Internet Media Planning: An Optimization Model

Janghyuk Lee¹

and

Laoucine Kerbache²

HEC School of Management, Paris
1 rue de la Libération
78351 Jouy-en-Josas, France

¹ Assistant Professor of Marketing
email: lee@hec.fr
phone: +33-1 3967 9440

² Associate Professor of Logistic and Production
email: kerbache@hec.fr
phone: +33-1 3967 7212

Acknowledgements: We are gratefully indebted to Vijaya Chebolu for her research assistance and particularly appreciate financial support from "La Fondation HEC" and data from KoreanClick.
ABSTRACT

Of the various media vehicles available for advertising, the Internet is the latest and the most rapidly growing, emerging as the ideal medium to promote products and services in the global market. In this article, the authors propose an Internet media planning model whose main objective is to help advertisers determine the return they obtain from spending on Internet advertising. Using available data such as Internet page view and advertising performance data, the model contributes to attempts not only to optimize the Internet advertising schedule but also to fix the right price for Internet advertisements on the basis of the characteristics of the exposure distribution of sites. The authors test the model with data provided by KoreanClick, a Korean market research company that specializes in Internet audience measurement. The optimal durations for the subject sites provide some useful insights. The findings contrast with current Web media planning practices, and the authors demonstrate the potential savings that could be achieved if their approach were applied.

Key words: media planning; optimization, advertising repeat exposure, probability distribution.
1. Introduction

For businesses that sell goods and services, advertising often represents the first means to make
the public aware of them. Among the various media vehicles available for advertising, the Internet is
the latest and most rapidly growing (Bell and Tang 1998). Already a major communication channel in
many developed countries, the Internet has attracted the attention of marketing managers not only
because of its rapid adoption but also because it is an interactive communication medium that provides
a wide variety of size, location, and technology options (Novak and Hoffman 1997). Especially the
rapid evolution of data transmission speed on the Internet makes consumers spend more time on the
Internet compared to other traditional mass media. According to a recent report from Jupiter Research
(2004), the broadband Internet is challenging TV in Europe as 40% of consumers having broadband
access at home said that they were spending less time watching TV. Therefore, companies
increasingly have started to rely on Internet advertising to acquire new customers and improve their
brand image. According to an Internet advertising revenue report from the Interactive Advertising
Bureau (2004), U.S. companies increased their spending on Internet advertising from $907 million in
1997 to $7,267 million in 2003, with a 21% growth rate between 2002 and 2003. Although this
percentage increase includes 3% of total media spending, the study finds a high concentration of
advertising spending on major Web sites. In the fourth quarter of 2003, the top 10 Web sites
accounted for 71% of total advertising spending, and the top 50 Web sites encompassed 96%.

Internet advertising offers more accurate measurement and more flexible planning than do
traditional media (Drèze and Zufryden 2000). For example, each site can measure systematically the
size of its audience and the frequency of exposure. This improved accuracy enhances the transparency
of a return on investment (ROI) analysis because the direct impact of an Internet advertisement on
sales can be assessed and even linked to a commercial Web site. The Internet also allows for content
modifications and schedule flexibility. If, for example, an Internet advertising campaign was
unsuccessful in its early stages, its content and schedule could be modified for the rest of the
campaign. For this overall approach, advertisers need a decision-making tool such as media planning.
Media planning determines the subject, timing, and location of advertisements. Decisions regarding the establishment of a media plan involve understanding and then integrating marketing objectives, the dynamics of the market, target audience, and the available media vehicles with their associated costs and characteristics. Because these data often are incomplete and uncertain, media planning problems tend to be probabilistic in nature. However, in contrast to traditional media publicity vehicles, such as magazines, newspapers, and television, the Internet provides much more data that enable advertisers to measure exactly the exposure of consumers to the displayed advertisements. In turn, Internet data make it possible to understand the impact of publicity on the consumer and increase its effectiveness.

To measure the impact of advertising, researchers need information about the number of consumer exposures and consumers’ attention levels to a particular advertisement. The past decade has witnessed the widespread adoption of media models for estimating the reach and frequency distribution of exposures of traditional media. Thus, several methods, both public and proprietary, are now available to media planners to use to estimate the proportion of the target audience that will be exposed once, twice, and up to n times by a combination of n insertions in m media vehicles (Little and Lodish 1969, Aaker 1975, Rust 1985).

Considering the innovation and technological benefits that the Internet has over traditional media, there is an undeniable need to adapt previous models of media planning effectively to the Internet environment, in which a visitor can be exposed to an advertisement many times during a fixed period. In this article, we propose an Internet media planning model to deal with this issue, through which we focus specifically on Web page view statistics and the repeat exposure effect. The main objective of this model is to help advertisers assess the ROI of spending on Internet advertising through the use of available market data such as Internet page views and advertising performance data.

We organize this article as follows: In section 2, we present the theoretical background related to exposure distribution and the repeat exposure effect. Section 3 is devoted to developing our Internet media planning model. In section 4, we present a real case study and analyze the results by exploring various scenarios that yield interesting managerial insights and demonstrate the robustness of our
2. Theoretical Background

The main concern of advertisers is how many people they can reach and how often. Advertising agencies attempt to achieve optimal planning about the number of placements and the choice of media to maximize the “reach/frequency” of a campaign with a given budget if other things (e.g., attractiveness, creativity) are equal. In the following, we address the major conceptual issues central to our research: exposure distribution and the repeat exposure effect.

2.1. Exposure Distribution

In traditional media such as magazines and television, the question of how to capture the exposure distribution of people across various media was explored first by Metheringham (1964), who used a binomial distribution to capture the reach and frequency of exposure for a single vehicle (i.e., magazine) with a fixed number of insertions and then developed the beta binomial distribution (BBD) to integrate varying probabilities of exposure that represent heterogeneity across consumers.

A more flexible model that treats more than one vehicle simultaneously is based on Dirichlet multinomial distribution (DMD) (Leckenby and Kishi 1984, Danaher 1988a). Rust and Leone (1984) extend the DMD with a hypergeometric adjustment that accommodates the case of an unequal number of insertions in all vehicles. Similarly, Danaher (1988b) develops a log-linear model to handle three or more vehicles at a time, which outperforms BBD and DMD in error measurements. Despite its enhanced performance, its computational burden remains a problem, which Danaher (1989) eased by using an approximation of the minimum deterioration of performance.

However, on the Web, the nature of advertising exposure changes drastically because of consumers’ unlimited access to advertising, in contrast with their passive, limited exposure in the case of predetermined advertising schedules such as those used on television. Therefore, it is no longer valid to apply distributions such as BBD and DMD that fix the total number of insertions; each insertion on a Web page can generate unlimited exposures. Because Web advertisements are posted
on a specific page for a fixed duration, the focus of exposure distribution shifts from estimating the number of exposures according to the number of insertions to determining the number of exposures an insertion generates during a fixed period of time.

The number of exposures to a Web page during a fixed period is a stochastic process that follows a Poisson distribution for average exposures. Because each consumer may have a different level of average exposure, exposure frequencies can be fit by the negative binomial distribution (NBD) (Ehrenberg 1959), a mixture of the Poisson and gamma distributions. The Poisson distribution represents the exposure rate for a fixed duration, whereas the gamma distribution captures the heterogeneity of the exposure rate. Because the number of future purchases a priori depends on whether a customer is active, Schmittlein, Morrison, and Colombo (1987) refine the NBD by introducing the probability of being active as represented by the Pareto distribution that mixes the exponential and gamma distributions. The “death” rate of the customer is captured by the exponential distribution, whose heterogeneity is embedded in the gamma distribution. In this research, we use the NBD because we assume that all users stay alive for the relatively short (e.g., four weeks) duration of an advertising campaign. This assumption significantly reduces the computational complexity of the media planning optimization model.

2.2. Repeat Exposure Effect of Advertising

Advertisements can influence the consumer’s three-stage (generation, consideration, selection) brand choice process. They also can alter the content of brand information on two dimensions—accessibility and value—stored in memory, including the brand name, the value and valence of brand-attribute beliefs, and the valence of brand attitudes (Nedungadi, Mitchell, and Berger 1993). By increasing the accessibility of product-attribute beliefs and brand attitudes (Berger and Mitchell 1989) and activating brand information, advertising repetition can enhance the performance of advertisements.

The two-factor theory proposed by Berlyne (1970) suggests that the impact of exposure frequency is mediated by two factors: habituation (learning) and tedium. Habituation can improve an advertisement’s performance, whereas tedium deteriorates it. If the tedium factor overwhelms its
counterpart after the number of exposures passes a threshold, repeat exposures may take the form of inverted-U curves, in which two opposing psychological processes operate simultaneously: positive habituation and negative tedium. Similar explanations of an inverted-U curve function for repeated exposure have been proposed for attitudes (Cacioppo and Petty 1979) and learning (Pechmann and Stewart 1989). Pechmann and Stewart (1989) use the terms “wearin” and “wearout” in their elucidation of the inverted-U curve response to advertising and suggest that wearin occurs during approximately the first three exposures, after which positive thoughts outnumber negative thoughts. The wearout stage begins with approximately the fourth exposure, when message recipients start to become bored and consequently generate negative repetition-related thoughts, which undermine the persuasive impact of the advertisement.

Krugman (1972) provides a different perspective on the effects of frequency. He proposes a three-hit theory, which posits that an advertisement reaches its maximum effectiveness at the third exposure. The first exposure elicits a cognate response to the nature of the stimulus. The second exposure is more evaluative and personal and raises questions about the meaning of the ad. But the third exposure represents the true reminder because the viewer has already gone through his or her cognitive process. Krugman further argues that additional exposures simply repeat the third-exposure effect without incremental improvements. Thus, the three-hit theory could be graphically depicted as an S-shaped or concave response curve with a plateau after the third exposure. Some previous research also supports the claim that attitudes, purchase intentions, and positive cognitive responses peak at the third exposure in the case of television advertising (Cacioppo and Petty 1979, Calder and Sternthal 1980, Belch 1982).

For the Internet, Drèze and Hussherr (2003) find positive repeat exposure effects for three major measures: aided brand awareness, unaided advertising recall, and brand awareness. They test the repeat exposure effects for a sample of 807 respondents who were surveyed both before and after their exposure to Internet advertising of 10 brands. The number of exposures ranged 0–9 times during 24 hours. They detect a statistically insignificant forgetting phenomenon.
2.3. Previous Approaches to Media Planning

In the 1960s, linear programming emerged as an appropriate modeling and optimization tool for allocating advertising to various media. Miller and Starr (1960) and Day (1962) established the criteria to apply linear programming principles for selecting media through questions about when (time) and where (space) advertisements should appear according to the budgetary constraints. Additional constraints that guarantee a minimum spending level of a media class or an individual medium can be taken into account, as can the minimum exposure of specific market segments. These models rely on the rationale that advertising creates an advertising exposure that in turn creates sales; that is, the purchase intention of consumers can be elevated by their enhanced awareness and positive attitude toward the brand. Along these lines, Lee (1962) developed a linear programming model to optimize the number of advertisements of a given time length to ensure the required level of awareness. Starsch (1965) extended the model to select markets in which the advertisement should appear and integrated frequent disparities between the sales potential, which is specific to each market, and the circulation of candidate media that serve those markets. Brown and Warshaw (1965) introduced the notion of nonlinear response to advertising, in which the response prompts diminishing, S-shaped returns in an exponential form with a saturation level (Vidale and Wolfe 1957). According to this theory, the number of advertisements used per period can be modeled as a decision variable, fractioned into regions that have different response levels.

To refine and extend previous models, Little and Lodish (1969) developed MEDIAC, in which they include the market segments, segment-specific sales potentials, and exposure probabilities of each media option. The decision variables are boolean variables that indicate the insertion of a given advertisement in a specific medium at a specific time. On the basis of the insertion variable, the authors can assess the total market response as a function of the level of current and previous exposures of consumers and their sales amount, weighted by segment. Subsequently, Aaker (1975) proposed a media planning model with a different approach. His ADMOD model focuses not on the aggregate vehicle audience but rather on sample populations selected from the various segments and thereby examines the likely impact—measured as a change in cognition or purchase intention—of a particular
insertion on each consumer in the sample. The change in cognition is assessed as a function of the number of exposures, depending on the media schedule, which then is extrapolated to the real population to provide the total expected results (e.g., profit generated by the media schedule). A binomial distribution captures the distribution of exposures because the ADMOD model includes a limited number of ad insertions. Finally, Rust (1985) suggested a television planning model (VIDEAC) that provides standard data such as program availability, cost, and rating (by segment) directly to advertising agencies. In VIDEAC, the exposure distribution is estimated by the BBD to capture the heterogeneity of the population exposure rate.

3. Model Development

Compared with those of previous media planning models, which were developed mainly for television, radio, and magazine advertising and have a discrete format with a limited number of insertions, the decision variables of our model represent the duration of advertising (weeks) on selected Web sites. The goal of our model is to assess the repeat exposure effect of advertising on the Web, where most sites attract visitors repeatedly. For the sake of modeling simplicity, we start by considering continuous decision variables and do not limit the number of insertions a priori. However, we examine other scenarios subsequently.

In our model, the objective function directly measures the number of individuals (i.e., consumers recalling the ad message) who can be influenced by advertising. It consists of two parts: the repetition function and the exposure frequency distribution. In addition, it assesses performance at the level of exposure frequency. For ad message recall, the objective function sums the total number of subjects who recall the ad message, obtained from the probability of message recall after being exposed k times and the number of visitors exposed k times. Thus, the function maximizes the number of influenced consumers by choosing the duration of sites according to the exposure disparity in the number of unique visitors and their repeat exposure distribution across selected sites.

The repetition function can be obtained from pretest results that calibrate the probability of ad performance in terms of exposure frequency. In previous media planning models, the repetition function seems to show diminishing returns at high exposure levels. Little and Lodish (1969) adopt a
function in an exponential form that describes the fraction of sales potential realized as a function of exposure level, in which there is a minimum fraction for no ad exposure and an upper bound of 1 for achieving the full potential of sales. In contrast, ADMOD (Aaker 1975) uses a repetition function with lower and upper bounds that are linked through a power function of exposure frequency. The VIDEAC model (Rust 1985) adopts a simple form of the square root of the number of exposures for the repetition function.

In our model, we use the repetition function to represent the probability of ad performance (message recall rate) and assume it to be a logit function of the exposure frequency,

\[ p(X = k) = \frac{1}{1 + \exp[-(a + bk)]}. \]

Its lower bound depends on the value of the constant a for the case of no exposure, and it increases monotonically. The shape of this repetition function depends mainly on the coefficient of exposure frequency b. The greater the coefficient b, the greater is the performance difference in the low range of exposures because the slope of the repetition function gets steeper. As the number of exposures tends toward infinity, the probability approaches 1.

For exposure frequency probability, we first describe it by a Poisson distribution with the mean exposure rate of \( \lambda \),

\[ f(X = k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}. \]

To estimate the exposure distribution flexibly over a varying time duration, the exposure frequency probability changes to

\[ f(X(t) = k|\lambda) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, \]

which includes the duration variable t to represent extended or shrunken duration. Because it is realistic to incorporate the heterogeneity of the exposure rate \( \lambda \), which varies across individuals following a distribution of \( g(\lambda) \),

\[ f(X(t) = k|\lambda) = \int g(\lambda) d\lambda, \]

which we incorporate by using the gamma distribution. The exposure frequency distribution can be estimated by a mixture distribution of Poisson and gamma that leads to a NBD with two parameters, \( \gamma \) as the shape parameter and \( \alpha \) as the scale parameter, in addition to the duration variable t, so that

\[ f(X(t) = k) = \sum_{\gamma} \frac{(\lambda t)^k e^{-\lambda t}}{k!} \frac{\Gamma(\gamma + k)}{\Gamma(\gamma) \Gamma(\alpha + t)} \left( \frac{\alpha}{\alpha + t} \right)^k. \]

The number of consumers exposed \( k \) times can be obtained by multiplying the total population (M) by the exposure frequency distribution.
Thus, our objective function to assess the number of individuals influenced by the advertising becomes a product of the probability of ad performance and the number of individuals exposed \( k \) times, \( \sum_{k=1}^{K} p(X = k) * M * f(Y = k) \), for a single site at which the decision variable is the ad duration \( t \).

In case of \( N \) different sites, the objective function becomes \( \sum_{i=1}^{N} \sum_{k=1}^{K} p(X = k) * M * f(Y = k) \), and we find a set of the duration for selected sites \( (t_i) \) that maximizes the objective function. However, the repetition function and frequency exposure distribution may be specific to each Web site that has its own parameters.

Along with the preceding objective function, our Internet media planning model includes the following set of constraints:

- The amount of budget allocated to an ad campaign. On the Internet, a frequently applied method to fix ad rates is based on the total frequency of exposures. Among ad practitioners, this rate is called the “cost per thousand exposures” (CPM).\(^1\) In our model, this rate is noted \( r_i \) because it may be specific to each site \( i \). Therefore, the expression of the ad campaign cost, obtained from the cost rate and the total number of exposures of the listed sites, \( \sum_{i=1}^{N} \frac{r_i}{1000} \sum_{k=1}^{K} M * k * f(Y = k) \leq A \), should be smaller than the campaign budget \( (A) \).

- The maximum duration of selected site(s). Even for cost-attractive sites, an advertiser cannot run an ad campaign for more than a certain period of time. This duration constraint often is related to the timing of the ad campaign. In our model, we easily incorporate it as \( t_i \leq T \ \forall i \).

Our model can summarized as follows: The objective function,

\[
\sum_{i=1}^{N} \sum_{k=1}^{K} p_i(X = k) * M * \Gamma(g_i + k) \frac{\alpha_i}{\alpha_i + t_i} t_i^{g_i} \left( \frac{t_i}{\alpha_i + t_i} \right)^k,
\]

is subject to the budget constraint,

\[
\sum_{i=1}^{N} \frac{r_i}{1000} \sum_{k=1}^{K} M * k * f(Y = k) \leq A.
\]

\(^1\) The Interactive Advertising Bureau (2004) reports that 47% of ad revenues were generated on the basis of CPM or impressions (including sponsorship) in 2003 for the U.S. market.
\[
\sum_{i=1}^{n} \frac{r_i}{1000} \sum_{k=1}^{K} M \cdot k \cdot \frac{\Gamma(\gamma_i + k)}{\Gamma(\gamma_i)k!} \left( \frac{\alpha_i}{\alpha_i + t_i} \right)^{\gamma_i} \left( \frac{t_i}{\alpha_i + t_i} \right)^{x_i} \leq A,
\]

and the time duration constraint,
\[
t_i \leq T \ \forall i,
\]
where the decision variable is \(t_i = \) advertising duration of site \(i\), and the other parameters are as follows:

- \(p_i(X = k) = \) ad performance of site \(i\) at \(k\) exposures,
- \(\gamma_i = \) shape parameter of the NBD capturing the exposure frequency of site \(i\),
- \(\alpha_i = \) scale parameter of the NBD capturing the exposure frequency of site \(i\),
- \(r_i = \) advertising fee rate (CPM), and
- \(A = \) total advertising budget amount.

Our is clearly a nonlinear programming optimization model with continuous variables and a complex objective function. The objective function is nonlinear, and therefore, the search for optimal or nondominated solutions will be computationally time consuming, especially for large problems.

Because of the particular characteristics of the Internet, advertisers cannot use previous media planning models to choose the optimal combination of slots from a predetermined schedule. Our model instead enables advertisers to plan an ad campaign in a continuous manner to reduce or increase the duration of advertising on the listed sites to maximize their objectives, as represented by the objective function. Also, our model fully assesses the marginal contribution of ad duration for each site and captures the ad effect across the range of exposure frequency. Finally, this model can be viewed as a platform that may be modified to incorporate other key issues of media planning, such as segmentation and the interaction effect.

To incorporate segments, we can split the exposure frequency distribution into \(S\) segments. Each segment has its own parameters, \(\gamma\) and \(\alpha\), of the NBD distribution, which enables us to assess the exposure frequency with different patterns (shape and scale),

\[
f_{\gamma_i}(X(t_i) = k) = \frac{\Gamma(\gamma_i + k)}{\Gamma(\gamma_i)k!} \left( \frac{\alpha_i}{\alpha_i + t_i} \right)^{\gamma_i} \left( \frac{t_i}{\alpha_i + t_i} \right)^{x_i}.
\]

This process is exactly the same as that used to
estimate the exposure frequency of sites, except that we split it by segment and add a weight \( w_s \) to represent the size of the segment, \( \sum_{i=1}^{N} \sum_{s=1}^{S_i} \sum_{k=1}^{K} p(X = k) \cdot M \cdot w_s \cdot f_s(X(t_r) = k) \).

The incorporation of the interaction effect is another feature of our model. For traditional media, the interaction effect of the ad copy and the vehicle (e.g., magazine, television or radio program) is of interest (Ray and Sawyer 1971, Ray and Strong 1971). In our model, we incorporate this effect by providing a specific repetition function that depends on the site or even on the segment,

\[
p_n(X = k) = \frac{1}{1 + \exp[-(a_n + b_n k)]}.
\]

4. MODEL APPLICATION

In the preceding section, we developed a media planning model adapted to handle Internet-specific characteristics, such as high repeat exposure and decision making over a continuous duration. By using real Internet page view statistics and a repetition function that shows the ad performance by exposure frequency, we obtain useful insights with regard to enhancing the efficiency of Internet advertising. Our findings contrast with current Internet advertising practices, presented previously, which reflect highly concentrated ad planning devoted to a limited number of popular portals or search engines that provide a wide reach and high repeat exposures.

4.1. Data Description

Our Internet data are provided by KoreanClick (www.koreanclick.com), a Korean market research company that specializes in Internet audience measurement. KoreanClick maintains a panel of Internet users, selected on the basis of stratified proportions in South Korea, between 10 and 65 years of age. Candidates for the panel are contacted by a random digit dialing method. After the person agrees to be a panel member, he or she receives authorization from KoreanClick by both e-mail and regular mail to register as a panel member. The panel member is counted as an effective member if he or she connected to the Internet at least once during the four preceding weeks. The Internet usage behavior of the panel member is measured by a module, called “iTrack,” that captures the use of the active Internet browsers by the panel member at his or her home or office.
There are several major performance indicators of Internet usage, including page view, visitor, unique visitor, and reach (Novak and Hoffman 1997). A page view is the act of browsing a specific Web site. When a visitor accesses a Web page, a request is sent to the server hosting the page; a page view occurs when the page is fully loaded. At this point, an “impression” takes place because the consumer is exposed to the page contents, including advertisements. The page view measurement is equivalent to exposure in the case of traditional media such as television, radio, and magazines. Used mainly to determine advertising fees, page view illuminates the volume of browsing on the Web site.

The visitor indicator reflects a person who recorded at least one page view of a specific site using the Internet browser, whereas the unique visitor indicator relates to the net count of visitors when multiple visits by the same person are eliminated. Finally, reach is the number of unique visitors among the total Internet users during a fixed period. The reach explains the capacity of the Web site in terms of how widely it covers the total Internet population.

KoreanClick provides reliable Internet data that minimize the measurement problems raised by Drèze and Zurfyden (1998). The iTrack module identifies the visitor at all points of the Internet as long as it is installed; however, it can miss some visit data if the panel member accesses the Internet from a public place, such as schools or Internet cafés. This loss of information probably is marginal compared with the panel member’s major Internet usage at either the workplace or home. Furthermore, iTrack captures visit data that are cached by the proxy server of the panel member. If Web pages have a frame, iTrack counts page views of only the destination page and does not double count it as page view of the framed page.

We gathered data for four weeks, March 3–30, 2003. The effective panel members number 4149 for week 1, 4195 for week 2, 4125 for week 3, and 4148 for week 4. We retain a total of 3492 panel members who were effective during the entire four-week period. Of these effective panel members, 36% are women, and they average 32 years of age. To measure their Internet usage behavior, we use page views of the index pages (similar to the cover page of magazine) of ten selected Web sites in three major categories: community portal, news, and search engine (see Table 1).

<Insert Table 1: Site Profile around here>
We selected these ten sites for their popularity among all types of Internet users and because they have a relatively wide reach (if a Web site has a small reach, its page view data become more volatile). All sites experienced similar gender proportions among their visitors except news sites, which receive visits from more men. Portal sites and search engines are highly frequented; for example, the community portal site 1 reached more than 80% of the total Internet users during the week of March 3.

To measure the average page views, we divided all page views by the total Internet users, which represents a measure similar to the gross rating points frequently used by traditional media. The average page views indicate the overall exposure rate of the given site to all Internet users. However, the most important indicator is the average page views per visitor, which captures the exposure frequency among members who actually visited the site. This indicator tends to increase when the site reaches a wider range of the Internet population. For example, sites with very wide reach (e.g., sites 1, 8, and 9) record more than 20 page views per visitor.

In addition to the exposure frequency distribution, we use a repetition function of Internet advertising to complete our media planning model. Lee and Briley (2004) report on a repetition function of Internet advertising that measures the repeat exposure effects of the ad recall rate for a high exposure frequency using 10 online ad performance surveys of 10,667 observations. They find a statistically significant message recall function in terms of the exposure frequency and the probability of ad message recall after k exposures, $P(X = k) = \frac{1}{1 + \exp[-(-1.141 + 0.187 \ln(k + 1))]}$.

This monotonically increasing function has a lower bound of 21.41% and a upper bound of 100%. Unlike ADMOD (Aaker 1975), our model does not need an upper bound of less than 100% because the logit form repetition function provides a plateau within the range of plausible exposures. For example, the message recall rate of this function is expected to be 53.77% after 1000 exposures. As we illustrate in Figure 1, the probability of message recall increases more rapidly in the low exposure frequency area (i.e., fewer than 10 exposures) than in the high exposure frequency area.

Because the performance of an Internet advertisement does not increase linearly with the increase of the exposure frequency (similar to other media), the increased page views per visitor should alert
the advertiser of its possible wasteful spending, according to the low advertising spending
effectiveness among consumers who experience high repeat exposures.

4.2. Exposure Frequency Distribution

We apply the NBD to capture the exposure frequency distribution. As we mentioned in the model
development section, the NBD is a mixture of Poisson and gamma distributions. Whereas the Poisson
distribution estimates the distribution of events (ad exposures) over a fixed duration with one
parameter $\lambda$ to represent the mean of events, and thereby captures the distribution of ad exposures in a
discrete manner, the gamma distribution introduces the heterogeneity of consumers’ average exposure
rate $\lambda$. When the Poisson is mixed with the gamma, it becomes an NBD with two parameters: $\gamma$ as the
shape parameter and $\alpha$ as the scale parameter. The mean exposure rate therefore is computed as $\gamma/\alpha$.
The NBD provides two major advantages because of its flexible nature.

The NBD Fit of One-Week Data

The first flexibility of the NBD is its reasonable fit of Internet page view data, even though the
exposure rate heterogeneity is embedded. On the basis of the maximum likelihood principle, we
obtain parameter estimates of our ten listed sites, as we show in Table 2. We use the commercially
available software MATLAB for the maximum likelihood estimation and compare it to parameter
estimates obtained with MS Excel Solver. Both software programs provide similar values for the two
parameter estimates, so for our remaining analysis, we use the estimates obtained from MATLAB.

To check the goodness of the NBD fit for our sample of $N = 3492$, we proceed with the
Kolmogorov-Smirnov (K-S) test (Massey 1951) instead of the Pearson chi-square test, which is
inappropriate for a sample of large observations because its value is too sensitive to the number of data
points. The K-S test, a nonparametric test, compares the goodness of fit of a sample distribution $S_n(x)$
with that of a population by measuring the absolute distance between the two distributions. The
maximum absolute distance between the sample and the population cumulative distributions is $d =
\text{maximum } |F_0(x) - S_n(x)|$. If the distance is smaller than the critical value at a significant level of $\alpha$%,
the sample provides an appropriate fit to the population at that significance level. The distance should get smaller as the size of the sample N increases. In our case, we suppose that the exposure distribution of the population of Internet users, \(F_0(x)\), follows an NBD.

The critical value to test the goodness of fit is given by \(1.22/\sqrt{N}\) for \(\alpha = 10\%\), \(1.36/\sqrt{N}\) for \(\alpha = 5\%\), and \(1.63/\sqrt{N}\) for \(\alpha = 1\%\). The critical values are 2.06\%, 2.30\%, and 2.76\%, respectively. As we show in Table 3, the goodness of NBD fit is acceptable at \(\alpha = 10\%\) except for sites 1 and 9. Using the conservative standard of \(\alpha = 1\%\), all 10 sites have an appropriate fit of the NBD. The K-S distance tends to correlate with the variance of both the exposure distribution of sites and the reach. For sites with a narrow reach, the NBD can minimize the K-S distance if it effectively captures those nonvisitors that represent the greatest distribution density. However, sites with a wide reach have more dispersed distribution and greater variation among visitors. For these, the NBD must fit not only nonvisitors but also visitors across their exposure frequencies. This greater variation in the exposure frequency distribution is indicated by the larger distance of K-S.

**Extension of the Duration**

After the NBD captures the exposure frequency for a fixed duration, it can generate the exposure frequency distribution for a flexible duration with the same distribution parameters, \(\gamma\) and \(\alpha\). In turn, the duration can be used as a decision variable in the optimization model. To modulate the duration variable, we must have reasonably stable page view data or else incorporate additional parameters to correct the exposure distribution. However, in our application, the page view data are stable because in South Korea, the Internet infrastructure is highly advanced and the market is mature. During our data collection period, 49.4\% of the total population used the Internet, which shows that Internet usage had reached a mature level. Therefore, we can apply the NBD across flexible periods. In addition, as we show in Table 4, the NBD weekly parameter estimates are stable for the four-week period. As a consequence, the average exposure frequency and its variance are very similar.

<Insert Table 4: NBD Parameter Estimates for Week 1, 2, 3, and 4 around here>
As we expected, when we extend the duration of the exposure frequency distribution from one to four weeks using the same parameters of the NBD, it generates more errors. The maximum distance of K-S increases substantially from 1.27% to 7.77%. As in the one-week case, sites with wider reaches generate more errors. Although we lack a solid explanation of the estimation errors and their direction (i.e., under- or overestimation), the K-S distance for the four-week extension provides information about the range of errors that may be generated when researchers use one-week parameter estimates across multiple weeks (see Table 3).

Ad Efficiency Curve

Because our model is flexible enough to measure ad performance by varying the campaign duration, we must check the efficiency of advertising across our listed sites. The efficiency curve of an advertisement can be obtained as the combination of the cost and the effectiveness function of a decision variable. For example, Danaher and Rust (1994) report an efficiency curve as a function of gross rating points that can measure effectiveness, as a ratio to cost, in terms of reach, effective reach, incremental sales, or awareness, depending on the goals of the campaign. For our purposes, because the advertiser pays only for valid exposures on the Internet, we can measure the efficiency of an ad by computing the cost of increasing the effectiveness measure by a unit as a function of the campaign duration.

We obtain the cost function by multiplying the ad rate by the number of exposures. The ad rate is given by $r_i$ (on a CPM basis), which means that the advertiser pays $r$ for 1000 exposures at site $i$. The number of exposures is computed by summing the number of the total population exposures according to the exposure frequency distribution, $\sum_{k=1}^{K} M * k * f(X(t) = k)$. The cost function can be written as

$$\frac{r_i}{1000} \sum_{k=1}^{K} M * k * f(X(t) = k).$$

The effectiveness function, $\sum_{k=1}^{K} p(X = k) * M * f(X(t) = k)$, is the sum of the values of the effectiveness measure, which represents message recall and can be computed by multiplying the probability of recall at $k$ exposures, $p(X = k)$, by the number of consumers exposed $k$.
times, \( M^*f(X(t) = k) \), as a function of campaign duration \( t \). As a result, the ad efficiency curve,

\[
\frac{1000 \sum_{i=1}^{k} p(X = k) \cdot f(X(t) = k)}{\sum_{i=1}^{k} k \cdot f(X(t) = k)}
\]

, can be structured as a function of campaign duration.

In Figure 2, we present a graph that displays the efficiency curve of the number of consumers who recall the ad message when $1 is spent.

< Insert Figure 2: Advertising Efficiency Curve around here>

< Insert Table 5: Advertising Efficiency around here>

As the figure shows, ad efficiency deteriorates as the duration of the campaign increases, mainly due to the limited marginal increase of ad effectiveness that cannot compensate for the ad cost increase for high repeat exposures. Across sites, those with lower average exposures per visitor tend to enjoy higher ad efficiency. As we show in Table 5, site 4 outperforms all other sites because it has the fewest average page views per visitor (5.6), which minimizes spending for consumers who experience high repeat exposures. In contrast, ad efficiency is low among sites with higher average page views per visitor (e.g., sites 1, 8, and 9), because the advertiser must pay to expose the same consumers to the advertisement repeatedly and therefore experiences low returns.

However, no one site is systematically dominated by another site that performed better in the short term. Because each Web site has a different type of exposure frequency distribution, each provides a different efficiency curve projection as a function of duration. For example, site 1 records a better ad efficiency than site 2 for the first 3.5 weeks, but site 2 outperforms site 1 for longer periods because it accommodates more visitors in the effective range of low exposure frequency. From Figure 2, advertisers could imagine a horizontal line of iso-efficiency that compares ad efficiency across sites. If an advertiser wants to limit the budget per visitor, it can fix the duration of listed sites at that specific iso-efficiency line. An optimum set of site durations can be aligned along one such iso-efficiency line.

4.3. Optimization Results

The specific data for our optimization model are as follows:

- The total Internet population (M) is estimated as 23,658,097, according to KoreanClick.
The advertising cost is $1 per 1000 exposures (CPM, r), close to the market price.

We apply the same ad performance function of repeat exposure \( p(X = k) \), as reported by Lee and Briley (2004), to all ten sites.

The value of our objective function represents the number of panel members who recall the ad message after being exposed \( n \) times. Those who recalled the ad message without being exposed (24.21% of the total population) are not taken into account in our objective function.

Our resulting Internet media planning model, which combines the exposure frequency distribution of the listed sites and the message recall rate function of repeat exposure, is solved using the specialized Lingo 8.0 software to obtain the optimal set of listed site durations that maximizes the number of consumers who recall the ad message. We present the optimization results for three budget levels ($300,000, $500,000, and $700,000) in Table 6, along with the marginal increase (dual prices) of message recall visitors.

As we discussed previously, this nonlinear programming model has a complex objective function that requires a relatively long computation time; for this model, it took a computer powered by an Intel Celeron 2.2 GHz processor with 224 MB RAM almost ten minutes to run it.

The optimal durations of the listed sites provide some useful insights. First, all ten sites should focus on maximizing the number of message recalls, not on an optimum determined by the limited number of sites with wide reach. The selection of sites depends largely on the shape of the ad performance function. As the marginal return decreases, sites could enter the optimal set if their duration is fairly short, but unless there is a minimum duration requirement, all sites can be used to maximize the ad performance.

Second, the optimal ad duration is much longer for sites with low average exposures per visitor (sites 4, 5, 6, and 10) than for those with high averages (sites 1, 8, and 9). This result is a logical consequence of the ad efficiency curve presented previously, in that ad efficiency is much greater on sites with low average exposures per visitor than on site with high ones.

Third, the efficiency of an ad campaign deteriorates substantially, mainly due to diminishing returns, as the ad budget increases, as exemplified by the situation in which the price of the budget
decreases from 17.53 ($300,000) to 8.19 ($700,000). These findings indicate that the advertiser can add another 17.53 consumers who recall the ad message by increasing the budget by $1 when the initial budget is $300,000 but can add only 8.19 consumers when the starting budget is $700,000. This finding enables advertisers to compute both their ROI for the point at which ad performance will begin to deteriorate substantially.

Our findings appear to be in conflict with the current practice of Internet advertising. Advertisers in Europe selected 1.9 to 2.4 sites for a campaign for an average duration of 6–8 weeks. That is, advertisers tend to limit their number of sites. According to our results, these advertisers are wasting their budget substantially, because they have concentrated their campaign on a small number of sites for a long period, which generates too many consumers who are exposed too many times. To evaluate this potential waste, we compare ad performance (number of message recall visitors) for three cases: all sites are programmed (optimal), five sites are, and only three sites (one from each category) are, as in Table 7. To determine the performance difference according to the campaign duration and in line with current ad campaign practice, we divide the three-site cases into two subgroups each: with a four-week constraint on the maximum duration and without. In all cases, we established a budget of $500,000 and determined the optimal set of site durations to maximize the number of message recall visitors.

At first glance, the combinations of only three sites are largely dominated in performance by the optimal solution of all ten sites. When all sites are used, the advertiser can capture approximately 17 million visitors who recall the ad message. In contrast, for the combination of sites 1, 4, and 8, it captures approximately 10 million visitors, a drop of 39.8%, and for the combination of sites 3, 5, and 9, it captures only 7.5 million visitors, only 57.9% of the optimal case. Our findings regarding ad performance therefore demonstrate the terrible amount of waste that takes place in current Internet ad spending practices that limit the number of used sites. As the number gets smaller, the ad performance

---

2 These results are based on Internet ad campaigns from 3130 sites in 14 countries in Europe (LemonAd 2002). Its Internet link is unfortunately no longer available.
deteriorates because of the greater exposures in a less efficient, high repeat exposure zone. The reduced number of message recall consumers represents the magnitude of waste in the three-site cases, and the unattractive dual price reflects their inefficiency. The optimal solution with the ten sites suggests 10.94 as the dual price of budget spending; that is, the advertiser captures 10.94 visitors who recall the ad message for any extra $1 in ad budget spending. Each combination of the three sites costs the advertiser twice as much in ad budget than the optimal case to capture the same number of ad recall visitors. According to the Interactive Advertising Bureau (2004), a phenomenon of high ad spending concentration has occurred among Web sites with wide reach, in which the top ten Web sites account for more than 70% of the total ad spending. This phenomenon reinforces the magnitude of potential waste that runs rampant in current Internet ad practices.

5. Managerial Implications and Model Limitations

The main objective of our research is to provide marketing managers and ad agencies with an optimization tool that maximizes the ROI of their advertising budgets by highlighting the optimal combination of sites and the ideal campaign duration for each site. The advantages of using our model for Internet media planning, especially the flexibility of exposure distribution, are multifold. Because each site needs only two parameter estimates for the NBD to generate the exposure frequency of any duration, the computational burden is greatly reduced for a large number of Web sites; it is not necessary to generate the exposure frequency at each step of the optimization according to the combination of chosen sites. In addition, it minimizes complexity when an advertiser wants to conduct segment-level media planning for which it is necessary to obtain additional parameter estimates for each segment.

The simplicity of the data is another advantage of our model. The exposure frequency data that we use can be obtained easily from any market research company that keeps a user panel. Various types of exposure frequency data can be generated from the raw panel page view data, and the parameters of the NBD can be estimated easily by standard software such as MS Excel Solver or other packages such as MATLAB. In addition, the pretest of ad effectiveness becomes more and more
affordable on the Internet. Companies such as DynamicLogic and DoubleClick offer this service for less than $2,000 per ad.

Finally, our model helps advertisers calculate their ROI for Internet advertising by providing concrete numbers about ad performance and efficiency. Our model enables advertisers not only to optimize their Internet ad schedule but also to fix the right price for their Internet advertisements on the basis of the characteristics of the exposure distribution of sites. Our findings contrast with the current Web pricing practices, because the ad rate should be based on the average exposures per visitor rather than on its reach.

Despite these major advantages, our model does not include some aspects that should be addressed to refine its performance. First, we do not take into consideration exposure duplication across sites. As a result, our objective function may overestimate ad performance. The magnitude of this overestimation may increase when the duplication rate of chosen sites increases or the planning unit is restricted to integer values. The complex nature of Internet media planning, which requires varying duration variables and multiple sites, does not allow us to use a simple weight between two sites to reduce the duplication, as Headen, Klompmaker, and Teel applied (1976). A possible solution may be to compute the overlapped exposure distribution among sites. Park and Fader (2004) find a substantial improvement in predicting the intervisit behavior of two-site cases when they use a Sarmanov family of multivariate distributions (e.g., exponential timing process and gamma mixing distribution). Exposure distributions in a canonical form should be developed to correct the ad performance by assessing the width and depth of overlapped exposures across sites.

Second, our model does not address the forgetting effect. Whereas MEDIAC (Little and Lodish 1969) integrates the forgetting effect as a memory constant by updating the exposure level at each period, ADMOD (Aaker 1975) does not incorporate it directly. In our model, the forgetting effect may not need to be included due to the relatively short campaign durations, for which the forgetting effect is not statistically significant (Drèze and Husherr 2003). However, more refined research on repeat exposures of an Internet advertisement with varying conditions, such as context and time lap, should be conducted to enhance our model performance.
Third, our Internet media planning optimization model has a complex nonlinear objective function. If for small problems (as the one solved here), the computational time is not an issue, the search for optimal or nondominated solutions will be computationally time consuming for large problems. In this case, we would need to revert to the development of a heuristic approach to solve the problem in reasonable computational time. This is part of an ongoing research project.

6. Conclusion

The results of our Internet media planning model provide useful insights that can enhance the efficiency of Internet advertising. An advertiser must consider as many sites as possible because advertising on a wide selection of sites minimizes wasteful spending. If a campaign is concentrated on only two or three sites, the advertiser must extend the campaign duration of those chosen sites. This extension substantially penalizes the efficiency of the campaign, because it becomes more expensive to get visitors to recall the ad message. If the advertiser uses media planning tools developed for traditional media, it must carefully choose the proper indicator to select its sites. In the case of traditional media, an advertiser would prefer sites with a wide reach and high average exposures, as long as the ad rate is the same. But in the case of Internet, the advertiser must pay attention to another indicator: the average exposures per visitor. Because the pricing practice for an Internet advertisement is based on the number of exposures (page views), buyers should consider the efficiency issue first. In turn, because the ad effectiveness function of exposure frequency decreases marginally, the choice of a Web site with high average exposures per visitor minimizes the efficiency of the campaign. Therefore, the advertiser must favor those sites with low average exposures per visitor, as long as these sites meet the minimum reach requirements.
REFERENCES


Table 1 Site Profile

<table>
<thead>
<tr>
<th>Site</th>
<th>Type</th>
<th>Reach</th>
<th>Average Page Views</th>
<th>Average Page Views/Visitor</th>
<th>Male Visitors</th>
<th>Female Visitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Community portal</td>
<td>81.8%</td>
<td>20.7</td>
<td>25.3</td>
<td>61.5%</td>
<td>38.5%</td>
</tr>
<tr>
<td>2</td>
<td>Community portal</td>
<td>14.4%</td>
<td>2.0</td>
<td>13.7</td>
<td>61.4%</td>
<td>38.6%</td>
</tr>
<tr>
<td>3</td>
<td>Community portal</td>
<td>17.3%</td>
<td>2.1</td>
<td>11.9</td>
<td>62.9%</td>
<td>37.1%</td>
</tr>
<tr>
<td>4</td>
<td>Community portal</td>
<td>11.2%</td>
<td>0.6</td>
<td>5.6</td>
<td>64.2%</td>
<td>35.8%</td>
</tr>
<tr>
<td>5</td>
<td>News</td>
<td>9.7%</td>
<td>0.8</td>
<td>8.6</td>
<td>78.2%</td>
<td>21.8%</td>
</tr>
<tr>
<td>6</td>
<td>News</td>
<td>8.5%</td>
<td>0.7</td>
<td>8.6</td>
<td>77.8%</td>
<td>22.2%</td>
</tr>
<tr>
<td>7</td>
<td>News</td>
<td>5.3%</td>
<td>0.7</td>
<td>13.8</td>
<td>81.9%</td>
<td>18.1%</td>
</tr>
<tr>
<td>8</td>
<td>Search engine</td>
<td>49.2%</td>
<td>10.1</td>
<td>20.5</td>
<td>62.5%</td>
<td>37.5%</td>
</tr>
<tr>
<td>9</td>
<td>Search engine</td>
<td>60.3%</td>
<td>12.6</td>
<td>20.9</td>
<td>61.8%</td>
<td>38.2%</td>
</tr>
<tr>
<td>10</td>
<td>Search engine</td>
<td>30.2%</td>
<td>3.2</td>
<td>10.5</td>
<td>67.1%</td>
<td>32.9%</td>
</tr>
</tbody>
</table>

Table 2 NBD Parameter Estimates (MS Excel Solver and MATLAB)

<table>
<thead>
<tr>
<th>Site</th>
<th>$\gamma$ Excel Solver</th>
<th>$\gamma$ MATLAB</th>
<th>$\alpha$ Excel Solver</th>
<th>$\alpha$ MATLAB</th>
<th>Pageview Mean Excel Solver</th>
<th>Pageview Mean MATLAB</th>
<th>Pageview Variance Excel Solver</th>
<th>Pageview Variance MATLAB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.482</td>
<td>0.486</td>
<td>0.023</td>
<td>0.024</td>
<td>20.6</td>
<td>20.3</td>
<td>898.3</td>
<td>870.6</td>
</tr>
<tr>
<td>2</td>
<td>0.040</td>
<td>0.040</td>
<td>0.020</td>
<td>0.020</td>
<td>2.0</td>
<td>2.0</td>
<td>99.3</td>
<td>100.0</td>
</tr>
<tr>
<td>3</td>
<td>0.052</td>
<td>0.052</td>
<td>0.025</td>
<td>0.026</td>
<td>2.0</td>
<td>2.0</td>
<td>83.1</td>
<td>78.3</td>
</tr>
<tr>
<td>4</td>
<td>0.042</td>
<td>0.042</td>
<td>0.067</td>
<td>0.067</td>
<td>0.6</td>
<td>0.6</td>
<td>10.1</td>
<td>10.0</td>
</tr>
<tr>
<td>5</td>
<td>0.031</td>
<td>0.031</td>
<td>0.037</td>
<td>0.037</td>
<td>0.8</td>
<td>0.8</td>
<td>23.7</td>
<td>23.8</td>
</tr>
<tr>
<td>6</td>
<td>0.027</td>
<td>0.027</td>
<td>0.036</td>
<td>0.036</td>
<td>0.7</td>
<td>0.7</td>
<td>20.7</td>
<td>20.7</td>
</tr>
<tr>
<td>7</td>
<td>0.014</td>
<td>0.014</td>
<td>0.019</td>
<td>0.019</td>
<td>0.7</td>
<td>0.7</td>
<td>40.1</td>
<td>40.0</td>
</tr>
<tr>
<td>8</td>
<td>0.165</td>
<td>0.167</td>
<td>0.017</td>
<td>0.018</td>
<td>9.9</td>
<td>9.4</td>
<td>603.0</td>
<td>537.1</td>
</tr>
<tr>
<td>9</td>
<td>0.222</td>
<td>0.227</td>
<td>0.018</td>
<td>0.020</td>
<td>12.1</td>
<td>11.2</td>
<td>675.4</td>
<td>567.8</td>
</tr>
<tr>
<td>10</td>
<td>0.102</td>
<td>0.104</td>
<td>0.033</td>
<td>0.035</td>
<td>3.1</td>
<td>3.0</td>
<td>97.6</td>
<td>89.7</td>
</tr>
</tbody>
</table>

Table 3 Kolmorov-Smirnov Distance

<table>
<thead>
<tr>
<th></th>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 4</th>
<th>Site 5</th>
<th>Site 6</th>
<th>Site 7</th>
<th>Site 8</th>
<th>Site 9</th>
<th>Site 10</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 week</td>
<td>2.52%*</td>
<td>1.17%</td>
<td>1.66%</td>
<td>1.46%</td>
<td>0.25%</td>
<td>0.22%</td>
<td>0.34%</td>
<td>1.03%</td>
<td>2.20%**</td>
<td>1.86%</td>
<td>1.27%</td>
</tr>
<tr>
<td>4 weeks</td>
<td>11.88%*</td>
<td>3.61%*</td>
<td>7.30%*</td>
<td>9.62%*</td>
<td>4.23%*</td>
<td>5.87%*</td>
<td>3.40%*</td>
<td>11.91%*</td>
<td>10.30%*</td>
<td>9.59%*</td>
<td>7.77%</td>
</tr>
</tbody>
</table>

*** Significant at $\alpha = 10\%$.
** Significant at $\alpha = 5\%$.
* Significant at $\alpha = 1\%$.  

26
Table 4 NBD Parameter Estimates for Weeks 1–4

<table>
<thead>
<tr>
<th></th>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 4</th>
<th>Site 5</th>
<th>Site 6</th>
<th>Site 7</th>
<th>Site 8</th>
<th>Site 9</th>
<th>Site 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma ) Week1</td>
<td>0.486</td>
<td>0.040</td>
<td>0.052</td>
<td>0.042</td>
<td>0.031</td>
<td>0.027</td>
<td>0.014</td>
<td>0.167</td>
<td>0.227</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>0.498</td>
<td>0.038</td>
<td>0.058</td>
<td>0.046</td>
<td>0.032</td>
<td>0.029</td>
<td>0.015</td>
<td>0.180</td>
<td>0.235</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>0.483</td>
<td>0.044</td>
<td>0.054</td>
<td>0.044</td>
<td>0.032</td>
<td>0.033</td>
<td>0.018</td>
<td>0.185</td>
<td>0.234</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>0.467</td>
<td>0.038</td>
<td>0.053</td>
<td>0.042</td>
<td>0.030</td>
<td>0.030</td>
<td>0.017</td>
<td>0.186</td>
<td>0.220</td>
<td>0.105</td>
</tr>
<tr>
<td>( \alpha ) Week1</td>
<td>0.024</td>
<td>0.020</td>
<td>0.026</td>
<td>0.067</td>
<td>0.037</td>
<td>0.036</td>
<td>0.019</td>
<td>0.018</td>
<td>0.020</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>0.025</td>
<td>0.019</td>
<td>0.029</td>
<td>0.075</td>
<td>0.040</td>
<td>0.037</td>
<td>0.020</td>
<td>0.018</td>
<td>0.021</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>0.024</td>
<td>0.022</td>
<td>0.026</td>
<td>0.068</td>
<td>0.037</td>
<td>0.037</td>
<td>0.025</td>
<td>0.018</td>
<td>0.020</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>0.024</td>
<td>0.018</td>
<td>0.028</td>
<td>0.067</td>
<td>0.037</td>
<td>0.040</td>
<td>0.026</td>
<td>0.019</td>
<td>0.020</td>
<td>0.038</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week1</td>
<td>20.3</td>
<td>2.0</td>
<td>2.0</td>
<td>0.6</td>
<td>0.8</td>
<td>0.7</td>
<td>0.7</td>
<td>9.4</td>
<td>11.2</td>
<td>3.0</td>
</tr>
<tr>
<td>Week2</td>
<td>19.8</td>
<td>2.1</td>
<td>2.0</td>
<td>0.6</td>
<td>0.8</td>
<td>0.8</td>
<td>0.7</td>
<td>10.0</td>
<td>11.3</td>
<td>3.0</td>
</tr>
<tr>
<td>Week3</td>
<td>19.9</td>
<td>2.1</td>
<td>2.1</td>
<td>0.6</td>
<td>0.9</td>
<td>0.9</td>
<td>0.7</td>
<td>10.0</td>
<td>11.5</td>
<td>2.9</td>
</tr>
<tr>
<td>Week4</td>
<td>19.2</td>
<td>2.1</td>
<td>1.9</td>
<td>0.6</td>
<td>0.8</td>
<td>0.8</td>
<td>0.6</td>
<td>9.6</td>
<td>11.1</td>
<td>2.8</td>
</tr>
<tr>
<td>Variance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week1</td>
<td>870.6</td>
<td>100.0</td>
<td>78.3</td>
<td>10.0</td>
<td>23.8</td>
<td>20.7</td>
<td>40.0</td>
<td>537.1</td>
<td>567.8</td>
<td>89.7</td>
</tr>
<tr>
<td>Week2</td>
<td>810.6</td>
<td>111.9</td>
<td>72.6</td>
<td>8.6</td>
<td>20.4</td>
<td>21.5</td>
<td>37.9</td>
<td>565.9</td>
<td>553.8</td>
<td>83.0</td>
</tr>
<tr>
<td>Week3</td>
<td>838.0</td>
<td>97.0</td>
<td>82.3</td>
<td>10.0</td>
<td>24.6</td>
<td>25.4</td>
<td>30.0</td>
<td>555.3</td>
<td>578.1</td>
<td>81.6</td>
</tr>
<tr>
<td>Week4</td>
<td>804.1</td>
<td>117.8</td>
<td>70.1</td>
<td>10.1</td>
<td>22.2</td>
<td>20.3</td>
<td>26.0</td>
<td>508.4</td>
<td>577.0</td>
<td>75.7</td>
</tr>
</tbody>
</table>

Table 5 Advertising Efficiency

<table>
<thead>
<tr>
<th>Duration (week)</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
<th>3.5</th>
<th>4</th>
<th>4.5</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>61.88</td>
<td>33.85</td>
<td>23.52</td>
<td>18.09</td>
<td>14.73</td>
<td>12.45</td>
<td>10.81</td>
<td>9.58</td>
<td>8.63</td>
<td>7.88</td>
</tr>
<tr>
<td>Site 2</td>
<td>47.91</td>
<td>27.67</td>
<td>20.00</td>
<td>16.01</td>
<td>13.63</td>
<td>12.08</td>
<td>11.00</td>
<td>10.21</td>
<td>9.62</td>
<td>9.15</td>
</tr>
<tr>
<td>Site 3</td>
<td>116.91</td>
<td>70.86</td>
<td>52.10</td>
<td>41.65</td>
<td>34.92</td>
<td>30.19</td>
<td>26.67</td>
<td>23.96</td>
<td>21.80</td>
<td>20.04</td>
</tr>
<tr>
<td>Site 4</td>
<td>218.17</td>
<td>140.61</td>
<td>106.42</td>
<td>86.64</td>
<td>73.56</td>
<td>64.21</td>
<td>57.14</td>
<td>51.60</td>
<td>47.12</td>
<td>43.41</td>
</tr>
<tr>
<td>Site 5</td>
<td>82.45</td>
<td>48.93</td>
<td>35.62</td>
<td>28.32</td>
<td>23.68</td>
<td>20.48</td>
<td>18.15</td>
<td>16.40</td>
<td>15.04</td>
<td>13.96</td>
</tr>
<tr>
<td>Site 7</td>
<td>75.78</td>
<td>44.97</td>
<td>32.77</td>
<td>26.09</td>
<td>21.88</td>
<td>18.99</td>
<td>16.92</td>
<td>15.37</td>
<td>14.17</td>
<td>13.22</td>
</tr>
<tr>
<td>Site 8</td>
<td>47.23</td>
<td>26.20</td>
<td>18.38</td>
<td>14.29</td>
<td>11.80</td>
<td>10.16</td>
<td>9.02</td>
<td>8.18</td>
<td>7.55</td>
<td>7.06</td>
</tr>
<tr>
<td>Site 9</td>
<td>47.07</td>
<td>25.83</td>
<td>18.01</td>
<td>13.91</td>
<td>11.42</td>
<td>9.77</td>
<td>8.61</td>
<td>7.76</td>
<td>7.13</td>
<td>6.63</td>
</tr>
<tr>
<td>Site 10</td>
<td>82.76</td>
<td>48.03</td>
<td>34.48</td>
<td>27.13</td>
<td>22.47</td>
<td>19.26</td>
<td>16.92</td>
<td>15.14</td>
<td>13.76</td>
<td>12.65</td>
</tr>
</tbody>
</table>
Table 6: Optimal Site Duration with Three Budget Amounts

<table>
<thead>
<tr>
<th>Duration (weeks)</th>
<th>Budget Amount ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>300K</td>
</tr>
<tr>
<td>Site 1</td>
<td>0.17</td>
</tr>
<tr>
<td>Site 2</td>
<td>0.33</td>
</tr>
<tr>
<td>Site 3</td>
<td>0.42</td>
</tr>
<tr>
<td>Site 4</td>
<td>1.10</td>
</tr>
<tr>
<td>Site 5</td>
<td>0.62</td>
</tr>
<tr>
<td>Site 6</td>
<td>0.62</td>
</tr>
<tr>
<td>Site 7</td>
<td>0.33</td>
</tr>
<tr>
<td>Site 8</td>
<td>0.22</td>
</tr>
<tr>
<td>Site 9</td>
<td>0.22</td>
</tr>
<tr>
<td>Site 10</td>
<td>0.48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># of Message Recall</th>
<th>Dual Price (Ad Efficiency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1505080</td>
<td>17.53</td>
</tr>
<tr>
<td>17787410</td>
<td>10.94</td>
</tr>
<tr>
<td>19667450</td>
<td>8.19</td>
</tr>
</tbody>
</table>

Table 7 Ad Performance Comparison

<table>
<thead>
<tr>
<th>All Sites</th>
<th>Sites 1, 3, 5, 6, 7</th>
<th>Sites 1, 4, 8</th>
<th>Sites 3, 5, 9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max 4 weeks</td>
<td>Max 4 weeks</td>
<td>Max 4 weeks</td>
</tr>
<tr>
<td>Site 1</td>
<td>0.26</td>
<td>0.50</td>
<td>0.53</td>
</tr>
<tr>
<td>Site 2</td>
<td>0.59</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Site 3</td>
<td>0.73</td>
<td>2.82</td>
<td>3.27</td>
</tr>
<tr>
<td>Site 4</td>
<td>1.95</td>
<td>0.00</td>
<td>7.83</td>
</tr>
<tr>
<td>Site 5</td>
<td>1.11</td>
<td>5.42</td>
<td>4.00</td>
</tr>
<tr>
<td>Site 6</td>
<td>1.12</td>
<td>5.86</td>
<td>4.00</td>
</tr>
<tr>
<td>Site 7</td>
<td>0.60</td>
<td>4.31</td>
<td>4.00</td>
</tr>
<tr>
<td>Site 8</td>
<td>0.36</td>
<td>0.00</td>
<td>0.86</td>
</tr>
<tr>
<td>Site 9</td>
<td>0.36</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Site 10</td>
<td>0.83</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Message Recall Individuals

<table>
<thead>
<tr>
<th></th>
<th>17787410</th>
<th>9852930</th>
<th>9846052</th>
<th>10706070</th>
<th>10682060</th>
<th>7483794</th>
<th>7428267</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underperformance</td>
<td>-44.6%</td>
<td>-44.6%</td>
<td>-39.8%</td>
<td>-39.9%</td>
<td>-57.9%</td>
<td>-58.2%</td>
<td></td>
</tr>
<tr>
<td>Dual Price (Ad Efficiency)</td>
<td>10.94</td>
<td>5.45</td>
<td>5.19</td>
<td>5.53</td>
<td>5.21</td>
<td>4.02</td>
<td>3.29</td>
</tr>
<tr>
<td>Underperformance</td>
<td>-50.2%</td>
<td>-52.6%</td>
<td>-49.5%</td>
<td>-52.4%</td>
<td>-63.3%</td>
<td>-69.9%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 Repeat Exposure Effect of Message Recall Rate
Figure 2 Advertising Efficiency Curve