European Journal of Operational Research 250 (2016) 579–589

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision support

Investigating the effects of mailing variables and endogeneity on mailing decisions

Nadine Schröder*, Harald Hruschka[1]

Department of Marketing, University of Regensburg, Universitätsstraße 31, D-93040 Regensburg, Germany

ARTICLE INFO

Article history:
Received 24 April 2013
Accepted 22 September 2015
Available online 9 October 2015

Keywords:
(D) OR in marketing
Specification of mailings
Mixture of Dirichlet Processes
Endogeneity
Dynamic optimization

ABSTRACT

Determining the optimal amount of mailings being sent to customers is crucial. However, this decision depends on various aspects. First, it is important to specify relevant mailing variables. By distinguishing different types of mailings and considering their sizes, we set our study apart from the majority of existing studies. To deal with unobserved heterogeneity we estimate a Mixture of Dirichlet Processes (MDP) whose components are Tobit-2 models. A policy function approach is used to take endogeneity into account. We investigate whether and how consideration of endogeneity leads to different managerial implications. To this end, we determine mailings by dynamic optimization for three individual customers which are prototypical for the segments discovered by the MDP model. We find out that mailings should be avoided altogether more frequently for all three customer types according to the model which takes endogeneity into account.

1. Introduction

In order to induce purchases, mail order companies rely on mailings that they target to their customers. These mailings may differ from one another w.r.t. the season or clearings (Gönül, Kim, & Shi, 2000) or types of mailings such as main catalogs and regular mailings (Holland, 1993, chap. 9). Past studies suggest that the use of mailing variables is crucial when modeling purchase probabilities (e.g., Naik & Piersma, 2002). Gönül and Shi (1998) find that mailing variables have a positive influence on purchase probability. However, research also shows that mailings do not generally lead to positive effects on purchases or sales. In fact, if a company sends too many mailings, customers may become irritated and abstain from further purchases (e.g., Van Diepen, Donkers, & Franses, 2009).

If such situations arise, mailings are said to display saturation effects. These are generally characterized by high levels of mailings at which further increases are not accompanied by more purchases or higher sales. As a progression of this effect, super-saturation is possible, i.e., purchases or sales decrease if mailings are increased (Hanssens, Parsons, and Schultz, 2001, chap. 3).

We make use of different types of mailing variables in order to model the influence of these mailings and express (super-)saturation effects. Besides, we also include a model which takes endogeneity into account. We show that not controlling for endogeneity may lead to different optimal targeting suggestions.

Table 1 contains studies that model saturation effects. We do not only refer to mail order companies but to charities as well. To keep terminology consistent, we refer to solicitations from mail order companies and charities as mailings. Consequently, we call the response of a customer to a mailing purchase (with corresponding sales) no matter which industry is involved. Based on the specification of mailing variables, these studies can be categorized into three groups.

The first group consists of publications (Baumgartner & Hruschka, 2005; Hruschka, 2010; Hruschka, Baumgartner, & Semmler, 2003) which consider number of mailings (NOM) as one of the predictor variables and assumes that its effect is static. All three studies transform the respective variables in order to assess saturation effects. S-shaped response functions (Baumgartner & Hruschka, 2005; Hruschka et al., 2003) show positive but degressive effects of mailings on purchase behavior. Only Hruschka (2010) corrects for endogeneity and finds biased estimates when he compares to models that do not account for endogeneity. All three publications use cross-section data, which means that they have for each customer only one value for the dependent variables. Therefore, they cannot perform dynamic optimization and have to use a static optimization approach.

In the second group of publications NOM is one of the predictors, but timing of mailings (Naik & Piersma, 2002; Van Diepen et al., 2009) or customer behavior (Gönül et al., 2000; Rhee & McIntyre, 2008) are taken into consideration as well. Whereas Gönül et al. (2000) calculate NOM depending on the previous purchase Rhee and McIntyre (2008) differentiate their mailing variable further. In this approach,

* Corresponding author. Tel.: +49 9419432274; fax: +49 9419432828.
E-mail addresses: nadine.schroeder@wiwi.uni-regensburg.de (N. Schröder), harald.hruschka@wiwi.uni-regensburg.de (H. Hruschka).
[1] Tel.: +49 9419432277.

http://dx.doi.org/10.1016/j.ejor.2015.09.046
0377-2217/

Table 1
Different studies w.r.t. mailing specification and saturation.

	Mailing specification	Saturation expressed with	Industry	Unobserved heterogeneity	Endogeneity	Optimization
Hruschka et al. (2003)	Number of mailings (NOM)	Semi-log and logistic functions	Mail order	Finite mixture models	Not accounted for	Static
Baumgartner and Hruschka (2005)	NOM	Cubic smoothing splines	Mail order	Not accounted for	Not accounted for	Static
Hruschka (2010)	NOM	Log function	Mail order	Mixture of Dirichlet Processes (MDP)	Policy function and Instrumental variables	Static
Gönül et al. (2000)	NOM w.r.t. past behavior	Box–Cox hazard function	Mail order	Finite mixture models	Instrumental variables	Static
Naik and Piersma (2002)	NOM w.r.t. dynamics	Goodwill variable	Charity	Continuous mixture models	Not accounted for	No
Rhee and McIntyre (2008)	Distinction between NOM w.r.t. past behavior	Gamma distribution	Nonprofit	Continuous mixture models	Latent variable	No
Van Diepen et al. (2009)	NOM w.r.t. dynamics	Quadratic terms	Charity	Continuous mixture models	Policy function	No
Campbell et al. (2001)	Mailing content	Similarity and timing matrices	Mail order	Not accounted for	Not accounted for	Dynamic
This paper	Type and size of mailings partly w.r.t. dynamics	Quadratic terms	Mail order	MDP	Policy function	Dynamic

two versions of NOM are calculated. One depends on recent the other on prior mailings. Similar to the first group, saturation is examined by different transformations of mailing variables. Except for Naik and Piersma (2002) all authors correct for endogeneity. In contrast to Gönül et al. (2000), Van Diepen et al. (2009) find endogeneity biases. Rhee and McIntyre (2008) do not report whether such a bias exists. Of these studies only Gönül et al. employ optimization by using a static approach.

The paper of Campbell et al. (2001) is the only one which belongs to the third group which is characterized by also taking into account the content of mailings. To make saturation evident, these authors use similarity matrices based on pairwise comparisons between two mailings w.r.t. common mailing contents. They take the revenue of products that show up in both mailings as indicator for the greatest possible saturation if these mailings were mailed on the same day. Besides, they also control for timing of mailings. As a result, they calculate a saturation matrix by combining the similarity and timing matrices. The authors do not investigate endogeneity. In terms of optimization a dynamic approach is employed.

Our approach differs in several ways from related extant publications. Most previous studies consider only one (binary) mailing variable, which indicates whether a customer receives a mailing or not. We on the other hand include three different mailing variables. One refers to the type of catalogs, the remaining two take care of the size of mailings. Please note that mailings often consist of several catalogs (to target customers more individually) and hence different amount of pages. To measure size of an actual mailing we use a variable defined as the number of remaining pages, i.e., the number of pages not considering the main catalog. Besides, we use a second size variable which equals the total discounted number of previous pages across catalogs sent to a customer.

These extended mailing variables allow to derive optimal mailing rules specific to individual customers which not only tell whether or not a mailing should be sent to a customer, but also how large such a mailing should be in terms of total pages. To this end, we employ dynamic optimization because our model comprises dynamic effects by means of stock variables which are based on the size and sales of previous mailings received by a customer. To the best of our knowledge our study is the first in which a dynamic problem is solved which also deals with the optimal size of mailings and is not limited to the assignment of catalogs to individual customers. The use of policy functions further sets our study apart from most other publications on direct mailings which neglect the concept of endogeneity altogether.

Unobserved heterogeneity is of vital importance as can be seen from most of the studies which take it into account. Ignoring unobserved heterogeneity may lead to spurious state dependence and biased parameter estimates (Popkowski Leszczyc & Bass, 1998). Whereas the majority of researchers follow continuous mixture approaches, we use a Mixture of Dirichlet Processes (MDP) model. The MDP is capable to reproduce multimodal and skewed distributions and determines the number of latent segments alongside the estimation process (e.g., Hruschka, 2010). Unlike, e.g., Van Diepen et al. (2009) we apply the concept of unobserved heterogeneity to all our model parameters.

2. Modeling

2.1. Tobit-2 model

We employ a Tobit-2 model (Amemiya, 1985, chap. 10). Research done by Levin and Zahavi (1998) shows that such a model is preferable over regression or logit models which are widely used to analyze the effect of direct mailing (for a general overview of models used in direct marketing please refer to Bose & Chen (2009)). Tobit-2 models deal with two decisions of customers simultaneously, i.e., the response to a mailing and the amount of sales a customer spends. If customer i responds to a mailing τ the first dependent variable, response $R_{i\tau}$, is equal to one. Only in this case the second dependent variable, the amount of sales $A_{i\tau}$, assumes a positive value as can be seen from Eqs. (1) and (2) (Donkers, Paap, Jonker, & Franses, 2006; Van Diepen et al., 2009):

$$R_{i\tau} = \begin{cases} 1 & \text{if } R^*_{i\tau} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$A_{i\tau} = \begin{cases} A^*_{i\tau} & \text{if } R^*_{i\tau} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The two latent variables $R^*_{i\tau}$ and $A^*_{i\tau}$ are specified as linear combinations of explanatory variables contained in vectors $x_{Ri\tau}$ and $x_{Ai\tau}$ to which the respective error term $(\epsilon_{Ri\tau}, \epsilon_{Ai\tau})$ is added:

$$R^*_{i\tau} = x'_{Ri\tau}\beta_R + \epsilon_{Ri\tau} \tag{3}$$