



Interfaces with Other Disciplines

Considering endogeneity for optimal catalog allocation in direct marketing

Harald Hruschka*

Department of Marketing, University of Regensburg, Universitätsstrasse 31, D-93053 Regensburg, Germany

ARTICLE INFO

Article history:

Received 6 January 2008
Accepted 20 January 2010
Available online 25 January 2010

Keywords:

(D) Marketing
Direct marketing
Endogeneity
Hierarchical Bayesian modeling

ABSTRACT

The majority of catalog allocation models using historical data ignore endogeneity of past catalog decisions. We investigate two alternative approaches which either impose a relationship between the number of catalogs allocated to a customer and customer-specific coefficients of the sales response function or use instrumental variables. Heterogeneity across customers is modeled by cluster effects following a nonparametric distribution derived from a Dirichlet process prior. Models are estimated by Markov chain Monte Carlo simulation methods and evaluated by cross-validation predictive densities. Models which consider endogeneity imply much lower effects for sending a higher number of catalogs. These models also lead to optimal allocations which differ strongly from optimal allocations obtained for models which ignore endogeneity. Higher values of both posterior model probabilities and model average profits suggest to allocate catalogs based on the instrumental variables approach.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

It is well known that firms decide on values of marketing variables (e.g., price, advertising, sales force efforts, etc.) based on the (expected) response of customers (e.g. Gupta and Zeithaml, 2006). Under such circumstances marketing variables are said to be endogenous, i.e. they are not independent of disturbances of market response functions (Chintagunta et al., 2006a). Decision makers in direct marketing, e.g., allocate a higher number of catalogs to customers for whom they expect higher sales or a stronger response.

Still, the majority of publications estimate market response functions by methods which ignore endogeneity and therefore could provide biased estimates of the effects of marketing variables (Ebbes et al., 2009). Table 1 gives an overview of papers which have addressed endogeneity in marketing. Results are inconclusive w.r.t. price endogeneity if one limits attention to papers which account for heterogeneity of coefficients across households. Chintagunta et al. (2005) provide evidence for endogeneity bias, whereas in Yang et al. (2003) the endogeneity bias disappears if heterogeneity is accounted for. Chintagunta et al. (2006b) discover that advertising effects are underestimated in one of the two product categories considered. This contradicts Manchanda et al. (2004) who find that the effect of sales calls is overestimated if endogeneity is ignored. One has to bear in mind that the approach of Chintagunta et al. (2006b) is limited by using aggregate data and ignoring heterogeneity. Therefore this approach runs the risk to (partially) mistake

heterogeneity for endogeneity. Gönül et al. (2000), Rhee and Russell (2009), Donkers et al. (2006) and Van Diepen et al. (2009) investigate endogeneity of catalog allocation in direct marketing and simultaneously account for heterogeneity. Two of these papers (Rhee and Russell, 2009; Van Diepen et al., 2009) provide evidence for endogeneity biases.

Based on this literature review several research gaps can be identified. The results of previous studies are inconclusive, both w.r.t. the extent of endogeneity bias and its direction. This fact together with the scarcity of studies in direct marketing justify to carry out a study dealing with the problem of endogeneity in this marketing area. Previous studies demonstrate that ignoring heterogeneity runs the risk of finding an endogeneity bias even in its absence, because as a rule ignoring either heterogeneity or endogeneity causes model parameters to be biased in the same direction. Therefore both endogeneity and heterogeneity should be accounted for simultaneously. Moreover, previous studies are typically confined to only one modeling approach to account for endogeneity.

Most of the studies which consider heterogeneity (i.e., five of seven studies) adopt a parametric approach. Their authors assume that individual coefficients follow a multivariate normal distribution which precludes multimodality and skewness. The nonparametric approach of our study overcomes this limitation. The nonparametric approach is able to show, for example, that both low and high response coefficients are important which, of course, is not possible if the unimodal multivariate normal distribution is assumed.

We now give more details on the papers which account for endogeneity in a direct marketing context (see Table 1). The core

* Fax: +49 941 943 2828.

E-mail address: harald.hruschka@wiwi.uni-regensburg.de

Table 1
Studies on endogeneity in marketing.

Authors	Data level	Heterogeneity accounted for	Considered marketing instrument(s)	Endogeneity not accounted for causes model to
<i>Studies focusing on marketing instruments other than direct marketing</i>				
Chintagunta et al. (2005)	Individual	Yes	Price	Underestimate price effect
Chintagunta et al. (2006b)	Aggregate	No	Price, advertising	Underestimate price effect, underestimate advertising effect in one of two categories
Manchanda et al. (2004)	Individual	Yes	Sales calls	Overestimate effect of calls
Villas-Boas and Winer (1999)	Individual	No	Price	Underestimate price effect
Yang et al. (2003)	Individual	Yes	Price	
<i>Studies focusing on direct marketing</i>				
Donkers et al. (2006)	Individual	Yes	Catalogs	
Gönül et al. (2000)	Individual	Yes	Catalogs	
Rhee and Russell (2009)	Individual	Yes	Catalogs	Overestimate effect
Van Diepen et al. (2009)	Individual	Yes	Catalogs	Overestimate effect

of the model of Gönül et al. (2000) is a finite mixture of proportional hazard functions. To consider endogeneity these authors replace observed mailing by a mailing probability computed by a binary logit model with homogeneous coefficients. Their estimation does not show great differences between models with and without endogeneity. As Gönül et al. do not know the profit margin, they set it equal to the value at which the optimal solution equals the number of catalogs actually sent over a three-months tracking period. For this value of the profit margin Gönül et al. determine a mailing policy based on their model which increases total expected profit by about 6% relative to the actual mailing policy.

Donkers et al. (2006) consider endogeneity by adding a binary probit model for the mailing decision of a charity to a Tobit-2 model which deals with monetary response (amount of a donation). The whole model is of the finite mixture type with both probit and Tobit parts being segment-specific. Coefficients are estimated by maximizing the complete data likelihood which consists of the likelihood of the Tobit model and probabilities of (not) sending a mailing to an address. For a one year holdout data set donations do not differ much regardless whether mailing decisions are made based on models with or without endogeneity. The authors explain this result by the fact that mailing decisions made by the charity were almost random.

Rhee and Russell (2009) use binary probit models for both customer response (donate = yes or no) and mailing decision of a charity. Their probit models are of two types. The first type uses RFM variables. To consider endogeneity it includes residuals of first order autoregressive models as predictors. Probit models of the first type are estimated by maximum likelihood. Probit models of the second type are characterized by the fact that one of their predictors is customer response of the previous period. A customer-specific constant in the customer response model is also used as predictor in the mailing probit model. Rhee and Russell interpret this constant as latent propensity of a household to donate. As in this application a customer can only donate after she has received a mailing, the customer response model does not include a mailing decision variable as predictor. Coefficients of the second model type are assumed to be multivariate normally distributed and estimated by a Markov chain Monte Carlo (MCMC) algorithm. Pseudo R^2 values for a holdout data set indicate better performance of models of the second type.

In Van Diepen et al. (2009) a Tobit-2 model serves to model monetary response (amount of a donation to a charity). The authors assume that individual coefficients are multivariate normally distributed and use a MCMC method for estimation. To deal with endogeneity they follow the approach of Manchanda et al. (2004) by imposing a relationship between the decision model

marketing instrument considered (mailings) and coefficients of the monetary response function.

These studies on endogeneity in direct marketing suffer from the following shortcomings. They do not agree on the existence of an endogeneity bias. Rhee and Russell (2009) as well as Van Diepen et al. (2009) apply the widespread parametric approach based on the multivariate normal distribution to account for heterogeneity which overly restricts the distribution of coefficients. The finite mixture models of Gönül et al. (2000) and Donkers et al. (2006) are in principle capable to detect multimodality and skewness. Gönül et al. (2000) and Donkers et al. (2006) base the decision on the number of clusters on a heuristic, which consists in comparing values of the Bayesian information criterion (BIC) for an increasing number of clusters. This heuristic approach constitutes a weakness of the finite mixture methodology followed by these authors.

Our study aims both to fulfill requirements indicated by previous studies and to overcome shortcomings of previous work in several respects. We estimate the extent of the endogeneity bias and assess its relevance for catalog decisions. We account for endogeneity and heterogeneity simultaneously. We also adopt a nonparametric approach for heterogeneity which does away with heuristic rules often used to decide on the number of clusters. Additionally we evaluate different approaches to deal with the endogeneity problem by cross-validation.

The different models and their variants are evaluated by cross-validation predictive densities, which can be interpreted as cross-validated log likelihoods. To assess importance of endogeneity biases we test whether coefficients and elasticities of the number of catalogs allocated are greater for models which ignore endogeneity. Comparison of optimal solutions for models which either ignore or consider endogeneity indicates whether endogeneity has an effect on catalog allocation.

2. Sales response models

According to preponderant empirical evidence advertising effects are subject to diminishing returns to scale, i.e. sales increases caused by an additional unit of advertising become smaller at higher levels of advertising. The semi-log function¹ is able to reproduce diminishing returns to scale (Hanssens et al., 2001):

$$S_i = a_{0i} + a_{1i} \log C_i + a_{2i} \log PS_i + \epsilon_{1i} \quad (1)$$

S_i denotes sales with customer i , C_i the number of catalogs allocated to customer i , and PS_i previous year's sales with customer i (called

¹ We also investigated the double-log function, but confine the presentation to the semi-log function whose performance was much better in terms of cross-validation predictive densities.

Download English Version:

<https://daneshyari.com/en/article/481784>

Download Persian Version:

<https://daneshyari.com/article/481784>

[Daneshyari.com](https://daneshyari.com)