A Case-Based Analysis of the Effect of Offline Media on Online Conversion Actions

Damir Vandic
vandic@ese.eur.nl

Didier Nibbering
didieren@gmail.com

Flavius Frasincar
frasincar@ese.eur.nl

Erasmus University Rotterdam
PO Box 1738, NL-3000
Rotterdam, the Netherlands

ABSTRACT

In this paper, we investigate how offline advertising, by means of TV and radio, influences online search engine advertisement. Our research is based on the search engine-driven conversion actions of a 2012 marketing campaign of the potato chips manufacturer Lays. In our analysis we use several models, including linear regression (linear model) and Support Vector Regression (non-linear model). Our results confirm that offline commercials have a positive effect on the number of conversion actions from online marketing campaigns. This effect is especially visible in the first 50 minutes after the advertisement broadcasting.

Categories and Subject Descriptors

H.4.2 [**Information Systems Applications**]: Types of Systems—*Decision support*; H.4.m [**Information Systems Applications**]: Miscellaneous

Keywords

Offline Media, Marketing Campaign, Online Conversion

1. INTRODUCTION

The last few years online advertising has grown rapidly. One of the reasons is that online advertising gives the possibility of targeting specialized audience segments [8]. However, one of the main side effects is the ease of tracking the results of an online campaign. Besides the advantages of online media, we know that no website can compete with the sheer size and usage volume of the TV audience. An average American family watches 60 hours of television a week [1]. We also know that offline advertising, like television commercials, increases website visits by increasing the consumer awareness [7].

In the literature, we can find many studies that investigate the sales-advertising relationship [2, 6, 7]. In particular, the authors [7] demonstrate that offline advertising could increase website visitation. They empirically determine the factors that drive traffic in the Web space. Offline advertising is one of these factors. It appears to increase website visitation through its significant influence on consumer awareness, while online advertising directly leads to increased website traffic. However, the relationship between the offline part of an advertising campaign and the online website visitation has never been confirmed using a large real-life data set. In this research, we use a unique large dataset that contains log information on website activity and offline and online advertisement campaigns. We investigate how offline advertisement influences the number of visitors that reaches the main portal site through online search engine advertisements. In this context, a conversion action is defined as a Web site visit at a particular point in time and with a particular goal.

WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

2. THE ADVERTISING CAMPAIGN

The focus of this paper is a campaign from the potato chips manufacturer Lays, which consists of an online and offline part. Lays started in 2010 with an advertising campaign called "Maak de Smaak" (Dutch for "Make the flavor") that aimed to increase brand preference by involving consumers in the product development. Dutch consumers were asked to (1) devise a potato chips flavor for the new Lays Limited Edition potato chips, (2) to purchase the trial packs of the final flavors, and (3) to vote on their favorite flavor. For tasks (1) and (2), Lays launched a Web site that served as the main portal of the campaign.

The data set that we use consists of 600,000 so-called "unique conversion paths". These unique conversion paths contain ordered nodes, where each node represents an online action (e.g., a Web site visit). A unique conversion path corresponds to one or more conversion actions. For example, a user might go to the home page and submit a new flavor in the competition. In this case there are two conversion actions: (1) the Web site visit, and (2) the submission of a new flavor, which is a more valued conversion action. Complementary to this data, we also have detailed data of the level of advertisement, both for TV and radio.

From the data set, we extracted several important variables. We take the 'number of Search Engine Advertisement (SEA) conversion actions' as the dependent variable. These conversion actions represent visits by users that arrived through an search engine advertisement. As first explanatory variable we take 'time', as it is plausible that people are more likely to use the Internet at certain times of the day. We use two independent variables for the number of radio and television commercials. Furthermore, 'size of the reached radio audience' is also used as an independent variable. Last, in order to estimate the individual effects of different channels, we employ dummy variables for each TV channel. We convert the data set to a time series data set

where a record represents a slice of 5 minutes. For example, if y_t denotes the dependent variable, then y_0 represents the number of SEA conversion actions in the first 5 minutes, y_1 in the second 5 minutes, etc.

3. RESULTS

We start our analysis with a simple econometric model: the Ordinary Least Squares (OLS) linear regression model. We estimate this model on a training set, and then we use the estimated model to predict the number of SEA conversions on a test set. In order to determine whether offline advertisement has any significant effect on the number of online SEA conversion actions, we compare the fit of two models, i.e., one that contains offline advertisement levels as independent variable(s) and one that does not. For this comparison, we employ the Diebold-Mariano test (DM-test) [4] and select the optimal model specification for both models. We found that the most optimal specification includes 10 lags for each time-bound variable (e.g., the number of radio advertisement broadcasts).

The results from the DM-test show that the model that contains offline advertisement levels as independent variable(s), has a significantly lower mean absolute error (MAE) at a significance level of 5% ($p = 0.04$). From this we can conclude that information about offline advertisements improves the prediction of online SEA conversion actions. Furthermore, we can confirm that, based on the Mincer-Zarnowitz regression test [5], the model with offline advertisement variables is more efficient (i.e., the estimations are less biased) than the model that excludes the offline advertisement variables.

We also perform the same experiment with a Support Vector Regression (SVR) model. This model is known for its excellent performance in various fields [3]. In this case, we found that, at a 10% significance level, the model that includes offline advertisement information has a significantly lower MAE than the model that excludes offline advertisement information. In other words, we can draw the same conclusion as with the OLS model, only this time with less certainty. The Mincer-Zarnowitz regression test showed that the two models differ slightly in efficiency, i.e., the level of bias in the estimation, with the model that excludes the offline advertisement information being slightly more efficient.

Figure 1 shows an overview of the effect of the lag size in the OLS and SVR models. We can see that in the beginning adding more lag elements decreases the error. However, for both models, the error increases after a certain number of lag elements. For the OLS model this behavior is more visible as the MAE of the SVR remains fairly stable across the whole lag range. Figure 1 confirms also what our statistical tests have shown, i.e., that the SVR regression model is clearly superior compared to the OLS regression model.

4. CONCLUSIONS

In this paper, we analyzed the effect of offline media on online conversion actions. Using a large marketing campaign data set, we estimated OLS-models and SVR-models that predict the number of search engine advertisement conversion actions. The models showed that offline media significantly affects the online conversion actions. In addition, we found that the influence is the highest in the first 50 minutes after the broadcasting of offline media.

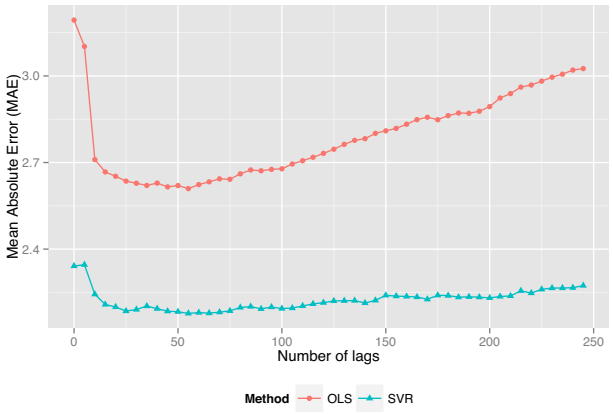

Figure 1: An overview of the Mean Absolute Error (MAE) for different lag values.

As future work we would like to investigate the role of the cost of offline media resources in the employed models. When we utilize the effects of offline media we have found in this paper with the costs of the offline media resources, we can provide the most cost-efficient allocation of the offline media resources. Another aspect that needs to be investigated is the consideration of information about the display advertising. In this way it is possible to devise a model that estimates the optimal combination of offline and online media in order to achieve the best cost-efficient advertisement result.

Acknowledgment

The authors are partially sponsored by the NWO Mosaic project 017.007.142: Semantic Web Enhanced Product Search (SWEPS).

5. REFERENCES

[1] S. Baran. *Introduction to Mass Communication: Media Literacy and Culture*. McGraw-Hill, 2010.

[2] D. Clarke. Econometric measurement of the duration of advertising effect on sales. *Journal of Marketing Research*, pages 345–357, 1976.

[3] R. Collobert and S. Bengio. Svmtorch: Support vector machines for large-scale regression problems. *The Journal of Machine Learning Research*, 1:143–160, 2001.

[4] F. Diebold and R. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263, 1995.

[5] J. A. Mincer and V. Zarnowitz. The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, pages 1–46. NBER, 1969.

[6] V. Rao. Alternative econometric models of sales-advertising relationships. *Journal of Marketing Research*, pages 177–181, 1972.

[7] R. Winer and J. Ilfeld. Generating website traffic. *Journal of Advertising Research*, 42(5):49–61, 2002.

[8] R. Zeff and B. Aronson. *Advertising on the Internet*. John Wiley & Sons, Inc., 1999.