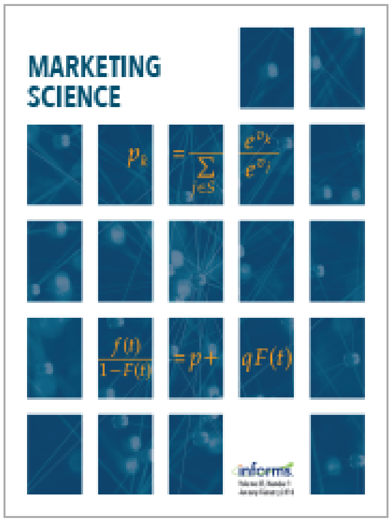

Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Optimal Internet Media Selection

Peter J. Danaher, Janghyuk Lee, Laoucine Kerbache,

MARKETING SCIENCE
Vol. 29, No. 2, March–April 2010, pp. 336–347
ISSN 0732-2399 | EISSN 1526-548X | 10 | 2902 | 0336

DOI 10.1287/mksc.1090.0507

Optimal Internet Media Selection

Peter J. Danaher
Melbourne Business School, Carlton, Victoria 3053, Australia, p.danaher@mbs.edu

Janghyuk Lee
Korea University Business School, An-am, Seong-buk, Seoul, South Korea, janglee@korea.ac.kr

Laoucine Kerbache
Department of OMIT and the Research Center GREGHEC, HEC School of Management, 78351 Paris, France, kerbache@hec.fr

In this study we develop a method that optimally selects online media vehicles and determines the number of advertising impressions that should be purchased and then served from each chosen website. As a starting point, we apply Danaher's [Danaher, P. J. 2007. Modeling page views across multiple websites with an application to Internet reach and frequency prediction. *Marketing Sci.* **26**(3) 422–437] multivariate negative binomial distribution (MNBD) for predicting online media exposure distributions. The MNBD is used as a component in the broader task of media selection. Rather than simply adapting previous selection methods used in traditional media, we show that the Internet poses some unique challenges. Specifically, online banner ads and other forms of online advertising are sold by methods that differ substantially from the way other media advertising is sold. We use a nonlinear optimization algorithm to solve the optimization problem and derive the optimum online media schedule. Data from an online audience measurement firm and an advertising agency are used to illustrate the speed and accuracy of our method, which is substantially quicker than using complete enumeration.

Key words: advertising; Internet marketing; media; optimization; probability models
History: Received: March 19, 2008; accepted: March 27, 2009; processed by Peter Fader. Published online in *Articles in Advance* July 23, 2009.

1. Introduction

The "dot com" crash of 2000 resulted in an immediate downturn in Internet advertising that lasted an additional two years. Subsequent to 2003, however, Internet advertising has grown at about 30% annually, to the point where online ad spend in the United States was over $21 billion in 2007 (IAB 2008). This growth has also been mirrored in Europe and Asia. Part of the reason for the slump and then rapid growth in online advertising has been a shift in the way advertising has been priced and sold. Advertising pricing models in traditional media are generally based on audience size, but the Internet's unique ability to track browsing behavior meant that early online ad campaigns were often priced on "pay-per-click." However, within a short time, click-through rates diminished to less than half of 1% (Manchanda et al. 2006). Indeed, the percentage of online ad campaigns priced on the basis of pay-per-click fell from 56% in 2000 to only 14% in 2007 (IAB 2002, 2008).

The rapid demise of the pay-per-click pricing model has been matched by an equally dramatic rise in online ad revenue allocated to paid or sponsored search, as provided by search engines such as Google and Yahoo!. Although paid search receives a lot of attention, over the past four years, the share of Internet ad spend devoted to this form of advertising has plateaued at about 40% (IAB 2005, 2006, 2007b, 2008). Instead, the silent, but steady, recent growth in online ad spend has been in display-type advertising (banners, pop-ups, skyscrapers, etc.), where ad prices are based on audience size, a pricing model conventionally termed cost-per-thousand (CPM). CPM-based online advertising has increased steadily since a low point in 2004 and has now attained almost 50% of total spend, to over $11 billion in 2007 (IAB 2008). This is more than was spent on radio advertising in the United States in that year (TNS Media Intelligence 2008).

The increasing spend on Internet display advertising is due to a realization of the benefits this advertising format has on awareness, recognition, and attitude formation compared with simple click-through. Manchanda et al. (2006, p. 99) review many industry and academic studies and conclude that "banner advertising has attitudinal effects and click-through is a poor measure of advertising response." Indeed, their own study shows a significant link between banner ad exposure and downstream purchase. Regarding paid search, Klaassen (2007) reports that it is largely confined to four major search engines (Google, Yahoo!, AOL, and MSN), who dominate the gross online ad spend for sponsored search. In contrast to these four websites, there are millions of others that accrue ad revenue based on CPMs. One of the major reasons for initial caution with CPM-based Internet advertising was a lack of comparability with traditional media (Smith 2003) and inconsistent measurement methods (Bhat et al. 2002, Coffey 2001). Today, these problems are largely resolved, and now it would be unusual for a major ad campaign not to include an online display component.

Despite this shift to and growth in CPM-based Internet advertising, there are currently no optimal media vehicle selection methods for the Internet in the published literature.[1] Therefore, the purpose of this study is to develop an optimal media selection method suitable for Internet campaigns using display advertising. Media selection on the Internet presents some challenges that do not arise in traditional media, and so our method is not a simple adaptation of previous methods. Instead, we use a previous audience size model that is customized to the Internet and additionally develop a budget function that is tailored to the unique way that online advertising "impressions" are priced and sold by Web publishers. For example, on the Internet it is possible to limit the number of times a person sees a particular ad. We find that our method is both accurate and computationally quick. Comparison with optimal schedules obtained by computationally slow complete enumeration shows that the schedules derived by our method are either the same or close to the true optimal schedule.

1.1. Audience Estimation and Optimal Scheduling for the Internet

A precursor to any media scheduling method is a requirement to develop reliable estimates of campaign audience size, such as reach and frequency (Meskauskas 2003). Danaher (2007) reviews several proprietary and nonproprietary methods to estimate the audience for online ad campaigns. One model, a generalization of the negative binomial distribution, was shown to be very accurate at predicting page views and then downstream audience size measures such as reach. However, obtaining good estimates of audience size is only part of the requirement for advertisers. Another key aspect is media selection, which optimally allocates the available budget across media vehicles so as to maximize an audience size measure (Bass and Lonsdale 1966, Rust 1986, Rossiter and Danaher 1998). This involves an examination of which media vehicles to select and how much to advertise in each. Hence, in this study, we build on previous work by Danaher (2007) in which he constructed a model for just Internet audience *exposure*. Here, we extend this work to Internet media *selection*, which requires more than just a model for media audience size. In particular, appropriate objective and budget functions must be developed. In turn, the selected media comprise the eventual media schedule.

The first comprehensive model for optimal media selection in traditional media was Little and Lodish's (1969) MEDIAC model, developed for TV media scheduling. Later developments were made by Aaker's (1975) ADMOD and Rust's (1985) VIDEAC models. Also see Danaher's 1991 review of operations research methods for optimal media selection. Heuristic optimization methods, such as the so-called "greedy algorithm," have proven popular and surprisingly accurate for optimizing media allocation (Rust 1986, Danaher 1991). There have also been a number of proprietary commercial software packages developed to optimize television advertising schedules. These include SuperMidas and X-pert (http://www.wpp.com). None of the proprietary methods has been benchmarked against the published methods, but the widespread uptake of commercial advertising optimization software is evidence that the advertising industry views them as extremely useful (Ephron 1998).

Even though optimal media selection is now commonplace in traditional media, there are no nonproprietary methods available for online media. As noted earlier, the two major forms of Internet advertising are based on paid search and audience size. Because paid search advertising is a relatively unique form of advertising, and depends mostly on the choice of target key words, website design and content, and how much the advertiser is willing to pay per click,[2] we do not consider this method of online advertising in this paper. Instead, we concentrate on Internet advertising scheduling for display ads, where the purpose is to reach a target group via a banner or some other

[1] By optimal media vehicle *selection*, we mean choosing websites and allocating advertising budget to the selected websites so as to maximize an audience measure, such as reach, while keeping within a budget constraint. There has been previous work on optimal placement and timing of banner ads within a single website. See, for example, Adler et al. (2002), Kumar et al. (2007), and Nakamura and Abe (2005). This placement and timing stage is done separately for each website but occurs subsequent to the media selection stage that we examine. Both stages comprise the eventual media schedule.

[2] In addition, Google, and more recently Yahoo!, position sponsored search advertisers on the match between what a Web user is looking for and whether a particular advertiser can fulfill that need (using a proprietary matching algorithm). For all major search engines, advertisers pay only when a user clicks on their displayed link. Such a revenue model is often termed "performance-based." See, for example, Seda (2004) for a detailed description of how to conduct search engine advertising.

form of visual advertisement. Our study focuses primarily on maximizing audience size measures across multiple websites. Here, we do not explicitly consider other media planning issues, such as the position, size, and animation of the ad, and creative fit with each website.

The rest of the paper is organized as follows. In the next section we develop objective functions based on a robust Internet exposure model. Section 3 then constructs a budget function that allows for a different share of advertising impressions in each website and possible ad restrictions at the individual level, both issues being unique to the Internet. We then describe the data used to illustrate our method in §4 and present the findings and applications in §5.

2. Development of Objective Functions

The starting point for any optimal media planning system is the formulation of an objective function. In traditional media, objectives to be maximized include reach, frequency, and effective frequency (Rust 1986), and the Internet is no different. Reach is the proportion of the target audience exposed at least once to the advertising campaign and is a very common advertising objective. Frequency is the average number of exposures per person among those reached (often expressed as gross rating points (GRPs) divided by the reach). Effective frequency is usually defined to be some threshold level of exposure, such as three (Krugman 1972, Naples 1979). That is, the objective might be to maximize the proportion of the target audience exposed at least three times. A model that forms the basis for measuring all of these objectives is the media exposure distribution.

2.1. Internet Advertising Exposure Distribution Models

A formal definition of the exposure distribution is as follows: let X_i be the number of exposures[3] a person has to website i, $X_i = 0, 1, 2, \ldots, i = 1, \ldots, m$, where m is the number of different websites. The exposure random variable to be modeled is $X = \sum_{i=1}^{m} X_i$, the total number of exposures to an advertising schedule. Therefore, X is a simple sum of random variables, but two nonignorable correlations make modeling it difficult (Danaher 1992), as will be seen later.

Danaher (2007) examined exposure distribution models for the Internet in detail.[4] For a single website, Danaher (2007) and Huang and Lin (2006) show that a negative binomial distribution (NBD) model fits the observed exposure distribution better than several other plausible models. Hence, for a single website, we also use the NBD with mass function

$$\Pr(X_i = x_i \mid r_i, \alpha_i, t_i) = \binom{x_i + r_i - 1}{x_i} \left(\frac{\alpha_i}{\alpha_i + t_i}\right)^{r_i} \left(\frac{t_i}{\alpha_i + t_i}\right)^{x_i}, \quad x_i = 0, 1, 2, \ldots, \qquad (1)$$

where r_i and α_i are the usual parameters for the gamma distribution. An additional parameter t_i, not incorporated by Danaher (2007), permits the NBD to be rescaled depending on the time interval used to estimate the parameters (see Lilien et al. 1992, p. 34). For instance, if one week of data is used to estimate the model, but an advertiser wants to predict the exposure distribution for a four week period, then all that changes is that t_i goes from one to four. This results from the additive property of the Poisson distribution, which underpins the NBD, and makes it extremely versatile for Internet exposure distribution applications.

For multiple websites, Danaher (2007) considered several possible models, including a multivariate generalization of the NBD. This multivariate NBD (denoted the MNBD) performed the best empirically as a model for predicting audiences to website ad campaigns, so we also use it in this study. The full details are given by Danaher (2007), but briefly, the MNBD uses a method developed by Sarmanov (1966) that combines several univariate distributions into a multivariate distribution, simultaneously taking into account pairwise and (possibly) higher-order interactions among the constituent random variables.[5] Park and Fader (2004) introduced the Sarmanov method to the marketing literature for an application to purchases across two websites, and Danaher and Hardie (2005) applied it to grocery store purchases between two categories (bacon and eggs).

The MNBD model for the full exposure distribution, with truncation after third-order terms, is

$$f(X_1, X_2, \ldots, X_m) = \left\{\prod_{i=1}^{m} f_i(X_i)\right\}\left[1 + \sum_{j_1<j_2} \omega_{j_1,j_2}\phi_{j_1}(x_{j_1})\phi_{j_2}(x_{j_2}) + \sum_{j_1<j_2<j_3} \omega_{j_1,j_2,j_3}\phi_{j_1}(x_{j_1})\phi_{j_2}(x_{j_2})\phi_{j_3}(x_{j_3})\right], \qquad (2)$$

[3] For Internet advertising, by "exposure" we mean an ad impression. Our data do not have explicit information on ad impressions, only page view impressions. Therefore, we treat an exposure and a page impression as equivalent.

[4] Further details on how univariate and multivariate exposure distributions are developed and applied in this setting are given in §A1 of the electronic companion, available as part of the online version that can be found at http://mktsci.pubs.informs.org.

[5] Also see in §A2 of the electronic companion, which discusses the sensitivity of the MNBD to pairwise correlations among websites, and the number of interaction terms to use in the Sarmanov model.

where in the case of the NBD,

$$\phi_i(x_i \mid t_i) = e^{-x_i} - \left(\frac{\alpha_i}{t_i(1-e^{-1})+\alpha_i}\right)^{r_i} \quad (3)$$

is called a mixing function and has the property that $\sum_{x_i} \phi_i(x_i) f_i(x_i) = 0$. The terms ω_{j_1,j_2} and ω_{j_1,j_2,j_3} are parameters measuring the second- and third-order associations, respectively, among the websites.

2.2. Objective Functions

Formulas for reach, average frequency, and effective reach can now be derived from this MNBD model, as we now demonstrate. Applying Equations (2) and (3) to model just reach, we have

$$\begin{aligned}\text{Reach} &= 1 - f(X_1=0, X_2=0, \ldots, X_m=0 \mid t_1, t_2, \ldots, t_m)\\ &= 1 - \left\{\prod_{i=1}^{m} f_i(x_i = 0 \mid r_i, \alpha_i, t_i)\right\}\\ &\quad \cdot \Bigg[1 + \sum_{j_1<j_2} \omega_{j_1,j_2}\left(1-\left(\frac{\alpha_{j_1}}{t_{j_1}(1-e^{-1})+\alpha_{j_1}}\right)^{r_{j_1}}\right)\\ &\quad \cdot \left(1-\left(\frac{\alpha_{j_2}}{t_{j_2}(1-e^{-1})+\alpha_{j_2}}\right)^{r_{j_2}}\right)\\ &\quad + \sum_{j_1<j_2<j_3} \omega_{j_1,j_2,j_3}\left(1-\left(\frac{\alpha_{j_1}}{t_{j_1}(1-e^{-1})+\alpha_{j_1}}\right)^{r_{j_1}}\right)\\ &\quad \cdot \left(1-\left(\frac{\alpha_{j_2}}{t_{j_2}(1-e^{-1})+\alpha_{j_2}}\right)^{r_{j_2}}\right)\\ &\quad \cdot \left(1-\left(\frac{\alpha_{j_3}}{t_{j_3}(1-e^{-1})+\alpha_{j_3}}\right)^{r_{j_3}}\right)\Bigg]. \end{aligned} \quad (4)$$

If the objective function is average frequency, then

$$\begin{aligned}\text{Av. Frequency} &= \frac{\text{GRPs}}{\text{Reach}}\\ &= \frac{\sum_{i=1}^{m} E[X_i]}{\text{Reach}}\\ &= \frac{\sum_{i=1}^{m} r_i t_i/\alpha_i}{\text{Reach}}. \end{aligned} \quad (5)$$

Last, if the objective function is effective reach at three exposures, then

$$\text{EffectiveReach} = \Pr(X \ge 3) = 1 - \sum_{x=0}^{2} f_X(x), \quad (6)$$

where $f_X(x)$ is the mass function for the distribution of total exposures $X = \sum_{i=1}^{m} X_i$, defined as

$$f_X(x) = \sum_{\{(x_1,\ldots,x_m):\, x_1+\cdots+x_m=x\}} f(X_1, X_2, \ldots, X_m), \quad x = 0, 1, 2, \ldots, \quad (7)$$

where the joint distribution $f(X_1, X_2, \ldots, X_m)$ is obtained from Equation (2).

Equations (4)–(6) are measures of just audience size. They are only part of the requirement for a system to optimally select websites. An equally important component is the budget function, and we now show how the Internet presents some unique challenges when it comes to formulating a budget function.

3. Development of the Budget Function

A very common measure of media cost that is well established in broadcast and print media is CPM (see, for example, Sissors and Baron 2002, Rossiter and Danaher 1998). CPM is the dollar cost charged by the broadcaster or publisher for 1,000 exposures among people in the target audience. CPM has now also become established in online media (for nonsearch advertising placements), but there is a slight modification. Instead of being the cost for 1,000 target audience people, it is the cost for 1,000 *impressions*.[6] This difference has important ramifications for Internet media planning, which makes it distinct from traditional media, as we now explain.

For traditional media, the 1,000 people exposed in a CPM measure are unique, whereas for online, the 1,000 impressions need not be served to 1,000 different people. The reason for this is that for online media the user controls the rate of ad delivery. For instance, if CNN.com puts the same ad on its home page for a week, then a twice-a-day visitor will be served 14 impressions of that ad during that week. Hence, depending on the incidence of repeat visits to a website, 1,000 impressions will result in an ad being delivered to much fewer than 1,000 people. It could be 100 people exposed 10 times, 50 people exposed 20 times, etc. (Chandler-Pepelnjak and Song 2004). This distinctive aspect of online CPMs must be allowed for in any optimal online media selection procedure. The incidence of high repeat visits to a website by the same person is of concern to advertisers who sometimes wish to limit the number of impressions delivered to visitors, known as "frequency capping." We now discuss media planning and scheduling issues peculiar to the Internet.

3.1. Share of Impressions

Websites that sell advertising on the basis of CPM usually set a time period of a week or a month in which the ad impressions are delivered to site visitors. For websites with a lot of traffic, ad impressions can accumulate very quickly—into the hundreds of

[6] The Internet Advertising Bureau defines an impression to be "a measurement of responses from a Web server to a page request from the user browser" (IAB 2007a). Included in the page served to a user is likely to be some form of advertising, such as banners, pop-ups, skyscrapers, etc.

thousands or millions. For example, at a typical CPM of $20, five million impressions in a week costs an advertiser $100,000, which is very expensive for a single website. Consequently, very few advertisers can afford to purchase every available impression, nor would they want to. It would be akin to buying every available commercial slot on a TV network for a week. Conversely, probably very few Web publishers would want to give exclusive coverage to a single advertiser even for a week. Therefore, Web advertisers must settle for a share of the available impressions per site for a fixed time period. The concept of share of impressions (SOI) has been an established measure in regard to Internet media planning for some time (see, for example, Ryan 2001).

We denote this share as s_i for website i, which can be thought of as the probability that a website visitor who requests a Web page has a particular ad delivered on that page. For instance, Dell might purchase a 5% share of impressions for a month-long campaign on MSN.com. This means that each page served to MSN.com visitors in that month has a 5% chance of containing a Dell advertisement. Accommodating a share of impressions into the reach estimate is straightforward because of the additive property of the Poisson distribution. All that changes is that the α_i parameter in Equation (4) is divided by s_i (Lilien et al. 1992, p. 34). A similar adjustment is required for the mixing functions in Equation (3).

An additional factor alongside the share of impressions is a possible guaranteed number of impressions during the campaign period. Figures of 100,000 to 500,000 impressions are common (Zeff and Aronson 1997). Such impression targets are usually easy for large websites to fulfill, but if the site falls short of its guarantee, then a "make good" is expected. This might entail an extended campaign period until the impression guarantee is achieved or a credit is applied to the next advertising purchase.

3.2. Frequency Capping

As mentioned earlier, online advertising has a particular quirk whereby the same person can be exposed to an ad many times at a rate determined by the user, not the publisher. Chandler-Pepelnjak and Song (2004) report the findings from 38 online ad campaigns and reveal that sales conversions and profits from banner ads are highest for somewhere between three and 10 impressions. They therefore recommend some form of frequency capping where possible. Frequency capping can be executed by using a cookie to identify whether a particular computer has previously visited a website and therefore been served an ad. Visitors that have been served an ad—say, 10 times—already will not be served any more ads for the duration of the campaign. Of course, some people do regularly clear their cookies (comScore 2007), but the majority do not, so that frequency capping can be used to increase the number of unique visitors to a site. This will increase the reach. It will also mean that advertisers will not waste money on the "...thousands of users who receive hundreds of impressions without any response" (Chandler-Pepelnjak and Song 2004, p. 4).

Operationalizing frequency capping in our framework is relatively straightforward. Recall that we can think of s_i as the probability a visitor will be served an impression on website i. If we additionally apply a frequency cap, then s_i becomes a conditional probability, with an impression being served conditional on the user having *not* exceeded the frequency cap. The probability of not exceeding the cap is denoted $\Pr(X_i \leq \text{cap})$ and can be obtained from the univariate NBD model as $\sum_{x=0}^{\text{cap}} \Pr(X_i = x)$. Now, the conditional probability for being served an ad impression, given that the frequency cap has not been exceeded is

$$s_i / \Pr(X_i \leq \text{cap}). \tag{8}$$

Notice that the cap is the same for each website. It is generally impossible to impose a cap on total exposures ($X = \sum_i X_i$) rather than X_i, because individual websites typically do not have information about the browsing behavior of visits to sites other than their own.[7]

3.3. Notation for Budget Constraint

As with conventional media planning, advertisers in the online environment must also work within a total budget, which we denote as B. This might be disaggregated into a budget per website, B_i, so that $\sum_{i=1}^{m} B_i \leq B$. All websites selling advertising based on audience size report a rate-card CPM, denoted c_i, for website i. In addition, they will give standard time intervals for an ad to be delivered on their website, such as a week, a month, or longer. Alternatively, they might report a fixed cost for a week or month of advertising and guarantee a certain number of impressions within that time period. Under this arrangement, an advertiser could set B_i to be the website's fixed cost and select a subset of sites that maximizes reach subject to $\sum_{i=1}^{m} B_i \leq B$.

For the NBD model, the expected number of impressions per person in the time interval $[0, t_i]$ is $r_i t_i / \alpha_i$. Hence, the expected total number of impressions among the target population to website i is $N r_i t_i / \alpha_i$, where N is the target population size. For a share, s_i, of the total impressions on website i in a time interval $[0, t_i]$, the expected number of ad impressions

[7] Frequency capping of total exposures should be possible when a third-party ad-serving company has control over the delivery of impressions for the entire campaign. The same general principle of adjusting the s_i values can still be applied in this situation.

delivered is $Ns_ir_it_i/\alpha_i$, at an expected cost per website of $(N/1{,}000)c_is_ir_it_i/\alpha_i$. The overall budget constraint is therefore

$$(N/1{,}000)\sum_{i=1}^{m} c_is_ir_it_i/\alpha_i \le B. \tag{9}$$

3.4. Statement of Optimization Problem

Applying the share of impressions adjustment and the budget constraint in Equation (9), as well as the reach model in Equation (4), the online advertising optimization problem for maximizing reach[8] by varying $(s_1, s_2, \ldots, s_m)$ can be stated formally as follows:

maximize

$$\begin{aligned}
\text{Reach} &= 1 - f(X_1=0, X_2=0, \ldots, X_m=0 \mid t_1, t_2, \ldots, t_m, s_1, s_2, \ldots, s_m) \\
&= 1 - \left\{\prod_{i=1}^{m} f_i(x_i = 0 \mid r_i, \alpha_i, t_i, s_i)\right\} \\
&\quad \cdot \left[1 + \sum_{j_1<j_2} \omega_{j_1,j_2}\left(1 - \left(\frac{\alpha_{j_1}}{s_{j_1}t_{j_1}(1-e^{-1})+\alpha_{j_1}}\right)^{r_{j_1}}\right)\right. \\
&\quad \cdot \left(1 - \left(\frac{\alpha_{j_2}}{s_{j_2}t_{j_2}(1-e^{-1})+\alpha_{j_2}}\right)^{r_{j_2}}\right) \\
&\quad + \sum_{j_1<j_2<j_3} \omega_{j_1,j_2,j_3}\left(1 - \left(\frac{\alpha_{j_1}}{s_{j_1}t_{j_1}(1-e^{-1})+\alpha_{j_1}}\right)^{r_{j_1}}\right) \\
&\quad \cdot \left(1 - \left(\frac{\alpha_{j_2}}{s_{j_2}t_{j_2}(1-e^{-1})+\alpha_{j_2}}\right)^{r_{j_2}}\right) \\
&\quad \left.\cdot \left(1 - \left(\frac{\alpha_{j_3}}{s_{j_3}t_{j_3}(1-e^{-1})+\alpha_{j_3}}\right)^{r_{j_3}}\right)\right]
\end{aligned} \tag{10}$$

subject to

$$0 \le s_i \le 1, \quad \text{and}$$
$$\frac{N}{1{,}000}\sum_{i=1}^{m} \frac{s_ic_ir_it_i}{\alpha_i} \le B. \tag{11}$$

If frequency capping is to be used, the only change to the optimization problem setup is to use Equation (8) where each s_i is replaced with $s_i/\Pr(X_i \le \text{cap})$, with the denominator probability obtained from the univariate NBD in Equation (1).

Equation (11) is the budget constraint for Internet media selection where websites permit an advertiser to purchase any share of impressions for a fixed time period, and the amount paid will vary depending on the share purchased. That is, the amount paid by the advertiser to the website is determined by the advertiser's chosen share of impressions, the website's CPM, and the expected total impressions for the campaign duration. This might be described as the "fully flexible" allocation method. However, other websites recommend or require that an advertiser have a minimum spend on their site and therefore state a fixed charge per period of advertising (such as a week or month). If we denote the fixed period cost per website as c_i^f, then an alternative optimization problem is to vary $(s_1, s_2, \ldots, s_m)$ so as to maximize reach, subject to $0 \le s_i \le 1$, and

$$\sum_{i=1}^{m} c_i^f I_{\{s_i>0\}} \le B, \tag{12}$$

where $I_{\{s_i>0\}}$ is an indicator function, being 1 if any share of impressions are allocated to site i and 0 otherwise. We call this the "fixed cost per website" allocation method.

3.5. Parameter Estimation

We follow Danaher (2007) and estimate the parameters of the univariate NBD models by the method of means and zeros. The bivariate and trivariate association parameters (ω_{j_1,j_2} and ω_{j_1,j_2,j_3}) are estimated by using Danaher's (2007) Equations (9) and (11), which equate the observed nonreach to that implied by the parametric model. Maximum likelihood estimation is also possible, but Danaher (2007) notes that it does not do as well at reach estimation, so we do not use it here.

4. Data

4.1. KoreanClick

Our Internet data are provided by KoreanClick (http://www.koreanclick.com), a Korean market research company that specializes in Internet audience measurement. KoreanClick maintains a panel of Internet users, aged between 6 and 75, selected on the basis of stratified sampling in South Korea. Potential candidates for the panel are selected by telephone random digit dialing. After the person agrees to be a panel member, he or she receives authorization from KoreanClick by both e-mail and postal mail to register as a panel member. The panel member is counted as active if they connect to the Internet at least once during the previous four weeks. The Internet usage behavior of each panelist is monitored by a module, called "iTrack," that captures all Internet activity by a panel member at their home or office. This method is very similar to that used by comScore (Coffey 2001) and Nielsen/Netratings in the United States and many other countries but has the added advantage of capturing panelists' Internet use at work as well as home. The only major difference is that KoreanClick's panel size is around 6,500, whereas comScore's and

[8] We work primarily with reach, but it is easy to change the objective function to be either average frequency or effective reach, as given, respectively, in Equations (5) and (6). See §5.2.

Netratings' are, respectively, one million and 30,000 people. This makes KoreanClick information less suitable for purchase transactions, where conversion rates are about 2%, but its sample size is ample for website audience estimation, especially for popular sites, as used in our application.

There are several major performance indicators of Internet usage, including page views, visitors, unique visitors, and reach (Novak and Hoffman 1997). A page view is the act of browsing a specific website page. When a visitor accesses a Web page, a request is sent to the server hosting the page and a page view occurs when the page is fully loaded. At this point, an impression or exposure is delivered, as the site visitor is (potentially) exposed to the page contents, including advertisements. Following Danaher (2007), we define a page view as equivalent to an exposure in the case of traditional media such as television, radio, and magazines.

We gathered data for two four-week periods, March 3–30 and June 2–29, 2003. The number of "active" panel members, who actually used the Internet in these periods, is 4,729 and 4,681, respectively, for March and June. We use March for model estimation and June for validation. According to KoreanClick, the "Internet population" in South Korea at the time these data were obtained was 23,658,097. However, because only 4,729 of the 6,468 total panelists actually used the Internet in March, the effective Internet population size for our data is $N = 23{,}658{,}097 \times (4{,}729/6{,}468) = 17{,}297{,}331$.

4.2. Website Description

To measure a panelist's Internet usage behavior, we use page views of the index pages (similar to the cover page of a magazine) of 10 selected websites in three major categories: community portal, news, and search engine, as shown in Table 1. We selected these 10 sites for their popularity among all types of Internet users and because they have a relatively high reach (if a website has a small reach, its page view data become more volatile). All sites attract similar gender proportions among their visitors except news sites, which receive more visits from men. Portal sites and search engines are highly visited; for example, the community portal site 1 (http://www.Daum.net) reached more than 70% of the total active Internet users in March 2003.

To measure the average impressions, we divided all page impressions by total Internet users, being a measure similar to gross rating points, which is frequently used by traditional media. The average page impressions indicates the overall exposure rate of the given site to all Internet users.

4.3. Website Costs

Table 1 also gives realistic CPM and fixed cost per week advertising rates for the 10 websites, as provided by Daehong, a Korean advertising agency. For example, MSN.co.kr (the Korean version of MSN.com) charges $10 per thousand impressions, or alternatively, $8,000 for an ad placement for one week on their website. For a CPM of $10, it is not difficult to calculate that $8,000 amounts to a purchase of $(8{,}000/10) \times 1{,}000 = 800{,}000$ impressions. Because the total impressions for MSN.co.kr in a week is 3,269,306, the share of impressions for a weekly ad posting is 0.2447. Hence, an advertiser buying space on MSN.co.kr for a week will have their ad delivered on about every fourth page impression that is served up to the site's visitors.

We additionally report frequency capping information in Table 1. The probabilities given in Table 1 are those used to adjust the s_i values, as shown in §3.2—namely, replace s_i with $s_i/\Pr(X_i \le \text{cap})$, where $\Pr(X_i \le \text{cap})$ comes from the univariate NBD model. For instance, the most popular website, Daum.net, delivers over 64 million impressions per week. If visitors are limited to at most 10 ad impressions of a particular advertisement, then 90.7% of

Table 1 Audience and Cost Information for Korean Websites

			March 2003						Frequency cap $\Pr(X_i \le \text{cap})$		June 2003	
Site no.	Domain home page	Category	Reach	Mean page Impressions	Avg. freq	Total impressions	CPM c_i, $	Weekly fixed cost c_i^f, $	cap = 10	cap = 5	Reach	Avg. freq
1	Daum.net	Portal	70.4	3.74	5.3	64,654,720	15	15,000	0.907	0.765	69.1	6.3
2	DreamWiz.com	Portal	12.0	0.42	3.5	7,259,028	8	9,000	0.993	0.978	10.3	3.6
3	MSN.co.kr	Portal	9.7	0.19	1.9	3,269,306	10	8,000	0.999	0.996	9.4	2.1
4	Chosun.com	News	7.8	0.20	2.6	3,477,751	10	7,000	0.998	0.992	7.2	2.6
5	JOINS.com	News	7.1	0.18	2.6	3,137,656	8	5,000	0.999	0.993	6.1	2.3
6	hani.co.kr	News	4.5	0.15	3.4	2,589,114	8	5,000	0.998	0.993	3.3	2.7
7	donga.com	News	5.2	0.13	2.5	2,267,304	10	5,000	0.999	0.995	4.6	2.1
8	Naver.com	Search engine	41.6	1.82	4.4	31,405,812	10	7,000	0.961	0.896	47.8	4.2
9	kr.yahoo.com	Search engine	52.0	2.17	4.2	37,571,416	12	5,000	0.957	0.874	49.0	4.1
10	empas.com	Search engine	25.3	0.69	2.7	12,009,395	10	12,000	0.994	0.971	24.2	2.5

visitors will not exceed the cap, meaning that about 9% will exceed the cap. Once this happens, these 9% of visitors will not receive any more of the designated ads (insofar as this is possible through the use of cookies) for the remainder of the week-long period. As would be expected, when the cap is lowered to five impressions, a greater percentage (33.5%) exceed the cap. For the other nine websites, exceeding the cap of 10 impressions is rare and is also not common for a cap of five, indicating that "frequency overload" for these sites is not especially problematic.

5. Results

5.1. Optimizing Reach

We wrote a program in FORTRAN to find the optimal solution that maximizes reach subject to either of the cost constraints in Equations (11) and (12). Given that we have a nonlinear objective function, a subroutine named NCONF from the IMSL (1997) library was used to solve the optimization problem.[9] This algorithm uses a successive quadratic programming method (Schittkowski 1983). The same algorithm and subroutine are used to maximize effective reach and average frequency. We set the campaign duration to be either one or four week(s) (thereby using data for either the first or all four weeks of March) and had total budget levels of \$50,000 or \$100,000. No frequency capping, as well as frequency caps of 5 and 10 impressions, was also used. In all cases, our method arrives at the optimum solution in less than 0.04 seconds using an IBM laptop running at 1.6 GHz.

Tables 2(a) and 2(b), respectively, report the optimal reach values for total budgets of \$50,000 and \$100,000. For example, a campaign with a budget of \$50,000, lasting one week with no frequency cap, achieves a reach of 21.8% when the purchaser has full flexibility and 20.4% when there is a fixed cost per website. Table 3 gives the SOI, number of impressions, and cost for these two optimal campaigns. Table 3 shows that for the fully flexible schedule, 9 of the 10 possible websites should be used, and the fixed-cost schedule uses 7 sites. Interestingly, in both cases, the most popular website (Daum.net) is not in the optimal schedule. This is likely due to its high CPM relative to its reach (see §A3 of the electronic companion).

As a point of comparison, a naïve media planner might choose to spend the entire budget on the website with the greatest reach. In this case, spending \$50,000 on Daum.net buys a SOI of 5.16% and achieves a reach of 15.5%. This is well below the optimum reach of 21.8% in Table 2(a). A less naïve, but intuitively reasonable, media schedule is to buy week-long campaigns in each of the top five sites (numbered 1, 2, 8, 9, and 10). The cost of doing this is within budget, at \$48,000, but the reach is only 19.5%, again lower than the optimal reach obtained by our method. Hence, it is apparent that using an intuitive heuristic, such as picking the most popular sites, does not guarantee the highest overall reach.[10] Part of the reason for this is the correlation in exposure among websites. For example, website 7 (donga.com) has moderate correlation with all the other news sites. Therefore, this site is not included in the optimal schedule which has a fixed cost per website, where three other news sites are already included.[11]

Also worth noting in Table 2(a) is that in every case the optimal reach for the fully flexible purchase is higher than when sites charge a fixed cost for a set time period. This is because allowing more flexibility in impression purchases permits the budget to be spread over more websites, thereby increasing the overall reach.

Table 2(a) also shows the benefit of frequency capping, with the reach increasing as the frequency cap is

Table 2(a) Reach for Different Campaign Durations and Frequency Capping Levels (Budget = \$50,000)

No. campaign weeks	Website purchase flexibility	Frequency cap (%)		
		None	10	5
1	Full	21.8	22.2	23.0
1	None	20.4	20.6	21.4
4	Full	22.7	26.5	31.4
4	None	21.6	24.8	28.6

Table 2(b) Reach for Different Campaign Durations and Frequency Capping Levels (Budget = \$100,000)

No. campaign weeks	Website purchase flexibility	Frequency cap (%)		
		None	10	5
1	Full	35.0	35.5	36.8
1	None	33.6	34.2	35.9
4	Full	36.8	42.7	48.5
4	None	35.6	41.6	46.3

[9] Implemention details for this subroutine are in §A1.3 of the electronic companion.

[10] Such intuitive schedules that choose just the dominant websites are common in both the United States (Klaassen 2007) and Europe (LemonAd 2002). Indeed, concentration of online advertising spend in fewer, but larger, websites is increasing (Klaassen 2007).

[11] As a matter of interest, we also obtained optimal schedules when it is assumed there is no association among websites (i.e., set $\omega_{j_1,j_2} = 0$ and $\omega_{j_1,j_2,j_3} = 0$ in Equation (2)). In every case the reach for this alternative optimal schedule is lower than when correlations are included. The difference increases for larger budgets, stronger correlations, and more websites. This justifies the inclusion of the bivariate and trivariate associations into the media exposure model. Also see §A2.1 in the electronic companion.

Table 3 Optimal Schedules for Korean Websites; Budget Is $50,000 Over a One-Week Period

		No frequency capping						Frequency cap = 5 impressions					
		Fully flexible			Fixed cost per website			Fully flexible			Fixed cost per website		
Site no.	Domain home page	SOI	Impressions	Cost	SOI	Impressions	Cost	SOI	Impressions[a]	Cost	SOI	Impressions[a]	Cost
1	Daum.net	0	0	0	0	0	0	0	0	0	0.0155	1,000,000	15,000
2	DreamWiz.com	0.0920	667,825	5,343	0.1550	1,125,000	9,000	0.0828	601,396	4,812	0.1550	1,125,000	9,000
3	MSN.co.kr	0.1419	463,401	4,634	0	0	0	0.0613	200,187	2,002	0	0	0
4	Chosun.com	0.0468	162,864	1,629	0.2013	700,000	7,000	0.0366	127,249	1,272	0.2013	700,000	7,000
5	JOINS.com	0.1322	414,729	3,318	0.1992	625,000	5,000	0.0941	295,320	2,363	0	0	0
6	hani.co.kr	0.0621	160,742	1,286	0.2414	675,000	5,000	0.0558	144,579	1,157	0	0	0
7	donga.com	0.0232	52,680	527	0	0	0	0.0176	39,803	398	0	0	0
8	Naver.com	0.0464	1,456,107	14,561	0.0223	700,000	7,000	0.0482	1,512,465	15,125	0.0223	700,000	7,000
9	kr.yahoo.com	0.0196	736,174	8,834	0.0111	416,667	5,000	0.0319	1,198,250	14,379	0	0	0
10	empas.com	0.0822	986,903	9,869	0.0999	1,200,000	12,000	0.0707	849,362	8,494	0.0999	1,200,000	12,000
		Reach = 21.8%			Reach = 20.4%			Reach = 23.0%			Reach = 21.4%		

[a]This is the total number of impressions, with the proviso that no one receives more than five.

lowered, as would be expected. Also demonstrated is the positive effect on reach if a campaign has longer duration. Going to four weeks from one week allows extra time for more unique visitors, thereby increasing the reach. If, in addition, a frequency cap is applied, the reach can increase substantially. For example, for a budget of $50,000 with a frequency cap of five in a four-week campaign, the fully flexible schedule achieves a reach of 31.4% compared with the no cap, one-week, fully flexible schedule optimal reach of 21.8%. The nearly 10-percentage-point increase is achieved *at no additional cost*.

The rightmost half of Table 3 gives the optimal schedule when a frequency cap of five is applied and the budget is $50,000. For the fully flexible schedule, the number of impressions decline for sites 2, 3, 4, 5, 6, 7, and 10, but increase for sites 8 and 9. It must be remembered that these impressions have the proviso that, where possible, no person receives more than five ad impressions. There is a large increase in the number of impressions purchased for the Korean version of Yahoo.com, as it has a lower probability of being below the cap (0.874) than most of the other sites. The same is true of Naver.com. Hence, the additional impressions purchased will be served to a larger pool of unique (new) visitors.

Table 2(b) mirrors 2(a) but with double the budget. Diminishing returns effects are apparent, because the one-week, no cap, fully flexible optimal schedule produces a reach of 35.0%, up from 21.8%, which is not a doubling of the reach. Again, the same effects of frequency capping and longer campaign duration are evident for the optimal schedules with a larger budget.

5.2. Optimizing Other Audience Measures

In addition to reach, we now obtain optimal schedules for three other objective functions, being effective reach at three exposures, effective reach subject to a minimum reach,[12] and average frequency. Maximizing effective reach subject to a reach minimum is very similar to maximizing $\Pr(X \geq 3)$ but adds a further constraint in addition to those given in §3.4—namely, that $\Pr(X \geq 1) \geq r_{\min}$, where $r_{\min}$ is the specified minimum reach.

Table 4 reports the optimal schedules. As seen from §5.1, the highest reach is obtained by spreading the budget over 9 of the 10 available websites. By contrast, maximizing effective reach requires the concentration of ad spend into vehicles that are the most cost effective at obtaining high repeat exposures. The second, fifth, and sixth sites fulfill these criteria. If an additional minimum reach threshold is added, the optimal schedule spreads the budget over more sites than for effective reach. We chose a minimum reach of 16%, being midway between the lowest reach in Table 4 (10.6%) and the highest (21.8%). Last, to maximize average frequency, the media strategy requires a heavy concentration in a handful of websites. Indeed, in this example, all the available impressions are purchased for sites 5 and 6, with the balance of the ad budget spent on donga.com. These are all cost-effective sites, where the cost of purchasing all the impressions is less than the total budget of $50,000.

5.3. Optimization by Complete Enumeration

A slow but sure method to optimally select online media vehicles is to laboriously calculate the reach for every possible campaign that stays within the budget. This method is known as complete enumeration and provides a benchmark to compare with the optimal reach obtained via our model-based method. We operationalize the complete enumeration method by using an empirically derived estimate of reach, as detailed

[12] We thank the area editor for this suggestion.

Table 4 Optimal Schedules for Alternative Objective Functions; Budget Is $50,000 Over a One-Week Period

		Objective function											
		Reach			Effective reach			Effective reach subject to reach ≥ 16%			Average frequency		
Site no.	Website	SOI	Impressions	Cost	SOI	Impressions	Cost	SOI	Impressions	Cost	SOI	Impressions	Cost
1	Daum.net	0	0	0	0	0	0	0	0	0	0	0	0
2	DreamWiz.com	0.0920	667,825	5,343	0.5492	3,987,457	31,900	0.4370	3,172,562	25,380	0	0	0
3	MSN.co.kr	0.1419	463,401	4,634	0	0	0	0	0	0	0	0	0
4	Chosun.com	0.0468	162,864	1,629	0	0	0	0.0365	126,795	1,268	0	0	0
5	JOINS.com	0.1322	414,729	3,318	0.3469	1,088,665	8,709	0.4294	1,347,515	10,780	1.0	3,137,656	25,107
6	hani.co.kr	0.0621	160,742	1,286	0.4533	1,173,879	9,391	0.4627	1,198,281	9,586	1.0	2,589,114	20,717
7	donga.com	0.0232	52,680	527	0	0	0	0.0195	44,322	443	0.1841	417,570	4,176
8	Naver.com	0.0464	1,456,107	14,561	0	0	0	0.0047	149,082	1,491	0	0	0
9	kr.yahoo.com	0.0196	736,174	8,834	0	0	0	0	0	0	0	0	0
10	empas.com	0.0822	986,903	9,869	0	0	0	0.0088	105,114	1,051	0	0	0
Reach			21.8%			14.7%			16.0%			10.6%	
Effective reach			1.50%			4.55%			4.38%			4.28%	
Average frequency			1.35			2.46			2.22			3.36	

in §A4 of the electronic companion. Briefly, we use the raw data to simulate the exposure distribution for share-of-impression values at each website while staying within the budget constraint. Because the simulated exposure distribution is not based on the MNBD model, this is a test of both model accuracy as well as the efficiency of the optimization algorithm.

For a total budget of $50,000, the complete enumeration method took over 68 hours for the fully flexible schedule (using increments of 0.0001 for each of the s_i). We did not even attempt to obtain the fully flexible optimal schedule using complete enumeration for a budget of $100,000, as the projected calculation time is 174 days. For the fixed cost per website schedule, complete enumeration took around 21 minutes, but this increases substantially to 22 hours when the campaign budget goes to $100,000 (as the number of feasible solutions increases by a factor of more than 60). For a budget of $50,000, the complete enumeration optimal schedule is exactly the same as for our model-based method, whereas for a budget of $100,000, the complete enumeration schedule is close to ours and has a reach higher by only 0.1 percentage points. Because our method is very much faster, taking no longer than 0.04 seconds, and yet produces the same or similar optimal schedules, it is apparent that our method has some advantages.

As a final test for our proposed method we undertook a validation check by comparing model-based reach estimates to those derived empirically across two time periods, March for estimation and June for the holdout period. Moreover, we obtained the optimal schedule and reach values in the holdout period. Section A5 of the electronic companion gives the details and shows that the optimal schedules and reach values across the two months are either the same or similar, thereby demonstrating reasonable validity for our method.

6. Discussion and Conclusion

The Internet presents challenges for online media planners that are not prevalent in traditional media. The main difference is the rate of delivery of advertisements, which is determined by the visit frequency to a site by a Web user. Although Web publishers have the ability to limit the number of ads delivered by frequency capping, this is not without its problems, because the same person can access a website from different computers. The purchase of online advertising is also different from other mass media, and so the budget function needs to be correspondingly suited to the new online arena. Like exposure distribution models, media selection methods for traditional media are not applicable to online media.

Instead, our method is to use a multivariate generalization of the NBD as an exposure distribution model that is specifically tailored to the Internet and is also adaptable to online media selection, where a share of impressions, a fixed number of impressions, or a purchase covering a fixed time period is required. That is, the MNBD is not only a good prediction model for Internet audience exposure, but it can also accommodate several different online advertising purchase methods. The downstream nonlinear optimization problem can be solved with a standard algorithm in a fraction of a second, and the resulting optimal solution is either the same or very similar to that obtained by a complete enumeration. However, the complete enumeration takes an order of magnitude longer, at best 21 minutes and at worst 68 hours. Moreover, in prediction validity tests, our method derives optimal schedules that are identical to those obtained in future time periods for moderate budget levels and are very close to the maximum achievable reach for larger budgets where there are many more feasible solutions. Therefore, the method we develop for optimal online media selection appears to have some merit.

Our findings also have some implications for media planners wishing to use the Internet. First, an obvious point is that reach usually increases as the number of websites in the campaign increases. Although this is well known for traditional media vehicles, current practice in online media purchases is for fewer sites (Klaassen 2007).

Second, where possible, online media planners should still attempt to buy a set number of impressions in a fixed time period, where the number of impressions is determined by the media planner, not the Web publisher. We call this fully flexible scheduling. A common method of online ad purchasing typically requires buyers to purchase impressions for a fixed period (such as a week or month) at a fixed rate, and then the site's CPM determines the number of impressions that will be delivered. This fixed cost per website arrangement gives the buyer less flexibility and always delivers lower overall reach.

Third, frequency capping always increases the reach, but another, less obvious way of increasing reach is to lengthen the campaign duration. Indeed, the combination of frequency capping and extending the campaign from one to four weeks increases the reach by almost 50% *for the same budget* in our example.

Our method can also be used to maximize average frequency and effective reach. Other aspects of online media scheduling, such as the timing and placement of ads within a website have already been considered by Adler et al. (2002), Kumar et al. (2007), and Nakamura and Abe (2005). Furthermore, targeting to specific demographic segments has not been demonstrated here, but it is not difficult to implement, as it requires that only a subset of the full data be obtained and the model refitted. We leave it to future research to embellish on these facets of optimal online media scheduling.

7. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mktsci.pubs.informs.org/.

Acknowledgments

The authors thank KoreanClick for providing the data. The second author's work was supported by a Korea University research grant, and the third author's work was partially supported by the HEC Foundation.

References

Aaker, D. A. 1975. ADMOD: An advertising decision model. *J. Marketing Res.* **12**(1) 37–45.

Adler, M., P. B. Gibbons, Y. Matias. 2002. Scheduling space-sharing for Internet advertising. *J. Scheduling* **5**(2) 103–119.

Bass, F. M., R. T. Lonsdale. 1966. An exploration of linear programming in media selection. *J. Marketing Res.* **3**(2) 179–188.

Bhat, S., M. Bevans, S. Sengupta. 2002. Measuring users' Web activity to evaluate and enhance advertising effectiveness. *J. Advertising* **31**(3) 97–106.

Chandler-Pepelnjak, J., Y.-B. Song. 2004. Optimal frequency: The impact of frequency on conversion rates, atlas digital marketing technologies. Retrieved March 5, 2007, http://www.atlassolutions.com/uploadedFiles/Atlas/Atlas_Institute/Published_Content/OptFrequency.pdf.

Coffey, S. 2001. Internet audience measurement: A practitioner's view. *J. Interactive Advertising* **1**(2) Article 9. http://www.jiad.org/article9.

comScore. 2007. Cookie-based counting overstates size of web site audiences. Press release (April 16). Retrieved May 1, http://www.comscore.com/press/release.asp?press=1389.

Danaher, P. J. 1991. Optimizing response functions of media exposure distributions. *J. Oper. Res. Soc.* **42**(7) 537–542.

Danaher, P. J. 1992. Some statistical modeling problems in the advertising industry: A look at media exposure distributions. *Amer. Statistician* **46**(4) 254–260.

Danaher, P. J. 2007. Modeling page views across multiple websites with an application to Internet reach and frequency prediction. *Marketing Sci.* **26**(3) 422–437.

Danaher, P. J., B. G. S. Hardie. 2005. Bacon with your eggs? Applications of a new bivariate beta-binomial distribution. *Amer. Statistician* **59**(4) 282–286.

Ephron, E. 1998. Optimizers and media planning. *J. Advertising Res.* **38**(4) 47–56.

Huang, C.-Y., C.-S. Lin. 2006. Modeling the audience's banner ad exposure for internet advertising planning. *J. Advertising* **35**(2) 123–136.

IAB. 2002. IAB Internet advertising revenue report. (June), http://www.iab.net/media/file/resources_adrevenue_pdf_IAB_PWC_2001Q4.pdf.

IAB. 2005. IAB Internet advertising revenue report. (April 27), http://www.iab.net/media/file/resources_adrevenue_pdf_IAB_PwC_2004_full.pdf.

IAB. 2006. IAB Internet advertising revenue report. (April 20), http://www . iab . net / media / file/resources_adrevenue_pdf_IAB_PwC_2005.pdf.

IAB. 2007a. Glossary of interactive advertising terms. http://www.iab.net/resources/glossary.asp#i.

IAB. 2007b. IAB Internet advertising revenue report. (May 23), http://www.iab.net/media/file/resources_adrevenue_pdf_IAB_PwC_2006_Final.pdf.

IAB. 2008. IAB Internet Advertising Revenue Report (May 23), http://www.iab.net/media/file/IAB_PwC_2007_full_year.pdf.

IMSL. 1997. IMSL MATH/LIBRARY user's manual. Visual Analytics, Houston.

Klaassen, A. 2007. Economics 101: Web giants rule "democratized" medium. *Advertising Age* **78**(15, April 9) 1–45.

Krugman, H. E. 1972. Why three exposures may be enough. *J. Advertising Res.* **12**(6) 11–14.

Kumar, S., M. L. Dawande, V. Mookerjee. 2007. Optimal scheduling and placement of Internet banner advertisements. *IEEE Trans. Knowledge Data Engrg.* **19**(11) 1571–1584.

LemonAd. 2002. Online Advertising Results: Year 2002 in Europe. Report, NetCrawling Deutschland GmbH, Hamburg, Germany.

Lilien, G. L., K. Philip, K. S. Moorthy. 1992. *Marketing Models*. Prentice Hall, Englewood Cliffs, NJ.

Little, J. D. C., L. M. Lodish. 1969. A media planning calculus. *Oper. Res.* **17**(1) 1–35.

Manchanda, P., J.-P. Dubé, K. Y. Goh, P. K. Chintagunta. 2006. The effect of banner advertising on Internet purchasing. *J. Marketing Res.* **43**(1) 98–108.

Meskauskas, J. 2003. Reach and frequency—Back in the spotlight, iMedia connection. (November 5), http://www.imediaconnection.com/content/2166.imc.

Nakamura, A., N. Abe. 2005. Improvements to the linear programming based scheduling of Web advertisements. *Electronic Commerce Res.* **5**(1) 75–98.

Naples, M. J. 1979. *Effective Frequency: The Relationship Between Frequency and Advertising Effectiveness*. Association of National Advertisers, New York.

Novak, T. P., D. L. Hoffman. 1997. New metrics for new media: Toward the development of Web measurement standards. *World Wide Web J.* **2**(1) 213–246.

Park, Y.-H., P. S. Fader. 2004. Modeling browsing behavior at multiple websites. *Marketing Sci.* **23**(3) 280–303.

Rossiter, J. R., P. J. Danaher. 1998. *Advanced Media Planning*. Kluwer Academic Publishers, Norwell, MA.

Rust, R. T. 1985. Selecting network television advertising schedules. *J. Bus. Res.* **13**(6) 483–494.

Rust, R. T. 1986. *Advertising Media Models: A Practical Guide*. Lexington Books, Lexington, MA.

Ryan, M. 2001. Dot coms in the driver's seat. AdRelevance Intelligence Special Report. Retrieved February 24, 2009, http://www.imagedomainweb.com/adrelevance/Reportb.htm.

Sarmanov, O. V. 1966. Generalized normal correlations and two-dimensional fréchet classes. *Doklady* (*Soviet Math.*) **168** 596–599.

Schittkowski, K. 1983. On the convergence of a sequential quadratic programming method with an augmented Lagrangian line search function. *Mathematik Operationsforschung Statistik, Ser. Optim.* **14** 197–216.

Seda, C. 2004. *Search Engine Advertising: Buying Your Way to the Top to Increase Sales*. New Riders Publishing, Indianapolis.

Sissors, J. Z., R. B. Baron. 2002. *Advertising Media Planning*, 6th ed. McGraw-Hill, Chicago.

Smith, D. L. 2003. Online reach and frequency: An update. MSN (April), http://www.mediasmith.com/storage/msn042003_1.pdf.

TNS Media Intelligence. 2008. TNS Media Intelligence reports U.S. advertising expenditures grew 0.2 percent in 2007. (March 25), Retrieved August 29, 2008, http://www.tns-mi.com/news/03252008.htm.

Zeff, R., B. Aronson. 1997. *Advertising on the Internet*. John Wiley & Sons, New York.