Original Article

How current targeting can hinder targeting in the future and what to do about it

Received (in revised form): 26th January 2009

Eddie Rhee

is an assistant professor of Business Administration at Stonehill College, Easton. He has also been a faculty at Santa Clara University. His research interests include customer scoring, customer contact strategy, customer relationship management and pharmaceutical marketing. He has published papers in the *Journal of the Academy of Marketing Science* and *Advances in Business and Management Forecasting*. He teaches Principles of Marketing, Marketing Research, Brand Management and Retail Management.

Shelby McIntyre

is a professor of Marketing at the Leavey School of Business, Santa Clara University. He has published in the *Journal of Marketing Research*, *Management Science* and the *Journal of Retailing*. He serves on the editorial board of the *Journal of Retailing* (2000–present). He teaches Marketing Information Systems, Marketing Research, Market Analysis, Brand Management, Strategic Product Marketing and Marketing Management.

ABSTRACT This study shows how current targeting of customers by a direct marketing firm can result in misleading models of customer response due to the truncation of observations for the customers who are not contacted, as well as the inevitable mis-specification of the explanatory variables in such models. These twin problems result in what technically is called *selection bias* and *endogeneity*, and are a direct result of standard industry practice when analysing such data. The current research corrects for these problems through a system of equations, including one for customer targeting and another for customer response, along with control variables. Hypotheses regarding the improvements the corrected approach would offer are that (1) the corrected approach would lower the inflated response forecast of the traditional approach; (2) the corrected approach would enable the customer targeting to be more efficient; and (3) the corrected approach would enable the customer targeting to be more effective. An application of the corrected approach to the customer database of a non-profit organisation shows that the corrected analysis supports the hypotheses and it further provides managerial implications.

Journal of Database Marketing & Customer Strategy Management (2009) 16, 15–28.

doi:10.1057/dbm.2009.5

Keywords: customer targeting policy; database marketing; targeting efficiency; targeting effectiveness; selection bias; endogeneity

Correspondence:

Eddie Rhee
Department of Business
Administration, Stonehill
College, 320 Washington
Street, Easton,
Massachusetts 02357, USA
E-mail: srhee@stonehill.edu

INTRODUCTION

A major advantage of database marketing without intermediaries is that the firm can collect fine detail about individual purchase such as what and how many products the

end customer has purchased, how much in dollar amount the customer has spent and how many days it has taken from the receipt of the marketing contacts until purchase. The firm can use such

information to model customer response probabilities, and thereby target more accurately the best customers for its current product offers. Such targeting has, however, subtle unintended consequences in that it essentially censors or truncates the resulting sales information for those who are not given the same opportunity to respond, and also often involves unobservable causal variables correlated with the included ones. Correcting response models for such degradation is the problem our research addresses.

The database marketing firm's customer targeting decision plays an important role, as the customer is heavily influenced by a new product offer with regard to making a purchase. For instance, the design and colour of products change frequently, and customers need to receive and browse a new catalogue or email in order to buy. In businesses such as direct mail of a non-profit organisation, consumers do not usually keep the old donation solicitation letter, and respond only when they receive a new one.

Firms have traditionally used customer purchase patterns, collectively known as RFM variables, in targeting the best customers.¹ Typically, Recency (R) is defined as the number of periods since the last purchase; Frequency (F) is defined as the total number of orders placed with the firm to date; and Monetary value (M) is defined as the dollar amount that the household has spent on all purchases from the firm over a standard period of time. The RFM values of the customers are updated on every targeting occasion. Conceptually, RFM variables are used for forecasting because past purchase behaviour is often the most reliable guide to future purchase behaviour.² For many database marketing categories, the likelihood of a future purchase is increasing in each RFM variable. Moreover, the predictive power of the three variables has the same rank order as the RFM name: R is the best predictor,

followed by F, and then M.³ Profiling a customer using RFM variables is widely used by database marketers and consultants as an easy and useful way of forecasting behaviour from a customer database.

By far, the most common forecasting method using RFM variables is the RFM coding method. After computing RFM values for all households, each variable is divided into five levels, defined by the quintiles of the household distribution. By cross-tabulating the coded RFM variables, each household can be categorised into one of 125 cells, each of which is assigned an empirically derived purchase probability often based on test mailings to that cell. RFM information is also inserted into more sophisticated predictive models. For example, RFM values are used as independent variables in a probit or logit response model.

In an effort to optimise the customer targeting effort, the firm strives to improve the efficiency (responses/piece) and effectiveness (revenue/piece) of targeting, and also assesses its current customer targeting policy for any modifications.

Potential degradation in analysing the customer database

Although the use of RFM values in the analysis of a customer database appears straightforward on the surface, two subtle but important problems arise as the firm inadvertently degrades the information content of the database.

The first problem is known as *selection bias*. This bias arises from the fact that the firm does not observe the customer's purchase behaviour on occasions when a product offer is not extended to that customer but is offered to others. As the purchase behaviour is observed only for the targeted occasions, the firm may inadvertently analyse only this part of the database. If the firm targets customers using its own (often non-random) selection rule, a model that only analyses the customers'

response to contacts generates biased results for targeting the customers, particularly those who were not contacted on some occasions based on models of those who did get an offer. In the literature, selection bias is known to arise when the firm uses a non-randomly selected sample to estimate behavioural relationships.⁴ This potential bias can only be controlled for by formally analysing the way in which the firm has targeted customers for marketing contacts (that is, emails, catalogues and direct mailing), in addition to the customer purchase analysis.

The second problem is known as the *endogeneity*. For instance, in studying a relationship between the consumer purchase and the marketing-mix variables such as price, promotion, advertising, and so on, the endogeneity problem exists with these variables, as both the marketing-mix variables and the purchase decisions are determined simultaneously as a function of the market situation, which is unobservable to the researcher.⁵ In formal statistical terms, the endogeneity with RFM yields incorrect parameter estimates in a predictive model due to *unobserved correlations* between the RFM variables and the error in the model.⁶ Hence, even after correcting for the selection bias problem, the potential endogeneity problem may still exist.

In the firm's customer targeting effort where the firm often uses RFM values for targeting, there might be other variables that the firm may use that are unobservable to the researcher but are correlated with RFM values. For instance, the firm may use the results of its customer experiment or a certain rule of thumb from its past experience for its selection rule. If a customer is not targeted for any reason, the R value (the number of periods since the last purchase) of this customer will be larger and F and M will be smaller. If the firm consistently ignores the customer for any reason, the RFM values of this customer will deteriorate regardless of the true

tendency to respond. Furthermore, even when the researcher considers all the variables that the firm is known to use for targeting, there may still be other non-quantifiable unobservable variables that are correlated with the RFM values.

In the customer purchase behaviour where the customer's purchase is often assumed to be related to the RFM values, there might be other unobservable variables that the customer may use such as the competing firm's product offers or word of mouth. If a customer does not make a purchase due to a competing firm's product offers, the R value of this customer will be larger and F and M will be smaller, and thus there will be a correlation between RFM values and the unobservable variables.

Marketing literature

In the literature, studies have begun to recognise each of these problems, but there seldom have been studies that recognised both the selection bias and endogeneity problems together, and very few have made corrections for both of them. Industry standard procedures (such as RFM coding and RFM response probit regression) and the early model of Bult and Wansbeek⁷ ignore these problems entirely.

Blumstein *et al*⁸ recognise the selection bias problem and propose a stochastic method for correcting the bias, but they do not deal with the endogeneity problem. Other studies by Bitran and Mondschein⁹ and Gonul and Shi¹⁰ provide an approach to dealing with RFM endogeneity. In addition, Chintagunta *et al*^{11,12} and Song and Chintagunta¹³ recognise the potential price endogeneity, but do not recognise the selection bias problem. Chen *et al*¹⁴ also recognise the potential endogeneity of a negotiated price in an automobile choice model, but again they do not directly deal with the selection bias problem. In these models, endogeneity correction was necessary, but no formal selection bias treatment was applied.

Rhee and McIntyre¹⁵ study the relative effect and the nature of the firm's prior and recent contact efforts on the customer purchase, but they address the problems in an indirect way. The current study extends and improves upon their approach by *directly* addressing the selection bias problem with a system of targeting and response equations, and the endogeneity problem by a control variable approach. Unlike their approach, this new approach provides measures of the magnitude and the seriousness of the selection bias and endogeneity problems in the house list.

Hypotheses

We set up several hypotheses about what improvements the corrected approach would offer the marketing manager.

Hypothesis 1: The corrected approach would lower the inflated response forecast of the traditional approach.

The traditional approach inadvertently analyses the customer response of only the targeted purchase occasions in the database, as the response (or not) can only be observed when the customer is targeted for customer contacts. As most firms target the best customers using a non-random selection rule, the response of the targeted customers tends to be higher than those not selected. The corrected approach recognises the firm's past targeting process and lowers the otherwise inflated response forecast.

Hypothesis 2: The corrected approach would enable the customer targeting to be more efficient.

As the traditional approach tends to inflate the response forecast, the firm would unnecessarily target more customers, but the actual response would be lower. If the firm uses the response forecast by the corrected approach, it would target fewer customers than the traditional approach would have

and it would result in higher response, and hence more efficient targeting.

Hypothesis 3: The corrected approach would enable the customer targeting to be more effective.

As the corrected approach would target fewer customers than the traditional approach, it would be less costly to target and would result in higher revenue, and hence more effective targeting.

The remainder of the paper is organised as follows. We first detail the corrected approach, discussing the components of the model that deal with the potential problems discussed above. The model is then applied to the customer database of a non-profit organisation. We show that the new approach performs better than a benchmark RFM response model and the model with selection bias treatment only. We also show that the corrected approach supports the hypotheses and enables the firm to evaluate the current targeting policy and to formulate a better policy. We conclude with a discussion of future research opportunities.

MODELLING FRAMEWORK

In the sample selection bias framework suggested by Heckman,⁴ economists use a system-of-equations approach for a number of different problems. For instance, Boyes *et al*¹⁶ study the problem that banks face in credit scoring. In assessing the bank customer's credit risk as to whether they will repay or default, the authors recognise that the credit risk of the customers in the random sample would be higher than that among customers who are granted the credit card loan. Thus, studying only the customers who are granted the credit card loan would provide biased results. They set up a system of equations to correct for the bias: an equation of whether the customer has been granted the credit card loan or

not, and an equation of whether the customer repaid or defaulted. This approach provides the bank with a credit risk profile, which is corrected for the potential bias due to the selected sample.

In the current study, a customer's likelihood of response to an offer and the firm's likelihood of targeting that customer for making an offer cause a selection bias, and an endogeneity exists in each of these cases. To deal with these problems, a system of two simultaneous equations is developed similarly. The two equations include (i) the firm's customer targeting of product offer and (ii) the customer's response to the offer. Aspects of this system of equations are detailed in the following sections.

The firm's customer targeting model

Following Heckman's⁴ procedure, the inclusion of the firm's customer targeting decision equation in the model corrects for the selection bias problem in the customer purchase history. The firm has targeted the customer seeming to be the most 'attractive'. More precisely, let the selection of a customer h by the firm at time t be defined as $Y_{S_{h,t}}$.

The selection rule can then be expressed as

$$Y_{S_{h,t}} = 1 \text{ if the customer is selected} \\ \text{meaning } U_{S_{h,t}} > 0 \text{ and } 0 \text{ otherwise.} \quad (1)$$

Above $U_{S_{h,t}}$ is the customer's attractiveness based on the following scoring model:

$$U_{S_{h,t}} = \alpha_0 + \alpha_1 \text{Demo}_h + \alpha_2 R_{h,t} + \alpha_3 F_{h,t} \\ + \alpha_4 M_{h,t} + \alpha_5 A_{h,t} + \alpha_6 R_{h,t}^* + \alpha_7 F_{h,t}^* \\ + \alpha_8 M_{h,t}^* + \alpha_9 A_{h,t}^* + \varepsilon_{S_{h,t}} \\ \varepsilon_{S_{h,t}} \sim N(0,1) \quad (2)$$

Note that $\varepsilon_{S_{h,t}}$ is an error term that represents any other unknown factors that may be related to the attractiveness of the customer. The variance of the error term is set to be one for model identification purpose, and hence this is a probit model. In our case, Demo is the customer's demographics; R is defined as the number of months since the most recent purchase; F is defined as the cumulative number of purchases; and M is defined as the total dollar purchase amount divided by the number of purchases, and thus is the historical average purchase level. As noted earlier, these three variables have been frequently used for modelling and conditional forecasting in the database marketing industry. A (Attempts) is defined as the number of marketing contacts delivered since the last purchase, and is analogous to the 'recent contacts' measure developed by Rhee and McIntyre¹⁵ to recognise that the firm's recent contact effort is directly related to its targeting decision at the current time.

The covariates in the customer targeting model are subject to the endogeneity problem. RFM values are endogenous in the system, as the firm may use other unobservable variables. For instance, the results of its customer experiment or a rule of thumb from its past experience of targeting decisions may have affected the RFM values. The A variable is also endogenous in the system, as the number of contacts since the last purchase is clearly the result of the firm's use of other unobservable targeting variables.

We employ a control variable approach in correcting the model for the potential endogeneity in RFM and A.¹⁷ We begin by modelling each endogenous variable as a function of its lagged value following the argument of Villas-Boas and Winer.⁵ We can use the lagged values, as they may be correlated with the error term in the previous time period but not with the errors in the current period. We first regress

the endogenous variables on their lagged values and compute the predicted value of the endogenous variables. Using these predicted values, we then obtain the residual value (endogenous variable minus predicted values) at each point in time. These residuals, R^* , F^* , M^* (residual of RFM) and A^* (residual of Attempts), are then placed in the model as the control variables. The theory underlying the control variable approach indicates that a significant non-zero coefficient on the residual variables (that is, α_6 , α_7 , α_8 , α_9) signals the presence of endogeneity in the data.¹⁷ Then, the estimated coefficient of the potentially endogenous variables (that is, RFM and A) will be more accurately determined, given the presence of these residual variables. Blundell and Smith¹⁸ show that this endogeneity correction approach of using residual values is asymptotically efficient in discrete and censored normal models. This theoretical consideration is particularly relevant here, as our model is both discrete (probit) and censored (those who do not respond are being censored out).

Customer purchase model

The customer is assumed to respond to the product offer if its attractiveness is above the threshold to that customer. Let the response of a customer h at time t be defined as $Y_{R_{h,t}}$.

The response can then be expressed as

$$Y_{R_{h,t}} = 1 \text{ if the customer responds meaning } U_{R_{h,t}} > 0 \text{ and } 0 \text{ otherwise.} \quad (3)$$

When some customers do not receive a marketing contact at a time period, they cannot purchase the product. Thus, $Y_{R_{h,t}}$ must be equal to 0 for these customers at this time period. We can only estimate the purchase model at a particular point in time over the customers who received a contact from the firm.

The attractiveness of the offer to the customer is defined as

$$U_{R_{h,t}} = \beta_0 + \beta_1 \text{Demo}_h + \beta_2 R_{h,t} + \beta_3 F_{h,t} + \beta_4 M_{h,t} + \beta_5 A_{h,t} + \beta_6 R_{h,t}^* + \beta_7 F_{h,t}^* + \beta_8 M_{h,t}^* + \beta_9 A_{h,t}^* + \varepsilon_{R_{h,t}} \quad (4)$$

$$\varepsilon_{R_{h,t}} \sim N(0,1)$$

where it is assumed that the attractiveness of the offer to the customer is related to the customer's Demo, RFM values and A. $\varepsilon_{R_{h,t}}$ is an error term that represents any other unknown factors that may be related to the attractiveness of the offer. The variance of the error term is set to be one for model identification purposes, and hence this is a probit model.

The covariates in the customer purchase model are also subject to the endogeneity problem. RFM values are endogenous in the system, as the customers may use other unobservable variables. For instance, the competing firm's product offers or a word-of-mouth encounter may have affected the RFM values. *Attempts* is also endogenous, as the number of contacts since the last purchase may also be related to the customer's use of other unobservable decision variables. The discussion of the endogeneity correction method applies to this model in exactly the same manner as in the firm's customer targeting model.

Properties of the errors

The relationship between the firm's customer targeting and customer purchase equations is modelled through a covariance matrix on $(\varepsilon_{S_{h,t}}, \varepsilon_{R_{h,t}})$. The covariance matrix is defined as

$$\Sigma = \begin{pmatrix} 1 & \sigma_{S,R} \\ & 1 \end{pmatrix} \quad (5)$$

The variance terms in the covariance matrix are ones, as the variance of the error terms is set to be one for model

identification purpose. Hence, the covariance value is equivalent to the correlation of the two equations. The covariance of the error terms ($\sigma_{S,R}$) measures the nature of the relationship between the firm's customer targeting and customer purchase likelihood, revealing the magnitude of the selection bias problem and whether there is value in the more complicated two-equation structure. If the covariance is statistically significant, it shows that there exists selection bias. If the value is positive, it implies that the firm has targeted the customers with higher purchase likelihood, as would be expected.

EMPIRICAL APPLICATION

Clearly, the corrected model is appropriate to any database marketing situation where prior targeting has developed a 'house list' that is to be frequently contacted again. A very 'clean' example of such a situation is direct mailing for a charitable organisation. The direct mail services industry handles most of the solicitation for charitable organisations, which generally 'outsource' this function to such 'experts'. It is estimated that charitable giving in the United States amounted to US\$260 billion in 2006, which constituted about 2.2 per cent of the gross domestic product that year.¹⁹

Fifty-nine per cent of Americans claim that they respond to charities based on the information that comes via direct mail.²⁰ Yet, the overall response rate to direct mail is 2.73 per cent,²¹ meaning that most direct mail is not responded to. Thus, firms in this industry need greater understanding of the factors that influence a donor's giving and what causes those behaviours to change.

Data description

The data consist of transaction records supplied by the Direct Marketing Educational Foundation (DMEF) of a non-profit organisation that uses direct mail to solicit contributions as the main vehicle

of its campaigns. Each record contains information on donor ID, postal code, donation history and mailing dates. As the contribution codes and mailing codes match, each contribution can be traced to a specific mailing and date.

The data set we use is a random sample of 895 households provided by the DMEF for our use. As, in these data, the majority of the customers respond within 3 months after the receipt of a mailing, the response behaviour is truncated at 12 months after the last mailing. It can safely be assumed that after 12 months the likelihood of the customer responding would be negligible. The data are based on 24 mailings during the period from July 1991 to January 1994. Thus, there are 21 480 observations in this data set. The first 14 mailings are used for estimation and the remaining 10 mailings are used for model validation.

An analysis of donations shows that the amount of money given by a household varies little over time. That is, if a given customer donates \$10 on one occasion, that customer is very likely to donate \$10 on every occasion. For this reason, it is possible to regard the amount of the donation as a stable characteristic of the household. We therefore concentrate only on the customer's response decision in this application.

There are three major solicitation types: R, S and a miscellaneous type. Type R refers to 'regular mailings', which are sent every three to six months. Type S refers to 'seasonal mailings', which are sent during December and January and other 'seasonal' times of the year. In some cases, the type of solicitation sent to a household is of some other type or is not recorded in the data set. These are referred to as Type M or 'miscellaneous mailings'. By separating out the Type M solicitations, we are able to study the characteristics of Type R and Type S without making unwarranted assumptions. As we show subsequently, the selection and response characteristics of Type R and Type

Solicitations are decidedly different. In our analysis of the donation data set, we modify the general model of customer targeting and purchase and take account of the three types of solicitations.

A postal code data set is used to obtain a demographic description of each household. This data set includes postal code, income index, percentage of households occupied by white, black and Hispanic persons, percentage of households with one or more children under 18, number of people in the household, household median age and median years of school for people aged 25 or more. Including gender information taken from the donation record, there are a total of nine demographic variables. In order to rule out any correlations among the demographic variables, we use principal components analysis to create a set of nine uncorrelated demographic variables. All nine principal component variables are used in the analysis.

Estimation

We calibrate the model by formulating the estimation problem using a Hierarchical Bayes approach, and by employing Markov Chain Monte Carlo (MCMC) methods to simulate draws from the posterior distribution of parameters.²² The prior distributions of all the coefficient population parameters are taken from the normal distribution and the prior distributions of the population precision matrix (inverse of the covariance matrix that contains $\sigma_{S,R}$) is taken from a Wishart distribution. All prior distributions are set to be diffuse, allowing the data to dominate the estimation of the parameters.

Assessing forecast accuracy

We assess the relative performance of the models by the Mean Absolute Deviation (MAD) of predicted versus actual response in the holdout data. To take account of uncertainty in the Bayesian estimation, the measures (MADs and coefficients) are

Table 1: Model fit

<i>Model</i>	<i>MAD</i>
Traditional model	0.3974
Selection bias-corrected model	0.3538
Selection bias and endogeneity-corrected model	0.3414

MAD is on the holdout sample.

MAD of the Corrected Models are significantly lower ($P < 0.01$).

computed using the last 10 000 simulated draws of parameters from the posterior distribution of the models estimated by MCMC. Each of these 10 000 draws of parameters is then used to predict the holdout sample, and the MAD is calculated across those 10 000 rounds of prediction. This methodology allows accurate generation of the standard deviation of the MAD statistic of each model. Therefore, it is possible to perform t -tests of the MAD in holdout from one model against the MAD in holdout of another model. Using holdout data is particularly important here, as the models being compared have different numbers of coefficients and we need to be sure that a larger model is not 'over fitting' the estimation data (Table 1).

Three models are compared for their relative performance. The traditional model, although it is standard industry practice, does not recognise, let alone correct for, the potential problem of selection bias and endogeneity. This traditional model achieves a MAD of 0.3974. The model that corrects for only selection bias reduces the MAD significantly to 0.3538 ($P < 0.01$). Finally, the model that corrects for both selection bias and endogeneity reduces the MAD even further to 0.3414 ($P < 0.01$).

Estimation results

The 'Traditional Model' and the 'Corrected Model' (corrected for selection bias and endogeneity problems) are compared side by side in Table 2.

Both models use precisely the same variables except that the corrected Model

Table 2: Traditional versus corrected model

	<i>Traditional model</i>		<i>Corrected model (eg, system of equations correcting for selection bias and endogeneity)</i>			
	<i>Customer response</i>	<i>SD</i>	<i>Customer response</i>	<i>SD</i>	<i>Customer targeting</i>	<i>SD</i>
Intercept	-0.5066*	0.0298	-1.1090*	0.1277	-0.7524*	0.0141
<i>Demographics</i>						
Non-white	-0.0203	0.0191	-0.0141	0.0191	0.0179*	0.0083
High income/education	-0.0119	0.0195	-0.0120	0.0196	-0.0168	0.0102
Non-white, children >18	0.0214	0.0222	0.0230	0.0221	0.0106	0.0116
Non-hispanic male	-0.0384	0.0240	-0.0458	0.0243	-0.0200	0.0126
Hispanic male	0.0103	0.0289	0.0179	0.0290	0.0126	0.0140
Household size, age	-0.0813	0.0584	-0.0795	0.0585	-0.0257	0.0181
Children under 18, age	0.0252	0.0565	0.0259	0.0557	0.0067	0.0211
Low income/education	0.0791	0.0668	0.0731	0.0685	-0.0015	0.0339
White and black (non-ethnic)	-0.2602*	0.1123	-0.2853*	0.1123	-0.0470	0.0601
<i>Type R</i>						
Recency	-0.4781*	0.1134	-0.5492*	0.1164	-0.1103*	0.0327
Frequency	0.4158*	0.0685	0.3911*	0.0677	-0.0849*	0.0240
Monetary	-0.1326*	0.0666	-0.2277*	0.0706	-0.1461*	0.0307
Attempts	0.1511*	0.0530	0.1537*	0.0564	-0.0698*	0.0238
<i>Type S</i>						
Recency	0.2605*	0.0682	0.0511	0.0787	-0.3213*	0.0367
Frequency	0.1570*	0.0434	0.2009*	0.0428	0.0809*	0.0228
Monetary	0.0793	0.0448	0.0638	0.0445	-0.0451	0.0235
Attempts	-0.2936*	0.0495	-0.2127*	0.0570	0.1959*	0.0240
<i>Type M</i>						
Recency	-0.4674*	0.1026	-0.8451*	0.1292	-1.1330*	0.0665
Frequency	0.4158*	0.0485	0.4322*	0.0497	0.0497	0.0322
Monetary	-0.4590*	0.0696	-0.6771*	0.0834	-0.6477*	0.0401
Attempts	0.0477	0.0486	0.2010*	0.0593	0.5604*	0.0364
<i>Endogeneity correction</i>						
Recency residual	—	—	0.0312	0.0955	0.0389	0.0381
Frequency residual	—	—	-0.5561	0.6232	-0.0677	0.2782
Monetary residual	—	—	-3.8670*	1.1770	-0.3018	0.3707
Attempts residual	—	—	0.0375	0.0778	0.5537*	0.0405
			<i>Estimate</i>	<i>St dev</i>		
Covariance			0.4537*	0.0912		

Note: Parameter estimates denoted by (*) are statistically different from zero ($P < 0.05$).

includes residuals for R, F and M as well as A. In addition, the 'corrected model' involves the second, or 'Customer Targeting', equation (the two columns to the far right), which is also based on the same set of variables. The most interesting comparison of the models is for the estimates for the 'Customer Response' coefficients, which both models provide.

Before comparing the customer response coefficients, however, attention is directed to the covariance statistic (see bottom of

Table 2), which is quite large at 0.4537 for the corrected model. This is analogous to the correlation between the error terms of the Customer Purchase and Customer Targeting equations. The high and statistically significant value implies that selection bias does exist in this case and the associated complication of having two equations is needed. It indicates the degree to which Customer Targeting has been at the 'most attractive customers'. It also explains *why* the MAD (in holdout) is

reduced by about 14 per cent (down from 0.3974 to 0.3414), which is a highly significant reduction ($P < 0.01$), as mentioned previously.

Given this understanding that (i) there exists substantial prediction improvement from correcting with the targeting equation and the control variables (as shown by MAD results) and (ii) *why* that improvement occurs (as shown by the covariance result), the next question relates to (iii) *where* is that improvement occurring (that is, with improved estimates of which variables)? To address this question, refer to the columns of coefficients for 'Customer Response' in each model. For the two models, the signs of the coefficients match across all of the variables, but a major difference between the response model results of the two models is in the magnitude of the intercept. The intercept of the corrected model is lower (-1.1090) than that of the traditional model (-0.5066), which implies that the probability of response becomes lower after correcting for the problem of selection bias. This result supports Hypothesis 1: *The corrected approach would lower the inflated response forecast of the traditional approach.* In addition, the statistical significance of the slope of the Monetary Residual variable implies that the endogeneity problem exists with M.

Figure 1 shows the predicted probability of the traditional model versus that of the corrected model. Using the estimated parameters and the predictors of the model, the probability of response is computed as a function of the deterministic part of the model (that is, estimated parameters $\times$ predictors) for each of the mailing types. For the Type R mailings, the predicted probability of response ranges from a value close to 0 per cent up to about 85 per cent. For the Type S mailings, the predicted probability of response ranges from about 10 per cent up to about 45 per cent. This difference in probability range is due to the nature of Type S mailings (that is, holiday and seasonal mailings).

Now comparing the models, for Type R mailings, the traditional model shows an inflated predicted probability of response that is higher than the corrected model by as much as 20 percentage points over most of the range. For the Type S mailings, the traditional model shows an inflated predicted probability of response that is higher than the corrected model by as much as 30 percentage points over most of the range. The traditional model's probability of response is inflated, as only the targeted mailing occasions in the database are analysed. If the firm uses the probability of response predicted by the traditional model, it will end up selecting more people than the corrected model would have, and it will result in a lower response rate. On the other hand, the corrected model recognises the firm's past targeting process and corrects for the inflated probability of response, and hence it will produce a more accurate prediction of the true response probability. The firm using the probability of response predicted by the corrected model will be more efficient in the selection of customers, and will reduce the amount of 'junk mail'.

Beginning with the demographics of the corrected model, most of the demographic variables are not statistically significant. The household constant variables such as demographics are usually not a good predictor of the response behaviour.

For the Type R mailings, the signs on R and F show the same direction as the industry prediction, but the frequent donors for the regular mailing in this data set tend to donate smaller amounts. Owing to the nature of the regular mailings (that is, mailings every three to six months), the customers who respond to these types of mailings are the regular donors, but are those who donate in smaller amounts at a time, which would show lower R and M and higher F. In addition, the customers who received more mailings since the previous response have a higher chance of

Figure 1: Illustration of traditional versus corrected model results.

responding to the current mailing. It indicates that wear-out phenomenon^{23–26} is not observed for the regular mailing and the regular donors want more frequent reminders.

The Type S mailing results show that the customers who have donated more frequently show a higher chance of responding to the mailing. This implies that the customers who donate to holiday and seasonal mailings are also regular donors. The coefficients of R and M variables are not

statistically significant, however. Unlike the regular mailing, how recent the last donation was and the donation amount do not show much explanatory power for the response. This result may be due to the fact that the holiday mailings are not as frequent as the regular mailings. The customers who have received fewer mailings since the previous response show a higher chance of responding to the current mailing. Here the wear-out phenomenon is observed and the holiday donors do not want too many mailings.

Assessing the efficiency and effectiveness

As shown in Figure 1 and discussed earlier, the Traditional model inflates the predicted probability of response (in the holdout sample), and thus for any probability-of-response cutoff value it selects more people for mailing, too many of whom end up not responding. To further test the hypotheses, two sets of measures are computed, efficiency (that is, response/ mailing) and effectiveness (that is, revenue/ mailing) on the holdout data, which contain 4240 mailings and 1280 responses. In the computation, each model (traditional and corrected) is used to select the customers to mail, based on an assumed series of cutoff probabilities of response ranging from 0.10 to 0.50. The selection is carried out by ranking the customers according to the *predicted* probability of response for the respective model and then assuming that the mailings were to all those over the cutoff probability (again by the respective model). Under each cutoff probability, the average number of responses (number of responses/ number of contacts) and average revenue (revenue/ number of contacts) are then computed.

The first table provides efficiency measures in terms of the average number of responses per mailing under specified cutoff probabilities. Under each cutoff probability assumption, the corrected model results in a larger average number of responses per mailing than the traditional model. The percentage improvement is 13.7–37.6 per cent across the range of cutoff values considered (all when compared to the traditional model). This result supports Hypothesis 2: *The corrected approach would enable the customer targeting to be more efficient.*

The second table provides effectiveness measures in terms of the average revenue per mailing. The corrected model results in larger average revenue per mailing than the traditional model. The percentage improvement is 2.4–57.2 per cent across

the cutoff values investigated (again compared to the traditional model). This result also supports Hypothesis 3: *The corrected approach would enable the customer targeting to be more effective.*

Managerial implications

Although effectiveness and efficiency are certainly key success metrics for customer contacting, the most managerially relevant metrics are overall revenue and the resulting profit contribution, which are the metrics we now address.

Revenue and profit comparison

A comparison of the traditional and corrected model on profit contribution is best illustrated by considering a direct mailing of 10 000 pieces sent according to the traditional model to those who are predicted to have greater than a 45 per cent likelihood of responding. The response rate would be 51 per cent (see Table 3, top panel) or 5100 people, and the average donation would be \$4.59 (see Table 3, bottom panel) for total revenue from the

Table 3: Efficiency and effectiveness comparison of traditional versus corrected model

Cutoff Pr(response)	#Responses/#Contacts		Percentage improvement
	Traditional model	Corrected model	
Efficiency (responses per mailing)			
0.50	0.56	0.68	20.8
0.45	0.51	0.70	37.6
0.40	0.46	0.66	43.0
0.35	0.42	0.59	40.3
0.30	0.39	0.56	43.9
0.25	0.35	0.52	45.3
0.20	0.33	0.47	44.2
0.15	0.31	0.41	29.8
0.10	0.31	0.35	13.7
Effectiveness (revenue per mailing)			
0.50	\$4.92	\$7.73	57.2
0.45	\$4.59	\$5.81	26.5
0.40	\$4.14	\$5.38	29.8
0.35	\$3.78	\$4.66	23.2
0.30	\$3.59	\$4.39	22.3
0.25	\$3.33	\$4.09	23.0
0.20	\$3.15	\$3.70	17.6
0.15	\$3.09	\$3.40	9.9
0.10	\$3.09	\$3.16	2.4

mailing of $5100 \times 4.59 = \$23\,409$. On the other hand, if the 10 000 pieces had been sent, instead, according to the corrected model, the mailing would have had a 70 per cent response rate, or 7000 people and an average donation of \$5.81 for total revenue of \$40 670.

Furthermore, if the mailing cost is \$2.00 per piece, the total mailing would have cost \$20 000 and thus the traditional model would have resulted in a \$3409 contribution yielding a return on investment (ROI) of 17 per cent ($3409/20\,000$). This compares to the expected results from the corrected model of \$20 670 in contribution, and thus $\text{ROI} = 101$ per cent ($20\,670/20\,000$). Thus, the corrected model leads to fairly dramatic predicted improvements in this instance.

Policy evaluation

In addition to the increase of efficiency and effectiveness and ROI of targeting, the corrected model further allows the firm to evaluate its current targeting policy on a more qualitative basis. If the current targeting policy is misaligned with the customer's response behaviour, the firm can identify where to improve its policy. For instance, as shown below first for regular mailings and then for special seasonal mailings, the firm's current targeting policy is not consistent with the customer response behaviour, which is why and how the firm would benefit from modifying its current policy.

The firm's customer targeting policy for Type R (regular) mailings with respect to the F and A variables contrasts with the customer's response behaviour. The signs on these variables are negative and significant for the customer targeting equation, but they are positive and significant for the customer response equation. The firm is therefore recommended to target more heavily (than in the past) those customers who have donated recently, more frequently but in smaller amounts, and the customers who have received more mailings since the

last response. Using the customer response of the corrected model for targeting will accomplish the modified mailing policy.

The firm's customer targeting policy for Type S (holiday and seasonal) mailings with respect to the Frequency variable is consistent with the customer response behaviour, as the signs for both are positive. The targeting policy with respect to the A variable contrasts with the customer response behaviour. In addition, the firm's targeting policy with respect to R is not supported by the response behaviour, as the slope is not significant for the customer response. For the holiday and seasonal mailing, the firm is therefore recommended to target more strongly the customers who donated more frequently and who received fewer mailings since the last response.

CONCLUSION

This study recognises the potential problems with using the customer database, and corrects the problems in a system of simultaneous equations through the covariance structure of the errors of the two equations and by including the control variables as additional covariates. The forecasting performance with the holdout data shows that the corrected model performs significantly better than alternative models. The corrected model would provide more efficient and effective targeting effort than the traditional model, and would further provide higher ROI and managerial direction for better customer targeting policy. The empirical results of the corrected model also show a noticeable difference in customer response behaviour between the regular mailings and holiday mailings.

Limitations and extensions

The study has several limitations, all of which provide avenues for future research. In this application with stable donation amounts by donors, the current model predicts only the firm's targeting and the customer's response. A future study with a new data set might

need to include the dollar amount of spending in the model. The current model is also limited to one kind of product, which is donation, but the catalogue companies sell a large array of products. In future applications to such data, the model would need to be extended to consider multiple-category products in the basket.

In addition, a better validation of the corrected model could be generated by a true field experiment, dividing the customers into two equal groups and mailing to one group according to the corrected model and the other group according to the traditional model. Such a true experiment, of course, would require control of the mailings.

ACKNOWLEDGEMENTS

The authors thank the Direct Marketing Educational Foundation for providing access to the data used in this study.

REFERENCES

- Hughes, A. M. (1996) Boosting response with RFM. *Marketing Tools* 3(2): 4–10.
- Schmid, J. and Weber, A. (1997) *Desktop Database Marketing*. Chicago, IL: NTC Business Books.
- David Sheppard Associates, Inc. (1999) *The New Direct Marketing: How to Implement a Profit-Driven Database Marketing Strategy*. Boston, MA: McGraw-Hill.
- Heckman, J. J. (1979) Sample selection bias as a specification error. *Econometrica* 47(1): 153–161.
- Villas-Boas, J. M. and Winer, R. (1999) Endogeneity in brand choice models. *Marketing Science* 45(10): 1324–1338.
- Davidson, R. and James, G. M. (1993) *Estimation and Inference in Econometrics*. New York: Oxford University Press.
- Bult, J. R. and Wansbeek, T. (1995) Optimal selection for direct mail. *Marketing Science* 14(4): 378–394.
- Blumstein, A., Canela-Cacho, J. and Cohen, J. (1993) Filtered sampling from populations with heterogeneous event frequencies. *Management Science* 39(7): 886–899.
- Bitran, G. R. and Mondschein, S. V. (1996) Mailing decision in the catalog sales industry. *Management Science* 42(9): 1364–1381.
- Gonul, F. and Shi, M. (1998) Optimal mailing of catalog: A new methodology using estimable structural dynamic programming models. *Management Science* 44(9): 1249–1262.
- Chintagunta, P., Dube, J. and Goh, K. (2005) Beyond the endogeneity bias: The effect of unmeasured brand characteristics on household-level brand choice models. *Management Science* 51(5): 832–849.
- Chintagunta, P., Kadiyali, V. and Vilcassim, B. (2006) Endogeneity and simultaneity in competitive pricing and advertising: A logit demand analysis. *Journal of Business* 79(6): 2761–2787.
- Song, I. and Chintagunta, P. (2006) Measuring cross-category price effects with aggregate store data. *Management Science* 52(10): 1594–1659.
- Chen, Y., Yang, S. and Zhao, Y. (2008) A simultaneous model of consumer brand choice and negotiated price. *Management Science* 54(3): 538–548.
- Rhee, S. and McIntyre, S. (2008) Including the effects of prior and recent contact effort in a customer scoring model for database marketing. *Journal of the Academy of Marketing Science* 36(4): 538–551.
- Boyes, W. J., Dennis, L. H. and Stuart, A. L. (1989) An econometric analysis of the bank scoring problem. *Journal of Econometrics* 40(1): 3–14.
- Blundell, R. and Powell, J. L. (2004) Endogeneity in semiparametric binary response models. *Review of Economic Studies* 71: 655–679.
- Blundell, R. W. and Smith, R. J. (1986) An exogeneity test for a simultaneous Tobit model. *Econometrica* 54: 679–685.
- Brown, M. (ed.) (2007) *Giving USA: The annual report on philanthropy for the year 2006*. Glenview, IL: Giving USA Foundation.
- Gardyn, R. (2003) Not giving up. *American Demographics* 25(8): 14–15.
- The Direct Marketing Association, Inc. (2006) *The DMA 2004 Response Rate Report*. New York: The Direct Marketing Association, Inc.
- Gelman, A., Carlin, J., Stern, H. and Rubin, D. (1996) *Bayesian Data Analysis*. New York: Chapman and Hall.
- Gonul, F., Kim, B. and Shi, M. (2000) Mailing smarter to catalog customers. *Journal of Interactive Marketing* 14(2): 2–16.
- Little, J. D. C. (1979) Aggregate advertising models: The state of the art. *Operations research* 27: 629–667.
- Simon, H. (1982) ADPULS: An advertising model with wearout and pulsation. *Journal of Marketing Research* 19: 352–363.
- Blair, M. H. and Rabuck, M. (1998) Advertising wearin and wearout: Ten years later – more empirical evidence and successful practice. *Journal of Advertising Research* 38(6): 7–18.