The Cost of Annoying Ads

DANIEL G. GOLDSTEIN
Microsoft Research
and
R. PRESTON MCAFEE
Microsoft Corporation
and
SIDDHARTH SURI
Microsoft Research

Display advertisements vary in the extent to which they annoy users. While publishers know the payment they receive to run annoying ads, little is known about the cost such ads incur due to user abandonment. We conducted two experiments to analyze ad features that relate to annoyingness and to put a monetary value on the cost of annoying ads. The first experiment asked users to rate and comment on a large number of ads taken from the Web. This allowed us to establish sets of annoying and innocuous ads for use in the second experiment, in which users were given the opportunity to categorize emails for a per-message wage and quit at any time. Participants were randomly assigned to one of three different pay rates and also randomly assigned to categorize the emails in the presence of no ads, annoying ads, or innocuous ads. Since each email categorization constituted an impression, this design, inspired by Toomim et al. (2011), allowed us to determine how much more one must pay a person to generate the same number of impressions in the presence of annoying ads compared to no ads or innocuous ads.

Categories and Subject Descriptors: J.4 [**Social and Behavioral Sciences**]: Economics

General Terms: Economics, Experimentation

Additional Key Words and Phrases: display, advertising, quality, compensating differential

1. INTRODUCTION

Display advertising is the prevalent way for publishers to monetize content on the Web. Publishers receive payment from advertisers for placing ads near their content or in their applications. Publishers are typically paid by the number of impressions they can deliver. Thus, they have an incentive to attract and retain users with valuable content, experiences, and applications, and have a disincentive to lose users due to annoyances.

Display ads vary in the extent to which they annoy users. Annoying ads are a source of tension for publishers since they both make money, through payments from advertisers, and cost money, through a decrease in page views due to users abandoning the site. This tension has led to conflict within publishing organizations

Authors' addresses: dgg@microsoft.com, preston@mcafee.cc, and suri@microsoft.com. This article is an adaptation of the following conference proceedings article and appears here with permission of the publishers: Goldstein, Daniel G., R. Preston McAfee, and Siddharth Suri (2013), "The Cost of Annoying Ads," Proceedings of the 22nd International World Wide Web Conference. This work was later expanded and presented in [Goldstein et al. 2014].

between salespeople, who have an incentive in the form of commission to sell any ads, and management, who are concerned with long-term growth of users and traffic. The continued long-term display of annoying ads may exert negative effects on the publisher, the user, and the advertiser, which we discuss in turn.

First, annoying ads can exert negative effects on publishers. Apart from the user abandonment effects we investigate in this paper, annoying ads might signal that the website, on which the ad is placed, lacks stability ("Why should I trust my email to a site that is so desperate for cash it accepts ads of such poor quality?"), reputability ("Why should I trust the objectivity of a site that is so in the pocket of advertisers it won't refuse any of them?"), or safety ("Why would I trust this publisher to protect me from phishing attacks, scams, malware, etc. if they are so indiscriminate about who they let advertise?").

Second, annoying ads can exert a negative impact on users. Ads with excessive animation can get in the way of the user consuming the publisher's content, undermining the very reason that brought them to the site.

Finally, annoying ads may harm the advertiser that created them. As will be shown, annoying ads are often characterized by exaggerated attempts to capture visual attention such as through fast-moving animation or bizarre imagery. While these manipulations do capture attention, they may also signal that the advertiser is desperate for business or low on resources, undermining the classical signal of quality that advertising is theorized to bring [Riley 2001]. Furthermore, experiments have shown that too much animation can result in lower ad recognition rates compared to ads with moderate or no animation [Yoo and Kim 2005; Burke et al. 2005]. In these ways, annoying ads may actually lower brand reputation and recall, two metrics advertisers typically strive to increase.

If annoying ads exhibit so many negative effects for publishers, users and advertisers, one may wonder why a publisher would run annoying ads at all. The answer may be that it is has been historically difficult to measure the monetary cost of annoying ads. The main contribution of this work is that we measure the compensating wage differential of annoying ads. That is, we measure how much more one must pay a user to do the same amount of work in the presence of annoying ads compared to innocuous ads or no ads. The compensating differential is important to measure because it captures some of the negative effects of advertising, which publishers need to heed as a lower bound when setting the price to run an ad.

Using two experiments, we compute the compensating differential for annoying ads. In the first experiment users rated a set of ads in terms of how annoying they found each ad. In the second experiment, we use those ads identified as more or less annoying, along with the recent methodological innovation of Toomim et al. [2011], to estimate the pay rate increase necessary to generate an equal number of page views in the presence of annoying ads, compared to innocuous ads or no ads. This estimate is the cost of annoying ads in our experiment. We chose categorizing emails as the task to proxy for using a publisher's site because users either implicitly or explicitly need to categorize their emails as spam or not spam in the presence of ads when using free web-based email services such as Yahoo! Mail, GMail, and Outlook.com.

2. RATING THE QUALITY OF ADS

We now describe the design and results of our first experiment, which served to identify sets of more and less annoying ads (henceforth "bad ads" and "good ads" for brevity) for use in the second experiment.

2.1 Method

The goal of this experiment is to rank a set of actual display ads in terms of annoyingness. After previewing all the ads, users were shown each ad individually, in random order, and asked to rate each ad on a 5-point scale with the following levels: 1) Much less annoying than the average ad in this experiment, 2) A bit less annoying than the average ad in this experiment, 3) Average for this experiment, 4) A bit more annoying than the average ad in this experiment, and 5) Much more annoying than the average ad in this experiment. If we view attributes of an ad as residing in a multidimensional space, the average ratings indicate how users project that multidimensional space onto a one-dimensional annoyingness scale. The 10 most and least annoying ads serve as the sets of "bad" and "good" ads in the next experiment.

3. MEASURING THE COST OF ADS

We use the method of Toomim et al. [2011] along with the sets of "bad" and "good ads" to measure the cost of annoying ads.

3.1 Method

The participants were 1223 Mechanical Turk workers who participated for a base pay of 25 cents and a bonus. Upon accepting the task, participants were randomly assigned to one of nine conditions: three pay conditions and three ad conditions. The pay conditions offered a bonus of one, two, or three cents per five emails classified (i.e., .2, .4, or .6 cents per email), and the ad conditions varied whether "bad ads", "good ads", or no ads were displayed in the margin as the task was completed. A chi-squared test found no significant difference in the number of participants beginning work across the nine conditions.

In all conditions, the task consisted of classifying the content of emails as "spam", "personal", "work" or "e-commerce" related. Emails were drawn from the public-domain Enron email dataset[1] with one email presented per page, along with accompanying ads, if any. In the "bad ads" condition, two ads randomly drawn from the 10 most annoying ads in our first experiment were displayed in the margins around the email being classified. The "good ads" condition was the same, except the ads were drawn from the 10 least annoying ads. In both conditions, ads were drawn randomly from their respective pools with each page load, and the URLs for the ads were such that ad blocking software would not filter them out. The "no ads" condition simply had whitespace in the margin. The footer included two buttons: one allowing them to submit and rate another email, and a second allowing them to stop categorizing and collect their payment.

[1] `http://www.cs.cmu.edu/~enron/` Identifying information such as email addresses, phone numbers and the name "Enron" were removed.

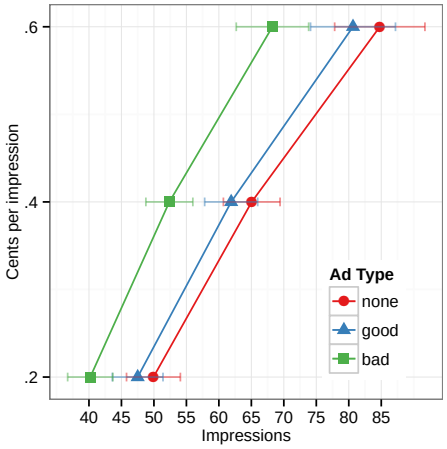

Fig. 1. Estimated impressions per condition based on the negative binomial model. Error bars are ± 1 standard error.

3.2 Results

Let an impression be one participant viewing one email (and its accompanying ads, if any), regardless of whether the participant classifies the email or quits before classifying it. The overall distribution of impressions per person is skewed with a mean of 61, a median of 25 and first and third quartiles of 6 and 57. Being bounded by 1 from below and effectively unbounded from above (only two participants reached the upper limit), these impressions constitute count data. In particular, they are overdispersed count data relative to the Poisson (observed variance / theoretical Poisson data variance is 228.7, $p < .0001$) and thus well suited to a negative binomial generalized linear model (GLM) [Venables and Ripley 2002]. The effects of the conditions on raw impressions are most easily seen in Figure 1, which also makes clear that the difference in impressions between the "good ads" and "no ads" conditions is not significant. Relative to a baseline of "bad ads", both the "good ads" condition and the no ads condition led to substantially more impressions (19% and 25% more impressions, respectively).

The model expressed in Figure 1 can be used to estimate the compensating differential of annoying ads in this experiment. Since the curves are slightly nonlinear, a range of compensating differentials could be calculated across the pay rate and ad conditions. To get a simple, single approximation we use the middle, "good ads" condition to estimate the effect of pay raises. We take the average of the .2 to .4 and .4 to .6 cent differences, giving an estimated increase of 16.58 impressions resulting from a .2 cent per impression pay raise. When summarizing the effect of ad quality, we use the number of impressions at the .4 cent pay rate. Moving from "bad ads" to no ads, impressions increase by 12.68. The pay raise required to achieve a 12.68 impression increase is .153 cents per impression ($= .2 * 12.68/16.58$) or \$1.53 CPM (cost per thousand impressions). That is, in this experiment, a participant in the "bad ads" condition would need to be paid an additional \$1.53 per thousand

impressions to generate as many impressions as a person in the condition without ads. Similarly, moving from the "bad ads" condition to the "good ads" condition resulted in an additional 9.52 impressions per person. It would require a pay raise of .115 cents per impression (= .2∗9.52/16.58) to generate 9.52 additional impressions, meaning that people in the "bad ads" condition would need to be paid an additional $1.15 CPM to generate as many impressions as in the "good ads" condition.

4. CONCLUSION

The main result of this paper is that annoying ads lead to site abandonment and thus fewer impressions than good ads or no ads. In what might be seen as good news for publishers, good ads and no ads led to roughly equal numbers of impressions.

We calculated the compensating wage differential in our experiment of bad ads to no ads to be $1.53 CPM, and bad ads to good ads to be $1.15. Some care must be taken in interpreting these numbers. While we picked a task—classifying emails—that should be familiar and common for most internet users, this task may not be representative of other internet tasks like reading news stories or searching for products to purchase. Abandonment rates may differ with different tasks and the effects of advertising may vary as well. While virtually every web service features competition, the switching costs vary from very low in consuming news to relatively high in changing email services. Because our users on Mechanical Turk have an outside option of working on an alternative task, we expect our results to be most applicable to situations involving lower switching costs. Nevertheless, we expect that our finding that annoying ads cost the user at least $1 CPM over more pleasant ads will be obtained in some other environments.

For these reasons, we suggest further studies be done on Mechanical Turk, as field experiments, and in laboratories to measure this differential on similar and different tasks. If studies across various domains with a variety of tasks and outside options arrive at similar differentials, more credence can be placed on these numbers. We view this work as a first step in this direction. If future work arrives at similar estimates across a variety of publishers, such estimates could serve as a useful lower bound for what a publisher should charge to run these ads. Moreover, it will be valuable to use the compensating differentials approach to price the various bad aspects of ads, including animation and poor aesthetics.

This work also suggests a variety of policy recommendations. Most directly, the $1 CPM user cost of bad ads has practical consequences for publishers, especially as bad ads often command lower CPMs. It is a reason that publishers should insist on a substantial premium for annoying advertisements. Moreover, a publisher could randomize which users see which ads and track both time spent on the page and the frequency with which users return to the site. This type of experimentation would capture longer term effects of annoying ads than those studied here. Also, publishers could give users an option to close or replace an ad. A replacement event would allow the publisher to infer that a user would prefer a random ad over the ad currently shown. Advertisers with a high closure rate should be charged more since more annoying ads would be closed or replaced faster than less annoying ads. Ad replacement would help the user by removing the annoying ad and the publisher by making it possible to charge for two impressions.

REFERENCES

Burke, M., Hornof, A., Nilsen, E., and Gorman, N. 2005. High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Transactions on Computer-Human Interaction 12,* 4 (December), 423–445.

Goldstein, D. G., Suri, S., McAfee, R., Ekstrand-Abueg, M., and Diaz, F. 2014. The economic and cognitive costs of annoying display advertisements. *Journal of Marketing Research*. In Press.

Riley, J. G. 2001. Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature XXXIX*, 432–478.

Toomim, M., Kriplean, T., Pörtner, C., and Landay, J. A. 2011. Utility of human-computer interactions: Toward a science of preference measurement. In *Proceedings of CHI 2011: ACM Conference on Human Factors in Computing Systems.*

Venables, W. N. and Ripley, B. D. 2002. *Modern Applied Statistics with S.* Springer, New York.

Yoo, C. Y. and Kim, K. 2005. Processing of animation in online banner advertising: The roles of cognitive and emotional responses. *Journal of Interactive Marketing 19,* 4, 18–34.