subjects reported that
### TABLE 2

<table>
<thead>
<tr>
<th>Percent Reporting Talking About Movies*</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Source and Direction**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information Source:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends' Opinion</td>
<td>36%</td>
<td>37%</td>
<td>38%</td>
<td>37%</td>
<td>38%</td>
<td>42%</td>
<td>54%</td>
<td>46%</td>
<td>48%</td>
<td>48%</td>
</tr>
<tr>
<td>Movie Review</td>
<td>18</td>
<td>32</td>
<td>34</td>
<td>34</td>
<td>26</td>
<td>27</td>
<td>13</td>
<td>24</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td>Advertising</td>
<td>46</td>
<td>31</td>
<td>28</td>
<td>29</td>
<td>36</td>
<td>31</td>
<td>33</td>
<td>30</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>Information Direction:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>60%</td>
<td>55%</td>
<td>63%</td>
<td>62%</td>
<td>67%</td>
<td>68%</td>
<td>68%</td>
<td>70%</td>
<td>75%</td>
<td>74%</td>
</tr>
<tr>
<td>Negative</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>12</td>
<td>12</td>
<td>8</td>
<td>11</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Neutral</td>
<td>30</td>
<td>40</td>
<td>37</td>
<td>31</td>
<td>21</td>
<td>20</td>
<td>24</td>
<td>19</td>
<td>16</td>
<td>19</td>
</tr>
</tbody>
</table>

*Question: During the last 7 days have you talked with anyone about some movie?**Question: For only those movies you have not seen, indicate what single factor or source most influenced your opinion on the movie. Indicate with a +, 0, − whether the factor was a positive, neutral, or negative influence on your opinion of the movie.

the information they received led them to hold a neutral and negative view of the movie.

In order to forecast the movie attendance, the proposed model, equations (1)–(5), was simplified to reflect the actual reports of subjects. Specifically, no subjects reported a negative trial of Gandhi (i.e., \( a^- = 0 \)) and no subjects reported that they would not recommend the movie to others after viewing it (i.e., \( z^- = 0 \)). Assuming no forgetting during the time frame over which the data were gathered, the model was simplified to:

\[
\frac{dx}{dt} = -ux - (k_1y^+ + k_3z^+)x - k_2y^-x, \tag{1a}
\]

\[
\frac{dy^+}{dt} = ux + (k_1y^+ + k_3z^+)x + (k_1y^+ + k_3z^+)y^- - k_2y^-y^+ - a^+y^+, \tag{2a}
\]

\[
\frac{dy^-}{dt} = k_2y^- (x + y^+) - (k_1y^+ + k_3z^+)y^- , \tag{3a}
\]

\[
\frac{dz^+}{dt} = a^+y^+. \tag{4a}
\]

Note that the reduced model contains five parameters, \( u, k_1, k_2, k_3, \) and \( a^+ \), and given time-series data on \( x, y^+, y^-, \) and \( z^+ \), these parameters can be estimated by using simultaneous least squares procedures (Theil 1971). Alternatively, making some distributional assumptions about the errors associated with the equations in the model, maximum likelihood estimation procedures can be developed to estimate the parameters (see, for example, Hauser and Wisniewski 1982). However, we adopted a simple approach to estimate the parameters. Since the equation (2a), associated with the positive-aware group, contained all the parameters, a discrete analog of this equation...
was used to develop the required estimates by using ordinary least squares procedures. Equation (2a) was further simplified to a regression analog with five coefficients:

\[ \Delta y^+ = u_x + k_1y^+x + k_3(z^+x + z^+y^-) + (k_1 - k_2)y^+y^- - a^+y^+. \]  

(2b)

Equation (2b) was applied to the Gandhi data for the first eight weeks of data holding the last two data points for prediction validation. The model provided a reasonably good fit with an adjusted \( R^2 \)-value of 0.85. Nevertheless, given five coefficients and eight observations, it was felt that a more parsimonious model was appropriate. Accordingly, the data were re-analyzed using the model: \( \Delta y^+ = u_x + k_3(z^+x + z^+y^-) - a^+y^+ \). This model assumes that \( k_1 = k_2 \approx 0 \). This assumption was based on the view that interaction between positive and negative awares tended to cancel each other while the effect of positive and negative word-of-mouth from tryers would dominate. Support for this view, though limited, is found in two studies on movie-related behavior. For example, in a field study, Burzynski and Baker (1977) observed that manipulating positive and negative comments from movie viewers as they exited the theater had an effect on those waiting in line to see the movie. Specifically, they reported that “some people (assumed to be positive awares since they were waiting in line to view the movie) exposed to negative comments immediately redeemed their tickets.” In a laboratory setting, Mizerski (1982) found that when students were given an evaluation of a movie by a fellow student, who had previewed a movie, this evaluation had a significant effect on the recipient's opinion of the movie.

Results using the parsimonious model proved to be a good predictor of change in the positive aware group (adjusted \( R^2 = 0.83, p = 0.02 \)). The parameter estimates and the standard errors, given in the parentheses, for the parsimonious model were: 

\( u = 0.11 (0.04), k_3 = 0.02 (0.02), \) and \( a^+ = 0.12 (0.11) \). Actual and estimated cumulative awareness and cumulative attendance for Gandhi for the eight estimation periods are shown in Figure 2. The actual and predicted cumulative awareness for the last two weeks of the study were 94 percent/95 percent and 94 percent/95 percent, respectively. Actual and predicted cumulative attendance for the last two weeks were 28 percent/33 percent and 30 percent/39 percent, respectively.

These empirical results are encouraging for the estimation and prediction of awareness. However, predicted values were systematically diverging from actual values for movie attendance. An explanation for this finding lies in the fact that a major examination was scheduled during the ninth week of the study and two popular

![Figure 2. Actual and Estimated Awareness and Attendance for the Movie Gandhi.](image-url)
motion pictures were introduced in nearby theaters during the last two weeks (prediction periods) of the study. These movies were *Spring Break*, a movie directed at college age viewers, and *High Road to China* starring Tom Selleck. Since the model does not explicitly account for external interventions (e.g., examinations and other movie releases), these factors likely depressed actual attendance values and accounted for the divergence of actual and predicted attendance.

**Optimal Advertising/Timing Policies**

In order to highlight the idiosyncrasies of negative word-of-mouth, as it relates to advertising and timing policies, we start by considering a case of pure negative word-of-mouth, then proceed to include positive word-of-mouth, and conclude by considering the pure positive word-of-mouth model.

**Case 1: Pure Negative Word-of-Mouth.** Consider the following scenario: the marketer/producer of the product (e.g., a movie) believes that the product is such that no consumer will adopt it twice. Moreover, any adopter will spread only negative information, i.e., will tell his friends to avoid the product and they concur. In addition, once informed, the time framework is so short that no forgetting occurs and there is no repeat purchase. This is clearly an extreme example (although not unheard of) but it helps to isolate the effects of negative information.

From the above scenario it is clear that the equations are reduced to the following system:

\[
\frac{dx}{dt} = -ux \ \text{positive} - (k_2y^+ + k_4z^-)x,
\]

\[
\frac{dy^+}{dt} = ux \ \text{positive} - a^-y^+ - (k_2y^- + k_4z^-)y^+,
\]

\[
\frac{dy^-}{dt} = (k_2y^- + k_4z^-)y^+ + (k_2y^- + k_4z^-)x,
\]

\[
\frac{dz^-}{dt} = a^-y^+.
\]

Note that all the word-of-mouth activity is negative. The only positive information flow is generated by advertising.

Though we model the maximization problem of a firm which produces a product which induces only negative word-of-mouth, the aim is clearly not to instruct the fly-by-night producer how to optimally market his product but rather to pinpoint the special problems negative information creates and how to deal with them.

In order to determine the policy of the firm, we have to make assumptions about its revenue/cost functions. Since the sales of the product are \(a^-y^+\) we let the gross profits of the firm (revenues net of all cost except advertising) be concave in the number of triers \(a^-y^+\). In addition we let the gross profits be negative if the adoption level of the product falls below a specified level of \(y^+\). This assumes the existence of fixed costs, such as the case, for example, with the revenues of a movie theater. If the occupancy level falls below a certain percent, losses are incurred. The fact that the gross profit function is concave allows for either linear or convex cost of production, and/or a downward sloping demand function. Following Arrow (1964) we implicitly assume that the firm has already chosen the optimal pricing path.

The firm can choose the time of the introduction of the product and withdrawal time, time of beginning and end of the advertising campaign and the advertising intensity. The optimal policy of the firm is as follows (for a proof, see Appendix A). The firm advertises at capacity for a period of time, i.e., set \(u(t) = \bar{u}\) where \(\bar{u}\) is the upper limit on advertising set forth by either managerial consideration or by practical
market consideration. It starts advertising (at $t_0$) before it introduces the product (at $t_1$). It ends the advertising campaign (at $t_2$) after the product introduction but before it withdraws the product (at $T$), as is shown in Figure 3.

The intuition behind this policy is as follows (again, for the formal proof the reader is referred to Appendix A). Since the producer/marketer knows that the word-of-mouth mechanism will be against the product and it does not start until consumers have actually adopted it, the producer does not introduce the product initially, but instead builds a stock of potential customers by advertising well before the product introduction. In that way only positive information is circulated. Once the desired level of potential customers has been achieved, the producer introduces the product. Word-of-mouth process takes time to be effective, since it is proportional to the number of people who have adopted the product (penetration level). Therefore for some time the producer enjoys a relatively stable adoption level (the level of the flow depends on the effectiveness of the producer’s advertising blitz). Gradually, the word-of-mouth process starts building up, and consequently the number of potential customers drops (as is depicted in Figure 4).

When the level of potential customers drops to a prespecified level of $y^+$ which depends on the profit schedule with relation to the adoption level) the producer withdraws the product since, at that level, its gross operating profits (revenues net of all costs except advertising) is zero. Naturally, since the marginal value of an additional potential customer at that time is zero, and since advertising is related to this marginal value, it is worthwhile to stop advertising before product withdrawal time ($T$).

Denote by $t_1$, the first time at which the level of positive potential $y^+$ has reached the threshold level of $y^+$. We prove (see Appendix B) that $t_1$ (the introduction time) is larger than $t_*$, i.e., the firm delays the introduction of the product until its stock of potential customers is larger than $y^+$ which is the level below which gross profits are negative. The intuitive explanation is as follows. Below the level $y^+$ gross profits (i.e., profits net of all costs except advertising) are negative. At $y^+$ profits are zero. Now consider $t_*$. If the firm introduces the product (or releases the movie) gross profits at $t_*$ will be zero. If the firm does not introduce the product (i.e., delays introduction) gross profits are still zero. The firm considers the negative word-of-mouth which will start if (and only if) it releases the product and thus delays introduction. Only at such time at which the gross operating profits will be large enough to offset the negative discounted losses due to the introduction will they introduce the product.

Denote by $t^*$ the time at which the positive potentials graph $y^+$ reaches its maximum. It is rather straightforward to show (see Appendix B) that at $t^*$ the introduction time, $y^+$ is still increasing while at $t_2$ the advertising termination time, $y^+$
is decreasing and thus \( y^* \) is in between these two points of time. The firm, therefore, terminates the advertising after sales have reached their peak.

**Case 2: Positive and Negative Word-of-Mouth.** This case will be dealt with less formally than the previous one partly since various assumptions can be made regarding the parameter configuration which will affect timing and advertising levels and partly because of the difficulty in solving analytically a control problem with four state equations. In this case we still adhere to the assumptions of the last case except that the positive trial rate \( a^+ \) is positive and thus the positive triers group \( z^+ \) is positive and adds a positive word-of-mouth information flow. Thus the negative effects are still large enough so that the repeat purchase is negligible. Addition of repeat rate will be considered in the next case. Adding the positive triers group does not change the main theme of the policy except for two points.

First, we have argued that \( t_1 \) (the product introduction time) is strictly positive by noting that as long as the number of positive potentials is below a certain level, profits are negative, which is inferior (for a given time) to having zero profits by delaying the introduction. This is so since at any future time the result of the early introduction is to lower profits since the negative word-of-mouth process (which is the only one) starts earlier. This is not the case when the positive triers add a positive flow of information. Then it is possible, if their contact rate is large enough, and if the group is large enough, to start the advertising simultaneously with the product introduction time. Although profits initially will be negative (as opposed to zero profits initially when delaying introduction), this will be compensated at future time by the increased profits due to the fact that the (relatively large) favorable word-of-mouth process started earlier. Thus the result, as shown in Figure 3, is still valid except that \( t_1 \) (product introduction time) might be zero in the extended case.

Second, we have argued that \( t_1 \) (the product introduction time) is strictly larger than \( \bar{t} \) (the threshold level time). In the extended case, it is possible that \( \bar{t} \) will be larger than \( t_1 \). The reason is that at \( \bar{t} \) the firm decided to delay the introduction because of the negative effect of the word-of-mouth. In this case, it will consider both negative and positive flows. If the positive effects are large enough the firm will introduce the product before \( \bar{t} \). Gross profits initially will be negative but this will be compensated by the future increased profits which are due to the earlier introduction and the subsequent larger word-of-mouth activity.

**Case 3: Pure Positive Word-of-Mouth.** In this case both the negative potential group and the negative trier group vanish. The repeat purchase model then collapses to the model proposed by Dodson and Muller (1978) (assuming identical contact and forgetting parameters and an additional switching parameter \( \theta \) from \( z \) to \( y \)).

Instead of solving the problem directly using control theory, a more indirect route was chosen. The problem was first converted into one that conformed to a "Most Rapid Approach Path" (MRAP) form, given by Spence and Starrett (1975) and then solved using their method. We first present a description of the MRAP idea.

Spence and Starrett (1975) studied a class of problems in which the control does not enter the objective function. First, they give a necessary and sufficient condition for a general autonomous, one state, one control, infinite horizon problem to be convertible into the above form. Then they characterize the solution to be as follows: the solution involves moving to a local maximum of the (converted) integrand as rapidly as possible, and then staying there forever. Intuitively, since the transformed objective function does not involve the control, there is no direct or indirect penalty in high level of the control, so that once the desired level of the state is identified (that depends on the initial conditions) a most rapid approach is clearly superior to any other approach.

The problem as posed in this section is not directly suitable to their approach but can be converted into a suitable one (see Appendix C). Dropping the positive sign,
assuming identical contact rates and forgetting parameters and an additional switching parameter \( \theta \) from \( z \) to \( y \), the model can be stated as

\[
\begin{align*}
\frac{dx}{dt} &= -ux - kx(N - x) + b(N - x), \\
\frac{dy}{dt} &= ux + kx(N - x) - (b + a)y + \theta z, \\
\frac{dz}{dt} &= ay - (\theta + b)z \quad \text{and} \quad s(t) = ay + gz.
\end{align*}
\]

Now, the problem is to maximize \( J = \int_0^\infty e^{-rt} (ps(t) - cNu) dt \) where \( r \) is the discount rate, \( p \) is the price of the product, and \( c \) is the cost of informing one person. As shown in Appendix C, the maximization problem can be converted to:

\[
J = \int_0^\infty e^{-rt} W(x(t)) dt \quad \text{where}
\]

\[
W(x) = B(r + b)(N - x) - c(N - x)[bN/x - k] + rcN \log(x/N)
\]

and

\[
B = \frac{pa}{r + b} \left[ 1 + \frac{(g - a)}{(r + b + \theta + a)} \right]. \tag{12}
\]

As explained in Appendix C, the solution approach now is to find one maximum of \( W(x) \) by differentiating \( W(x) \) with respect to \( x \) and equating to zero. The resultant quadratic equation in \( x \) gives the following solution:

\[
x^* = \frac{Nb}{(-r/2 + \left[ (r/2)^2 + b(r + b)B/c + bk \right]^{1/2}}. \tag{13}
\]

From equation (13), a straightforward comparison reveals that \( x^* < N \) if and only if

\[
(r + b)B/c + k - (r + b) > 0. \tag{14}
\]

The structure of the optimal policy is therefore clear. If (14) holds, then the optimal policy is to achieve the desired level of \( x^* \) as quickly as possible by setting \( u = \bar{u} \) until \( x^* \) is attained. However, if (14) does not hold, the optimal policy is to set \( u = 0 \); that is, not to enter the market.

It is possible to turn now to the interpretation of condition (14), which is a benefit-cost comparison. First note that in case \( k > r + b \), (14) holds and there is no need for a cost benefit comparison. Therefore, in what follows, it is assumed that \( k < (r + b) \).

Define \( \bar{C} \) to be:

\[
\bar{C} = c(1 - k/(r + b)). \tag{15}
\]

\( \bar{C} \) represents the effective cost of informing one person. This cost consists not only of the out-of-pocket cost of informing one person \( c \) but also takes into account that this one person will contact and inform others. Thus the word-of-mouth mechanism represents a free flow of information. Therefore, the effective cost of reaching a person is reduced by that amount of free flow \( k \), discounted to reflect not only the time preference \( r \) (since this flow is not instantaneous) but also the probability of forgetting —since this free flow will terminate when the person forgets.

In the same manner, it is possible to interpret \( B \) as it appears in equation (12). First note that if \( a = g \), \( B \) represents the earnings \( pa \), discounted at the rate of \( r + b \). In case \( a \neq g \), rewrite \( B \) as

\[
B = \frac{pa}{r + b} \left[ 1 + \frac{(g - a)}{(r + b + \theta + a)} \right].
\]

Thus these discounted earnings \( pa/(r + b) \) are reduced (or increased) by the difference between the repeat purchase rate and the trial rate. It is capitalized at a higher rate to take into account not only the time preference rate and the probability of forgetting, but also the probability of switching between the \( y \) group and the \( z \) group.
With the definition of $C$ and $B$ and their interpretation as effective cost and benefit given, it is now easy to see that (14) is indeed a cost benefit comparison since it can be written as

$$B/C > 1.$$  \hspace{1cm} (14')

Thus if $B/C < 1$, i.e., the effective cost of reaching a person is not less than the benefit expected from him, the firm had better not enter the market.

In case $B/C > 1$, i.e., the benefit is greater than the cost, the MRAP solution is optimal. The firm then tries to build the level of informed persons to $N - x^*$ as quickly as possible and then maintains that level forever. Denote that maintenance level by $u^*$, that is, with $x^*$ as defined in (13), $u^*$ has to satisfy $\dot{x} = 0$, i.e.,

$$u^*x^* + kx^*(N - x^*)/N = b(N - x^*).$$  \hspace{1cm} (16)

The left-hand side of (16) represents the people who are newly informed, either by direct mechanism or by the word-of-mouth mechanism. The right-hand side represents the people who forget. Thus the maintenance level $u^*$ equates the number of people who are newly informed to the number of people who forget. The optimal timing is depicted in Figure 5.

There are two more subcases which should be considered, now that $u^*$ is defined: first, if $u$ is infinite and second, if $u < u^*$. If $u$ is infinite, the mathematical optimal policy is to “jump” to the desired level $x^*$ instantaneously by making an “infinite injection” at $t = 0$. Then the policy is to maintain that level forever. If $u < u^*$, then the desired level of informed people cannot be reached and the optimal policy is to set $u = u$ for all times. This, however, does not, strictly speaking, meet the requirements of the MRAP since $x^*$ is not attainable. Nevertheless, it is easily checked that such a case can be added to the Spence and Starrett (1975) paper using similar arguments. It is of interest to note that the sales rate will be monotonically increasing if $a < g$ and will have an overshooting pattern if $a > g$. Thus, the patterns proven in Dodson and Muller (1978) when advertising is constant, can be proven here when the firm is advertising optimally (see Appendix C).

Summarizing the three cases, the implications are induced from the intuitive understanding of the optimal solution. There is a desired level of awareness to be reached, and it should be reached as quickly as possible. Now in the usual case, with repeat purchase, the policy calls for reaching the level as quickly as possible by advertising at capacity when releasing the product. Once the desired level is reached, maintenance advertising is called for, as it maintains the desired level at a constant state. In the case where negative information is dominant, there is a desired level of awareness to be reached as quickly as possible by advertising before the product release time. This helps delay the word-of-mouth flow and build quickly the aware group.
INTRODUCTION STRATEGY FOR NEW PRODUCTS

Before this desired level is reached, the product is released and the adverse word-of-mouth starts circulating. Once the desired level is reached, advertising is dropped and the awareness level drops. The existence and magnitude of the fixed costs determines whether and when the product is withdrawn.

Conclusions

In many product-marketing situations, the impact of product promotion and advertising efforts is enhanced by word-of-mouth effect—that is, by the recommendation of the product by current adopters to potential adopters. The importance of the word-of-mouth effect in the development of marketing strategies is well documented in marketing (Robertson 1971). In fact, the word-of-mouth effect has served as the primary underlying behavior rationale in the development of innovation diffusion models of new product acceptance (Mahajan and Muller 1979). Most existing innovation diffusion models, however, assume that individual experience with the product is always communicated positively through word-of-mouth. For certain innovations, this assumption is tenuous since communicators of the product experience may transfer favorable, unfavorable, or indifferent messages through word-of-mouth. To overcome this limitation, this paper examined a diffusion model explicitly incorporating negative word-of-mouth communication. An application was presented and optimal advertising timing policies for the introduction of a new product in the presence of negative word-of-mouth were delineated. Future extensions of the proposed model to include other marketing variables, market interventions, and competition will further contribute to the convergence between the diffusion and adoption models and thus enhance the practical and diagnostic power of the diffusion models.1

1 We would like to thank John Hauser and three anonymous referees for a number of helpful suggestions. This research was made possible by a generous grant from Procter and Gamble Company to the Marketing Department at the Edwin L. Cox School of Business.

Appendix A

To simplify notations, since the solution involves only three groups, $a^-$ will be denoted by $a$, and $y^+$ will be denoted by $y$. Without loss of generality we can assume that $t_0 = 0$ since because of the negative word-of-mouth it is clearly not optimal to introduce the product before advertising it. The firm chooses $t_1$, $t_2$, $T$ and $u(t)$ so as to maximize:

$$\int_0^{T} \left[ \frac{n}{cN} \sigma_N \right] e^{-\rho t} dt + \int_{t_1}^T \left[ \sigma(y) - cN \right] e^{-\rho t} dt$$

subject to the constraints

$$\frac{dx}{dt} = -ux, \quad \frac{dy}{dt} = ux \quad \text{for} \quad 0 < t < t_1,$$

$$x(0) = N \quad \text{and} \quad \frac{dx}{dt} = -ux - (k_d(N - x - y - y^-) + k_2 y^-) x,$$

$$\frac{dy}{dt} = ux - ay - (k_d(N - x - y - y^-) + k_2 y^-) y \quad \text{for} \quad t_1 < t < T,$$

$$\frac{dy^-}{dt} = (x + y)(k_d(N - x - y - y^-) + k_2 y^-),$$

$$x(t_1) = x_1, \quad y(t_1) = y_1, \quad y^-(t_1) = 0, \quad \text{and} \quad 0 < u(t) < \bar{u} \quad \text{for all} \quad 0 < t < T.$$

$\bar{u}$ is the upper limit on advertising set forth by either managerial consideration or by practical market consideration. $r$ is the discount rate; $c$ is the cost of reaching one person. $\pi$ is a concave profit function which satisfies that $\pi$ is increasing in $y$ and $\pi(y) = 0$ for some $y > 0$. The concavity of the gross profit function allows for either linear or convex cost of production, and/or a downward sloping demand function. Following Arrow (1964) we implicitly assume that the firm has already chosen the optimal pricing path.

Separating the problem into two time intervals depicts the fact that before product introduction time $t_1$, no word-of-mouth process starts and profits are zero. As soon as the product is introduced, negative word-of-mouth starts spreading.

PROPOSITION. The following policy satisfies the necessary conditions for optimality: $u(t) = \bar{u}$ for $0 < t < t_2$, and $u(t) = 0$ for $t_2 < t < T$ and $0 < t_1 < t_2 < T$. 

This content downloaded from 212.179.28.34 on Wed, 13 Aug 2014 07:21:37 UTC
All use subject to JSTOR Terms and Conditions
PROOF. Define two current value Hamiltonians for the two time intervals as follows:

\[ H_1 = -cN\nu - \mu_1ux + \lambda_1ux, \]
\[ H_2 = \pi - cN\nu - \mu_2ux - \mu_2x(k_4N - x - y - y^- + k_2y^-) \]
\[ + \lambda_2ux - \lambda_2ay - \lambda_2y(k_4N - x - y - y^- + k_2y^-) \]
\[ + \gamma(x + y)(k_4N - x - y - y^- + k_2y^-) \]

where \( \mu, \lambda, \) and \( \gamma \) are the current value multipliers associated with the state variable \( x, y \) and \( y^- \). In addition to the standard conditions for optimality, the multipliers and the Hamiltonians are continuous at \( t_I \) so that

\[ H_1(t_I) = H_2(t_I), \lambda_1(t_I) = \lambda_2(t_I) \text{ and } \mu_1(t_I) = \mu_2(t_I). \]

In addition, since \( T \) is chosen optimally, \( H_2(T) = 0 \). For details see Amit (1977, Chapter 3) which is an extension of Kamien and Schwartz (1981, §9). Since the Hamiltonians are linear in \( u \), we have for \( i = 1, 2 \):

\[ \frac{\partial H_i}{\partial u} = (\lambda_i - \mu_i)x - cN \]

either \( u = a \) and \( x(X; t) > cN \)

or \( u = 0 \) and \( x(X; t) < cN \)

or \( 0 < u < a \) and \( x(X; t) = cN \).

Following standard control theory (see, for example, Sethi and Thompson 1981) the multipliers satisfy:

\[ \frac{dy}{dt} = \gamma(x + y)(k_4N - x - y - y^- + k_2y^-) \]

First, it is clear that \( t_1 > 0 \) since as long as \( y(t) < y \) (initially \( y(0) = 0 \)), \( \pi < 0 \) if the product is introduced which is inferior at any time to \( \pi = 0 \) which is achieved by delaying the introduction. At any future time the result of the early introduction is even worse since the negative word-of-mouth process starts earlier.

During this time interval \( 0 < t < t_1 \) we have the following assuming it is worthwhile to enter at all, i.e., \( \lambda_1(0) - \mu_1(0) > c \), i.e., at zero the benefits from informing one person are larger than the cost of informing:

\[ d\lambda_1/\lambda/dx = \gamma(x + y)(k_4N - x - y - y^- + k_2y^-) \]

At \( t = T \), \( \partial H_2/\partial u = -cN \) since both \( \lambda(T) = \pi(T) = 0 \). Since \( H_2 \) is continuous and positive at \( t_I \), there must be a time, denoted by \( t_2 \), at which \( \partial H/\partial u = 0 \). Therefore, \( t_2 < T \). Lastly \( T \) is finite since during \( t_2 < t < T \), \( u = 0 \) and so \( dy/dt < 0 \). Since it approaches zero asymptotically, there exists a finite time, denoted by \( T \), at which \( y = \bar{y} \) and the operation terminates, since at that time \( H_2(T) = 0 \). Q.E.D.

Appendix B

Following Appendix A, we drop the superscripts from \( a^- \) and \( y^+ \). Let \( t_1, \tilde{t}, t^* \) and \( t_2 \) be the introduction time, first time at which \( y(t) = \bar{y} \), time at which \( \bar{y} \) is at maximum, and advertising termination time respectively.

PROPOSITION. The time points \( t_1, \tilde{t}, t^* \) and \( t_2 \) satisfy the following inequality: \( \tilde{t} < t_1 < t^* < t_2 \).

PROOF. Following Appendix A, the continuity of the Hamiltonian function at \( t_1 \) implies \( \pi(y(t_1)) = \lambda(t_1) ay(t_1) \). Since we assume that the value of an additional potential customer is positive at zero, i.e., \( \lambda(0) > 0 \) then \( \lambda(t) = \lambda(0)e^{\theta t} \) and so \( \lambda(t) > 0 \). Thus \( \pi(y(t_1)) > 0 \) and so \( y(t_1) > \bar{y} \). Since, for \( 0 < t < t_1, dy/dt > 0 \), then \( t_1 < \tilde{t} \). This also implies that \( t_1 < t^* \) since \( t^* \) is the time at which \( y \) reaches its maximum. At \( t = t_2, u = 0 \) and so \( dy/dt < 0 \). This implies that \( t^* < t_2 \) (we cannot prove strict inequality since \( dy/dt \) is discontinuous at \( t_2 \)). Q.E.D.

Appendix C

The problem is to maximize \( J = \int_0^\infty e^{-n} (ps(t) - cNu) \) dt where \( s(t) \) is the sales rate given by \( s(t) = ay + g2 \) and

\[ \frac{dx}{dt} = -ux - kx(N - x) + b(N - x), \]

\[ \frac{dy}{dt} = ux + kx(N - x) - (b + a)y + \theta z, \]

\[ \frac{dz}{dt} = ay - (\theta + b)z. \]

Substituting the expression for the sales into the objective function and substituting \( N - x - y \) for \( z \), we get:

\[ J = \int_0^\infty e^{-n} \{ pay + pg(N - x - y) - cNu \} dt. \]
Solving equation (1') algebraically for $u$ in terms of $x$ and $\dot{x}$ (and redefining $k$ to be $k/N$) yields:

$$u = b(N - x)/x - k(N - x)/N - \dot{x}/x.$$  

Substituting $N - x - y$ for $z$ in equation (2') and solving it algebraically for $y$ and substituting the last expression for $u$ yields:

$$y = [(b + \theta)(N - x) - \dot{x} - y]/(a + b + \theta).$$

Let $I = \int_0^\infty e^{-\gamma t}d\mu(t)$. Substituting for $y$ from the last equation and integrating by parts the last term, using the initial condition $y(0) = 0$, using the definition of $I$, and rearranging terms yields:

$$I = (1/(r + a + b + \theta)) \int_0^\infty e^{-\gamma t}[(b + \theta)(N - x) - \dot{x}]dt.$$  

Substituting $I$ and the expression for $u$ into the objective function $J$, and integrating by parts the terms containing $\dot{x}$, using the initial condition $x(0) = N$, yields:

$$J = \int_0^\infty e^{-\gamma t}W(x(t))dt$$

where $B$ is as defined in equation (12).

The objective function is now free of the control $u$. The problem, formally, is to maximize ($J$) subject to (1') and $0 < u(t) < u_0$.

This problem is time autonomous, infinite horizon, one state, one control, with the objective function free of the control. Thus it conforms to the MRAP formulation.

As mentioned before the solution is to approach as rapidly as possible a local maximum of $W(x)$ and then stay there forever. The problem now reduces to finding the extrema of $W(x)$. This is done by differentiating $W(x)$ and equating to zero. The result is a quadratic equation in $x$, where the one positive root is $x^*$ as given in equation (13).

Lastly, as for the sales rate, rewrite $s$ as: $s = a(N - x) + (g - a)z$. It is rather straightforward to check the $(x, u)$ and $(x, z)$ phase diagrams and to find that $x$ is decreasing and $z$ increasing, both monotonically. Thus, if $g > a$, the sales are monotonically increasing. Rewrite sales as: $s(t) = (a - g)y(t) + g(N - x(t))$. Differentiating, we get $\dot{s} = (a - g)\dot{y} - g\dot{x}$. Evaluating for $t > t_2$ yields: $\dot{s} = (a - g)\dot{y}$. It is clear that since in this period $\dot{y} < 0$, then $\dot{s} < 0$ if, and only if, $a < g$. Since $s(t)$ is a continuous function and $s(0) > 0$, then if $a > g$ an overshooting pattern must occur.

Q.E.D.

References


Broadcasting (1982), "Moving Those Movies" (March 29), 133.


