IMPROVING THE PROFITABILITY OF DIRECT MARKETING: A QUANTILE REGRESSION APPROACH

ZHANG XI

MPHIL

LINGNAN UNIVERSITY

2009

IMPROVING THE PROFITABILITY OF DIRECT MARKETING: A QUANTILE REGRESSION APPROACH

by

ZHANG Xi

A thesis
submitted in partial fulfillment
of the requirements for the Degree of
Master of Philosophy in Business
(Marketing and International Business)

Lingnan University

2009

ABSTRACT

Improving the Profitability of Direct Marketing:
A Quantile Regression Approach

by

ZHANG Xi

Master of Philosophy

Direct marketing is to target consumers who are most likely to respond. A number of target selection methods have been employed to select potential customers. These methods either only consider the customer response probability and ignore the profit issue or assume that the estimates of profit are homogenous across customers when considering the expected amount of profit. Furthermore, the traditional analytical techniques based on ordinary least squares (OLS) regression, which focus on the average customer, cannot examine the differences of various customer groups or account for customer heterogeneity in profitability estimates. Quantile regression, instead of the point estimate for the conditional mean, can be used to estimate the whole distribution, especially the upper tail which we are interested in. Quantile regression does not have strict model assumptions as OLS does and is not sensitive to outliers. To model consumer response profit in direct marketing, this thesis tested the endogeneity bias in the recency, frequency, monetary value (RFM) variables using the control function approach, made sample selection bias correction using Heckman’s procedure, and then adopted quantile regression to estimate customer profit and make forecast of the profit distribution of future values. Furthermore, we adopted the recentered influence function (RIF) regression methods proposed by Firpo et al. (2007) to perform unconditional quantile regression for customer profit estimation. The comparison of OLS, conditional and unconditional quantile regression shows that while OLS may induce possible misleading estimation results, conditional and unconditional quantile regression can provide more informative estimation results. The findings can help direct marketers augment the profitability of marketing campaigns and have meaningful implications for solving target marketing forecasting problems given the constraint of limited resources.

DECLARATION

I declare that this is an original work based primarily on my own research, and I warrant that all citations of previous research, published or unpublished, have been duly acknowledged.

ZHANG Xi

Date:

CERTIFICATE OF APPROVAL OF THESIS

IMPROVING THE PROFITABILITY OF DIRECT MARKETING: A QUANTILE REGRESSION APPROACH

by

ZHANG Xi

Master of Philosophy

Panel of Examiners:

______________________________ (Chairman)
Prof. Ling-Yee LI

______________________________ (External Member)
Prof. Xubing ZHANG

______________________________ (Internal Member)
Prof. Man-leung WONG

______________________________ (Internal Member)
Prof. Geng CUI

Chief Supervisor: Prof. Geng CUI

Co-supervisor: Prof. Tsang-sing CHAN

Approved for the Senate:

____
Prof. SEADE Jesus
Chairman, Research and Postgraduate Studies Committee

____
Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgement

I am especially grateful to my Chief-Supervisor, Prof. Geng CUI, and Co-Supervisor, Prof. Tsang-sing CHAN, for their thorough guidance on my thesis. I also appreciate the financial support provided by Lingnan University, which makes it possible for me to complete my studies. Furthermore, I wish to extend my appreciation to Dr. Ling Peng, Dr. Patrick POON, and other faculty members for their constructive advice on certain aspects of my thesis, Mr. Simon LO, Ms. Xue GAO and Ms. Xiaoning GUO for their research assistance, and Ms. Linda JO, and Ms. Vanessa CHAN for their kindly help during my studies. Finally, I want to thank my parents for their love and encouragement.

1 INTRODUCTION

1.1 Research Problem

Direct marketing is in nature targeting a group of consumers who are the most likely to respond and to bring the highest level of profit to the businesses. It typically only sends the promotional materials to 20% or less of its potential buyers from its list. Customer profitability analysis reveals that a large portion of profits comes from a relatively small number of customers and that not all the customers are equally profitable. In addition, the huge disparity of response profit across different segments of customers suggests heterogeneous response behavior to direct marketing campaigns of customers with different characteristics. Thus, the key problem in direct marketing is to find a target selection method which can both examine the heterogeneity of customer behavior and identify the most profitable customers.

The typical analytical technique for analyzing profit, ordinary least squares (OLS) regression, does not taken into consideration of such heterogeneity and focuses on the “average customers”. In addition, OLS regression can only estimate the central location of profit distribution and is highly influenced by outliers. In direct marketing, estimating the effects of explanatory variables on other parts (e.g., upper tail) of the profit distribution is more important for targeting purposes. Furthermore, both response and profit are highly skewed in their distribution and do not conform to the normality assumption of OLS regression. In this study, we propose a novel approach, quantile regression, to customer profitability analysis. We believe that quantile regression is able to account for heterogeneity in profit and examine the purchase behavior pattern of the most profitable customers whom are the target of direct marketing companies.

1.2 Rationale

Conditional quantile regression, introduced by Koenker and Bassett (1978), extends the OLS regression model to conditional quantiles (i.e. 90^{th} quantile) of the response variable. It has significant advantages over OLS regression, especially for customer profitability analysis in our case. It dose not have the strict model assumptions, such as assumptions for normality and homogeneity of variance, which are often violated in customer profit data. Its estimation, based on the weighted least absolute deviations (LAD) method, gives a robust measure of location and is not sensitive to outliers as OLS regression is. It naturally can be used to characterize the entire conditional distribution of a dependent variable given a set of explanatory variables. In our study, conditional quantile regression is used to examine the effects of a set of explanatory variables, such as RFM variables—Recency, Frequency, and Monetary Value, lifetime value and marketing intensity variables, and consumer transactional variables, on the conditional quantile of customer profitability. We also use the estimated conditional quantile regression model to predict the entire customer profit distribution of future values.

A further development of the conditional quantile regression is the unconditional quantile regression, proposed by Firpo et al. (2007). Unlike conditional mean, conditional quantiles do not average up to their unconditional population counterparts. Hence, unlike OLS regression, conditional quantile regression estimates cannot be used to assess the more general impact of changes of predictor variables on the corresponding quantile of the unconditional distribution of an outcome variable. The proposed unconditional quantile regression method consists of running a regression of the (recentered) influence function of the unconditional quantile of the outcome variable on the explanatory variables. Unconditional quantile regression is proved to be able to obtain effects of predictor variables at different quantiles of the unconditional distribution. In our study, we use unconditional quantile regression to obtain estimations at different quantiles of the unconditional profit distribution, especially focusing on the quantile in which managers have an interest, i.e. 90^{th} quantile.

1.3 Objectives and Significance of This Study

This study has several objectives. First, by quantile regression, this study attempts to explore how the high profit customers are different from the other, in terms of customer characteristics, reactions to marketing intensity, and purchase behaviors. Second, this study attempts to forecast customer future profit by conditional quantile regression. The predictions are not point prediction as OLS regression gives but probabilistic prediction of the whole future profit distribution of customers. Third, this study ranks the testing data set by the predicted profit values given by OLS regression and quantile regression. By decile analysis, this study compares the prediction results and assesses the model performance. Fourth, this study uses the new unconditional quantile regression in customer profitability analysis, and compares the estimates among OLS regression, conditional and unconditional quantile regression. This study further examines the differences between conditional and unconditional quantile regression and addresses the suitability of unconditional quantile regression in estimating the unconditional quantile customers.

Before adopting quantile regression analysis to a direct marketing dataset, we make correction of the potential endogeneity of RFM variables. We further make correction of sample selection bias in our data using the standard Heckman's procedure (1979). We then use conditional and unconditional quantile regression for customer profitability analysis. The estimated conditional quantile regression model is also used to make probabilistic forecast of the profit distribution of future values. The results show that the slope estimates of coefficients vary across different quantiles, indicating heterogeneity of variance. The differences of the signs and significance of predictor variables between OLS and quantile regression shows that OLS regression may miss the significance of some predictor variables and also give a misleading estimation of the effects. Comparison of the conditional and unconditional quantile regression estimated coefficients gives us a more accurate estimation of the effects of predictor variables on the outcome variable, conditionally or unconditionally.

This thesis is one of the few studies that adopt quantile regression in analyzing marketing problems. While OLS regression only estimates the expected profit, quantile regression gives estimation of the whole profit distribution, giving marketing managers more information of the purchase pattern of customers. Our proposed probabilistic forecast of the customer profit distribution by conditional quantile regression gives a more accurate interval prediction than OLS regression. The comparison of the conditional and unconditional quantile regression supports the findings of Firpo et al. (2007) and further shows the usefulness of unconditional quantile regression in empirical application, in our case, to solve direct marketing problems.

1.4 Organization of Our Study

The rest of the thesis is organized as follows. In Chapter 2, we review the literature of direct marketing modeling and common target selection methods, the features of customer profitability analysis and the limitation of OLS in such analysis, and then propose a new method—quantile regression. In Chapter 3, we discuss the basic theory and interpretation of quantile regression, the advantages of quantile regression, its development and empirical applications, its use for customer profitability analysis, including its suitability and relevant applications. We further introduce a new development of quantile regression—unconditional quantile regression. In Chapter 4, we discuss the issue of quantile regression forecasting and relevant applications. In Chapter 5, we explain our data and variables, perform descriptive data analysis, make correction of endogeneity and sample selection biases, and use conditional and unconditional quantile regression to analyze our data and discuss the empirical results. In Chapter 6, we discuss the findings, draw conclusions and suggest directions for further research.

2 LITERATURE REVIEW

2.1 Direct Marketing

In the last few decades, direct marketing has become an important field of marketing. In 2008, direct marketing accounted for approximately 10% of total US gross domestic product and 1.6 million direct marketing employees in the US. Their collective sales efforts directly support 9.3 million other jobs, accounting for a total of 10.9 million US jobs (Direct Marketing Association Statistical Fact Book 2009). In fact, nowadays more and more companies are using the information about their customers' preference and behavior, which is stored in their databases, to target their marketing efforts. Moreover, many companies are using direct marketing channels as their main strategy for interacting with their customers (Bult 1993).

Direct Marketing is a sales and promotion process in which the promotional materials and information are sent to individual customers via direct calling, mail, catalogue and so on (Bitran & Mondschein, 1996). Among these communicating channels, direct mail is the most important medium of the various direct marketing media. Advertising expenditures via direct mail increase annually. According to the Direct Marketing Association, in 2008 commercial and nonprofit marketers spent $176.9 billion on direct marketing, accounting for 52.1% of all ad expenditures in the US. These advertising expenditures generate approximately US $2.057 trillion in total incremental sales.

Given the rapid growth of direct marketing in recent years, the accurate prediction of consumer response to direct marketing campaigns has become a priority for many companies (Bodenberg & Roberts 1990). The prediction of customer response mainly focuses on identifying the potential buyers who can be called target customers or respondents. This can be done by analyzing data from previous campaigns or by organizing test mail campaigns from which models can be generated to select the customers who will be targeted. Therefore, much direct marketing research focuses on segmentation or target selection.

2.2 Target Selection

Target selection, also called list segmentation, can be seen as the process of dividing the market, i.e. the mailing list, into two distinct groups, viz. a group that should receive a mailing and a group that should not receive a mailing. Target selection is an important data mining problem for direct marketing. It is obvious that target selection is a crucial component for the development of a direct mailing campaign since a campaign can only be effective if the mailing reaches the proper targets who are the most likely to respond. Therefore, direct marketers have expended considerable effort towards target selection methods. See Roberts and Berger (1989), Bult and Wansbeek (1995), Jonker et al. (2004) and Otter et al. (2006) for overall reviews of these target selection methods.

According to Jonker et al. (2004) and Kaymak (2001), target selection methods can be divided into two groups: segmentation methods and scoring methods. Segmentation techniques aim to divide individuals into groups (segments) using a number of explanatory variables, such that each segment is expected to be more or less homogeneous with respect to these variables and their (expected) response to a direct mail offer. The segments that have the highest probability to respond are then selected to receive a mailing. The scoring approach assigns a separate score to each individual customer and the score is indicative of the likelihood of response of the customer. The customers are then ordered according to their likelihood of response based on the prediction of the target selection model. Only the customers who are the most likely to response (e.g. their score is above a threshold value) will receive the product offer. The next two subsections provide an overview of these selection methods.

2.2.1 Segmentation Methods

The most frequently used method is the so-called Recency, Frequency, and Monetary value model (RFM-model). Recency measures include the number of consecutive mailings without response and the time period since the last order. Frequency measures include the number of purchases made in a certain time period. Monetary value measures include the amount of money spent during a certain period of time. The RFM-model is a simple method that splits each RFM-variable into categories (subjectively chosen by the researcher) and assigns probabilities to each category of each characteristic in accordance with its differential response behavior (Bult & Wansbeek 1995). In addition, RFM information is inserted into predictive models. For example, RFM values can be used as explanatory variables in a probit or logit response model or an OLS regression model (Rao & Steckel 1995). Additional procedures drawn from the data mining literature, such as decision trees, neural networks and Bayesian networks, can also be used to link RFM values to buying behavior (Rhee & Russell 2008; Berry & Gordon 2000; Cui & Wong 2004; Cui et al. 2006)

However, a disadvantage of RFM-model is the limited number of selection variables used. Usually there are more household characteristics available than those used in the RFM-approach that have an effect on the probability of response (Bult & Wansbeek 1995). Another critical problem is the possible endogeneity of the RFM variables. RFM variables are based on past selection decision of firms and previous responses from households. For instance, if a household is not selected to receive a marketing offer (and the household has no way to respond to the offer otherwise), the recency will be larger and the frequency and the monetary value will be smaller, than the values of these same variables for a comparable household who received the solicitation. If the firm consistently ignores the household for any reason, the RFM values of this household will deteriorate regardless of the true propensity to respond. In this sense, the RFM variables may be endogenous and their parameter estimates biased due to the correlations between RFM variables and the error of the model (Rhee & Russell 2008; Cui et al. 2006).

2.2.2 Scoring methods

Individual scoring methods predict the probability of response or revenues of response per individual. There is a rich tradition in the database marketing literature of methods to rate customers by their expected response to marketing actions, such as logistic regression, neural networks and Bayesian networks (e.g., see Blattberg et al. 2001; Levin & Zahavi 2001; Cui & Wong 2004; and Cui et al. 2006). The probability models concern the binary classification problem, whether buy or not buy. Some of these methods are discussed hereafter.

Logit and probit models are developed to model such binary (1/0) process. The main difference between these two models is the way in which the disturbances ε are distributed. The logit model assumes they are logistically distributed and the probit model assumes the disturbances are normally distributed. The model assumes that every individual has a certain tendency to respond to a mailing received. This tendency is influenced by the explanatory variables. If the tendency is larger than zero, then the individual will respond, otherwise the individual will not respond. A drawback of the logit and probit models is that they assume symmetric costs of misclassification: the same weight is given to false negative and false positive misclassification errors. But the costs of misclassification are not symmetric. False negatives will be more expensive than false positive errors. Bult (1993) and Bult et al. (1996) used the asymmetric loss function to address this problem and a semi-parametric version of a logit model. Bult et al. (1996) expanded the asymmetric loss function by incorporating heterogeneity. A drawback of their model is that the different segments are determined a *priori* and revenue is modeled as the average of that from last year.

Tree generating techniques, such as Automatic Interaction Detection (AID), Chi-square Automatic Interaction Detection (CHAID) and Classification and Regression Trees (CART), have also been used to predict consumer responses. AID and CHAID determine for every available predictor the optimal split such that the within-group variance of the response is minimal. Variables with the lowest group variance are selected and subdivided. The sub lists are analyzed in the same way (Haughton & Oulabi 1997). An advantage of these methods is to avoid the double counting problem of the RFM model (Bult & Wansbeek 1995). CART results in a decision tree, where at each node there is a division. The groups are ultimately described as a combination of variables (group 1 has variable 1 smaller than x, variable 2 between y and z, etc) (Jonker et al. 2004). Magidson (1988) recommends not using AID, as this method only allows binary splits. CART and CHAID do not seem to differ much in performance, but CART is preferred when there are a large number of continuous variables, and one should opt for CHAID when there are many categorical variables (Jonker et al. 2004; Trasher 1991).

Artificial neural networks (ANNs), one such method that mimics human brain process, has been applied to modeling direct marketing responses. ANNs consist of many non-linear computational elements called nodes, and different nodes are arranged into different layers: the input layer, the hidden layer and the output layer. The task of neural network models is to determine relationships between the input (independent variables) and the output (dependent variables) by building network structures among them using the hidden layers (perceptrons) and hidden nodes (neurons) that resemble the human thinking process. ANNs are not subject to the assumptions of normality, linearity or complete data, thus are particularly useful for exploring complex models and noisy data. However, when Zahavi and Levin (1997) applied ANNs to modeling consumer responses to direct marketing, they found that the neural networks learned by the backpropagation method did not perform any better than logistic regression.

Besides these methods, topics on target selection have received much attention recently; see e.g. Banslaben (1992), Bult (1993), Bult and Wansbeek (1995), DeSarbo and Ramaswamy (1994), Magidson (1988), and Spring et al. (1999). For example, DeSarbo and Ramaswamy (1994) have proposed the Consumer Response-Based Iterative Segmentation Procedures (CRISP), which can simultaneously derive market segments and estimate models of customer response in each segment. By controlling for unobserved consumer heterogeneity among consumers, this model can help to improve the accuracy of classification. Several authors have tested a beta-logistic model that could update the estimated response probabilities over time and lead to more accurate predictions (Rao & Steckel 1995).

However, this stream of research on response modeling only considers response probability (yes/no). But a high response rate does not necessarily mean high profit. A direct marketing manager is also interested in profit maximization in addition to response maximization. To date, researchers have adopted various statistical methods to model revenues of response or profit, such as multiple regression. Many researchers have adopted a profit-maximization approach to identify the high profit consumers as well as low profit or unprofitable consumers so that precious marketing resources can be saved to augment profitability (Bult & Wansbeek 1995; Bitran & Mondschein 1996). Customers are selected until their predicted marginal revenue becomes zero.

Recently, researchers begin to take the amount of purchase into consideration. Simon (1987), for example, suggests taking the average amount of purchase from a random sample of the customers on the mailing list over a couple of years, and to use this as the expected value of a potential customer. Then the response to a positive reply is considered fixed as yet, and the response can be modeled again by a binary choice model. Rao and Steckel (1995) suggest using an ordinary least squares (OLS) model to determine the expected revenues and obtaining the total revenue as the product of the expected revenues and the probability of response. However, their empirical example is a binary choice model. Donkers et al. (2006) proposed a target selection rule for sending mailings, based on a model that jointly estimates incidence and quantity, while accounting for previous target selection. They used a probit model for response probability and a loglinear regression model for the amount of response.

Otter et al. (2006) jointly modeled both decisions and derived a number of profit maximizing selection methods. They empirically illustrated the methods using a data set from a charitable foundation. It appeared that modeling both aspects of the response yielded considerably higher profits relative to selection methods that were based on solely modeling the response probability. Although these researchers start to consider the amount of response, the exact way of incorporating revenue modeling into response modeling needs further research.

2.3 Customer Profitability Analysis

Recent rise in data-driven relationship marketing advocates the incorporation of “customer profitability analysis” for segmenting markets, allocating marketing resources and devising marketing mix strategies in a way that returns high levels of profits. The relationship marketing perspective advocates the emphasis on the more profitable customers.

The issue of customer profitability has attracted interest in both the management accounting and marketing literature. With the advent of activity-based costing in the 1990s, management accounting researchers have been interested in understanding the processes and factors that influence individual customer profitability, which can help vendors more effectively allocate customer management effort across customers and better target high-potential customers (Shields 1997). In the marketing literature, researchers have attempted to build customer profitability models in the direct marketing context (e.g., Beger & Nasr 1998; Mulhern 1999), in which customer profitability is evaluated solely on the transactions between the direct marketer and the customer.

Mulhern (1999) provided a demonstration of customer profitability in a business-to-business marketing situation involving pharmaceutical sales. The author argues that assessing the distribution of profitability is extremely important because it reveals the extent to which an organization depends on a small set of customers for its profits. Information on the distribution for concentration of profits can also be used for targeting marketing decisions. In the pharmaceutical sales example, the assessment of the distribution of customer profit can be made by observing a ranking of computed profits from highest to lowest (Figure 1). In such a situation a relatively small number of customers have a very high level of profit, while the balance of customers has a low, and in some cases, negative profit/loss. This type of distribution appears to be prevalent and business and consumer marketing contexts based on anecdotal evidence (Mulhern 1999).

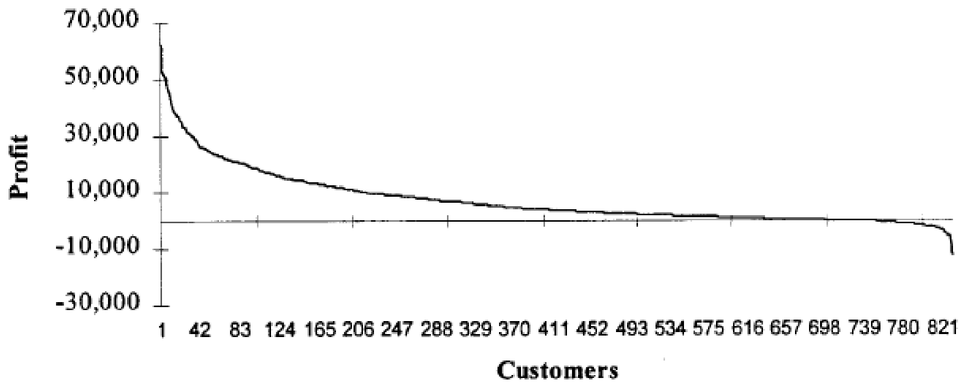


Figure 1 Customer Profit Ordering for Physicians: Highest to Lowest

Source: Mulhern, F.J. (1999). "Customer profitability analysis: Measurement, concentration, and research directions", Journal of Interactive Marketing, Vol. 13, No. 1, p. 32.

Assessing the distribution of customer profit can be done by observing a frequency distribution as shown in Figure 2. The frequency distribution shows a high frequency of low profit customers versus a low frequency of high profit customers generated. Figure 2 shows that a small group of customers generate a large portion of profit to the company. In general, customer profits exhibit long-tailed, skewed distributions with many outliers.

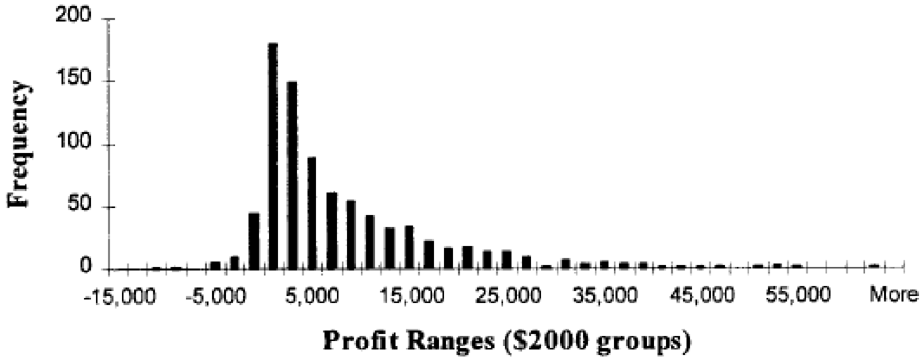


Figure 2 Frequency Distribution of Customer Profit for Physicians

Source: Mulhern, F.J. (1999). "Customer profitability analysis: Measurement, concentration, and research directions", Journal of Interactive Marketing, Vol. 13, No. 1, p. 33.

Another simple but useful tool for assessing the distribution of customer profit is to plot the cumulative percentile of customer profit against the cumulative percentile of the number of customers. Figure 3 shows this kind of plot for the physicians in Mulhern's (1999) empirical example. Note that the curve is quite bowed and surpasses the 100% line. It means each percentile to the right of the apex (representing about 15% of the customers in this example) reduces the overall profitability of the customer base. Profits are quite concentrated, as 20% of the customers account for 65.5% of the profits and half the customers account for 95.5% of the profits (Mulhern 1999). This further demonstrates that a large portion of profits comes from a relatively small set of customers and not all the customers are equally profitable. Thus, selecting the most profitable customers is extremely important to increase companies' profit.

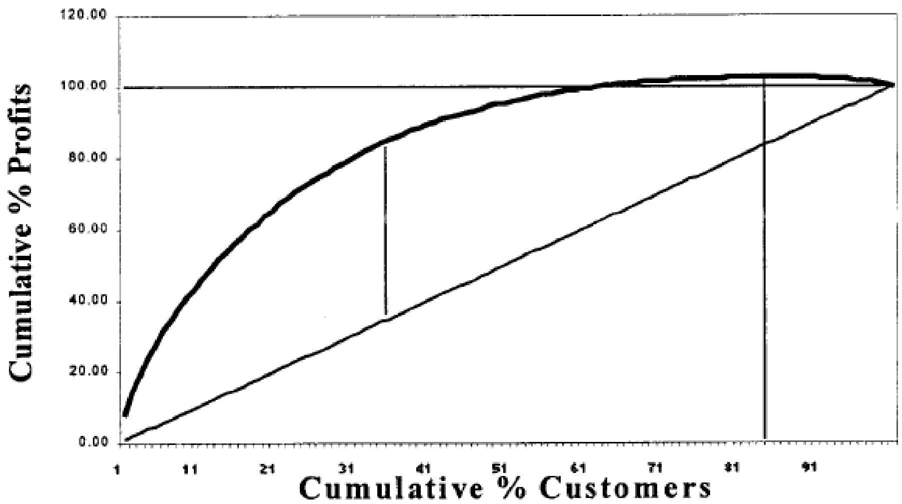


Figure 3 Inverted Lorenz Curve for Physicians

Source: Mulhern, F.J. (1999). "Customer profitability analysis: Measurement, concentration, and research directions", Journal of Interactive Marketing, Vol. 13, No. 1, p. 34.

The standard OLS regression is a common scoring approach used to obtain estimates for the conditional mean of some variable, given some set of covariates. Although very powerful, one drawback of this approach is that the estimate obtained is only one number used to summarize the relationship between the dependent variable and the independent variables. As demonstrated from Mulhern (1999), customer profit varies in great range across customers. This huge disparity may suggest heterogeneity of consumer response behavior. According to Jonker (2004), previous studies on target selection method do not take such heterogeneity into account, and it seems not very realistic that each individual will response in the same way, and hence it would have improved face validity if the methods had allowed for heterogeneity. Mulhern (1999) suggests further research to understand what factors determine the degree of disparity of profits across customers. Traditional OLS regression, which is under the conditional mean framework, is not able to measure and understand the disparity of profits across customers. In addition, OLS regression is influenced by the long-tailed, highly skewed profit distribution and is not capable of handing many outliers. Furthermore, the uniqueness of customer profit data usually violates several critical model assumptions of OLS regression. Thus the estimates by OLS regression may suffer serious bias. In the next subsection, we will discuss in detail the limitations of using OLS regression in customer profitability analysis.

2.4 Limitation of OLS regression in Customer Profitability Analysis

According to Hao and Naiman (2007), conventional regression analysis focuses on the population mean; that is, we summarize the relationship between the response variable and predictor variables by describing the mean of the response for each fixed value of the predictors, using a function referred to as the conditional mean of the response. Under ideal conditions, one of which is that the error is assumed to have precisely the same distribution whatever values may be taken by the values of the predictors, ordinary least square (OLS) regression is sufficient to describe the relationship between the covariates and the response distribution (Fitzenberger, Koenker & Machado 2002). Koenker and Hallock (2001) refer to this as a pure location shift model since it assumes that the predictor affects only the location of the conditional distribution of response variable, not its scale, or any other aspect of its distributional shape. An estimated model of such conditional mean function, supplemented perhaps by an estimate of the conditional dispersion of response variable around its mean, can be fully satisfied. In addition, using conditional-mean models leads to estimators (least squares and maximum likelihood) that possess attractive statistical properties, are easy to calculate, and are straightforward to interpret (Hao & Naiman 2007).

However, there are several limitations of OLS regression, especially in customer profitability analysis. First, when summarizing the response for fixed values of predictor variables, the conditional-mean model can not be readily extended to noncentral locations (Hao & Naiman 2007). For example, in customer profitability analysis, direct marketers have special interest in customers who generate high profits to companies (upper tail). The usual limited promotion marketing budget can not allow companies to target every customer who has a certain probability to respond, but can only allow managers to select the most profitable customers to maximize profit. OLS regression focuses only on the central location so that it is not sufficient to address the questions of target selection.

Second, the model assumptions of OLS regression are not always met in customer profitability analysis. In particular, the homoscedasticity assumption frequently fails. Focusing exclusively on central tendencies can fail to capture informative trends in the response distribution (Hao & Naiman 2007). Homoscedasticity refers to the assumption that the dependent variable exhibits similar amounts of variance across the range of values for an independent variable. In customer profitability analysis, it means that for different values of x, the disparity is the same across customers of different levels of profit. However, as we can see from our previous demonstration of the distribution of customer profits, the disparity varies greatly across customers. The homoscedasticity assumption can not hold up in this type of analysis.

Furthermore, heavy-tailed distributions, as shown in Mulhern (1999), commonly occur in customer profitability analysis, leading to a preponderance of outliers. A related assumption made in OLS regression is that the regression model used is appropriate for all data, which is called the one-model assumption. The conditional mean can then become an inappropriate and misleading measure of central location because it is heavily influenced by outliers (Hao & Naiman 2007). The traditional OLS regression is to identify outliers and eliminate them. However, outliers and their relative positions to those of the majority are important in customer profitability analysis. According to the 80/20 rule, the very small group of customers who generate a large portion of profit are very important so that eliminating them from analysis not only can not help reach companies' goal of maximizing promotion profit but also result in loss of valuable customers and damage companies' long-term interest. Hao and Naiman (2007) argue that in terms of modeling, OLS regression can not simultaneously model the relationship for the majority cases and for the outlier cases. Thus, OLS regression is inefficient to accomplish direct marketers' goal.

Fourth, another distinctive feature of OLS regression is the normality assumption. As shown by the histogram of customer profit distribution in Mulhern (1999), customer profit exhibits a long-tailed, heavily skewed distribution. This confirms that customer profit is not normally distributed. According to Hao and Naiman (2007), hypothesis testing of OLS regression that whether an explanatory variable significantly affects the dependent variable requires determination of the sampling variability of estimators. Calculated p-values rely on the normality assumption or large-sample approximation. Violation of these conditions may cause biases in p-values, thus leading to invalid hypothesis testing. In addition, calculation of confidence intervals for OLS regression predictions requires the assumption that the errors have a normal distribution with mean zero and variance σ^2. This means that if repeated measurements of *y* are taken for a particular value of *x* then most of them are expected to fall close to the regression line and very few to fall a long way from the line. However, as we demonstrate previously, the errors do not have a constant variance (heterogeneity) or a normal distribution, as identified from the distribution of response variable. This means that the estimated OLS regression would not be an appropriate or effective prediction of customer purchase behavior and hence can not capture the high profitable customers whom marketers want to target.

Fifth, according to Hao and Naiman (2007), the focal point of central location has long steered researchers' attentions away from the properties of the whole distribution. It is quite natural to go beyond location and scale effects of predictor variables on the response and ask how changes in the predictor variables affect the underlying shape of the response distribution. In customer profitability analysis, for the given values of *x*, the conditional mean spending of customers may not be the interest of managers. Companies may want to know the whole conditional spending distribution of customers, such as the small spenders (lower tail) and big spenders (upper tail) of the conditional distribution. The central location, the scale, the skewness, and other higher-order properties--not central location alone—characterize a distribution. Thus, conditional-mean models are inherently ill-equipped to characterize the relationship between a response distribution and predictor variables (Hao & Naiman 2007).

Due to the above limitations, using OLS regression for customer profitability analysis is highly problematic and questionable. Hence, an alternative method is the main urgency of direct marketing researchers. The central research question is whether the profitable customers show different characteristics from the less profitable customers in terms of their lifetime values, marketing responses, and demographic variables. It has been suggested that consumer heterogeneity should be recognized by allowing the model's parameters to vary across the population, as it is difficult for a single model to represent multiple distinct consumer segments (Blattberg & Sen 1976). There is no shortage of techniques for modeling consumer heterogeneity.

Researchers and practitioners have long dreamed of having separate models for separate segments to uncover the unobserved heterogeneity among consumers (Libai, Narayandas & Humby 2002; Shepard 1999). Conventional methods include applications of several non-parametric methods such as cluster analysis, discriminant analysis, automatic interaction detectors (e.g., CHAID). To determine whether the strength of the relationship differs from segment to segment, Shepard (1999) suggests coding a predictor variable into several dummy variables. For instance, instead of using age, one may use "young," "middle-aged," and "elder" as well as their interactions with other predictors to see if the parameters are different in their relationship to direct marketing response. Mulhern (1999) proposed a segment-based approach to customer profitability analysis. Libai, Narayandas and Humby (2002) proposed a stochastic segmentation method. However, these solutions may result in the loss of information. This study proposes to use a more direct approach to customer profitability analysis – quantile regression. In the next chapter, we will discuss the basic theory of quantile regression (QR), the empirical application of quantile regression in various areas, and its suitability for customer profitability analysis.

3 QUANTILE REGRESSION

3.1 Introduction

Ordinary regression analysis is undoubtedly the most popular and well-known statistical technique and is at the heart of many other statistical techniques as well. In ordinary regression, the relation between a dependent variable and a set of explanatory variables is described by the conditional expectation function. An alternative to conditional-mean modeling can be traced back to mid-18th century, an approach called conditional-median modeling. Median regression, least absolute deviation (LAD), can replace ordinary least squares (OLS). The median-regression model can be used to achieve the same goal as conditional-mean-regression modeling: to represent the relationship between the central location of the response and a set of covariates. When the distribution is highly skewed, the mean can be challenging to interpret while the median remains highly informative (Hao & Naiman 2007). However, the mean and median tell little about the other parts of the distribution, such as the tails. As far as the entire conditional distribution is concerned, it is not satisfactory to characterize only the conditional mean and/or median behaviors.

Koenker and Bassett (1978) proposed quantile regression, permitting estimating various quantile functions of a conditional distribution, among which the median (50th quantile) function is a special case. Each quantile regression characterizes a particular (center or tail) point of a conditional function; combining different quantile regression thus provides a more complete description of the underlying conditional distribution.

In analogy with classical linear regression methods, based on minimizing sums of squared residuals and meant to estimate models for conditional mean functions, quantile regression methods are based on minimizing asymmetrically weighted absolute residuals and intended to estimate conditional median functions and a full range of other conditional quantile functions. The basic motivation for using quantiles rather than simple mean regression is that the stochastic relationship between random variables can be portrayed much better and with much more accuracy (Buhai 2005). Koenker and Bassett (1978) argue that conventional least squared estimators may be seriously deficient in linear models constructed on some non-Gaussian settings, where quantile regression would provide more robust and consequently more efficient estimates. Actually, quantile regression is particularly useful when the conditional distribution does not have a “standard” shape, such as an asymmetric, fat-tailed, or truncated distribution. In the next section, we will discuss the basic theory underlying quantile regression.

3.2 Basic Model and Interpretation

Various reserchers have provided comprehensive introductions of quantile regression. Here, we adopt Buhai (2005)'s work on the overview of quantile regression. For the technical part of quantile regression, please refer to Koenker (2005) and Koenker & Bassett (1978).

3.2.1 Basic Model

The quantile regression model introduced by Koenker and Bassett (1978) extends the notion of ordinary quantiles in a location model to a more general class of linear models in which the conditional quantiles have a linear form. To briefly recall the ordinary quantile, consider a real valued random variable Y characterized by the following distribution function

$$F(y) = \Pr(Y \le y) \tag{1}$$

Then for any $\tau \in (0, 1)$, the τ -th quantile of Y is defined as follows:

$$Q(\tau) = \inf\{y : F(y) \ge \tau\} \tag{2}$$

The median is then $Q(1/2)$, the first quartile $Q(1/4)$ and the first decile $Q(1/10)$. The quantile function provides a complete characterization of Y, just like the distribution function F. The quantiles can be written as solutions to the following optimization problem: for any $\tau \in (0, 1)$, define the piecewise linear "check function"

$$\rho_\tau(u) = u(\tau - I(u < 0)) \tag{3}$$

where I(.) is the usual indicator function. The solution to the minimization problem is then

$$\hat{\alpha}(\tau) = \arg\min_{\varepsilon\in\Re} E[\rho_\tau (Y - \xi)] \tag{4}$$

The sample analogue of $Q(\tau)$ is based on a random sample $\{y_1,..., y_n\}$ of Y. The τ -th quantile can then be identified, in the spirit of (4) above, as any solution to:

$$\hat{\alpha}_\tau = \arg\min_{\xi\in\Re} \sum_{i=1}^{n} \rho_\tau (y_i - \xi) \tag{5}$$

Let x_i, i = 1...n, a K×1 vector of regressors. We can then write the equivalent of expression (1) as:

$$F_{u_\tau} (\tau - x_i^{'} \beta_\tau \mid x_i) = \Pr(y_i \le \tau \mid x_i) \tag{6}$$

which is essentially a different form derived from the more familiar:

$$y_i = x_i^{'} \beta_\tau + u_{\tau_i} \tag{7}$$

where the distribution of the error term u_{τ_i} is left unspecified, the only constraint being the (usual) quantile restriction $Q_\tau (u_{\tau_i} \mid x_i) = 0$.

Using as analogy the estimation of conditional mean functions as in

$$\hat{\beta} = \arg\min_{\beta\in\Re^K} \sum_{i=1}^{n} (y_i - x_i^{'} \beta)^2 \tag{8}$$

the linear conditional quantile function

$$Q_Y (\tau \mid X = x) = x_i^{'} \beta_\tau \tag{9}$$

can be estimated by solving the equivalent of expression (8) for this case:

$$\hat{\beta}_\tau = \arg\min_{\beta\in\Re^K} \sum_{i=1}^{n} \rho_\tau (y_i - x_i^{'} \beta) \tag{10}$$

3.2.2 Interpretation

Buchinsky (1998) and Buhai (2005) provide detailed interpretation of quantile regression estimation. A least squares estimator of the mean regression model would be concerned with the dependence of the conditional mean of Y on the covariates X. The quantile regression estimator tackles this issue at each quantile of the conditional distribution, providing thus a more complete description of how the conditional distribution of Y given X = x depends on x. In other words, instead of assuming that covariates shift only the location or the scale of the conditional distribution, quantile regression looks at the potential effects on the shape of the distribution as well (Buhai 2005). Then, how can the quantile's coefficients be interpreted? According to Buchinsky (1998), consider the partial derivative of the conditional quantile of *y* with respect to one of the regressors, say *j*. The derivative is to be interpreted as the marginal change in the *t*-th due to marginal change in the *j*-th element of *x*. If *x* contains K distinct variables, then this derivative is given simply by the coefficient on the *j*-th variable, β_j. Caution is required in interpreting this result. It does not imply that a person who happens to be in the τ-th quantile of one conditional distribution will also find himself/herself at the same quantile had his/her *x* changed. Actually, the interpretation of quantile regression coefficients is analogous to that of OLS regression coefficients. In the case of a continuous covariate, the coefficient estimate is interpreted as the change in the quantiles of the response variable corresponding to a unit change in the predictor (Hao & Naiman 2007).

3.3 Advantages

First, a quantile regression model can be used to characterize the entire conditional distribution of a dependent variable given a set of regressors. Since multiple quantiles can be modeled, it is possible to achieve a more complete understanding of how the response distribution is affected by predictors, including information about shape change. A set of equally spaced conditional quantiles (e.g., every 5% or 1% of the population) can characterize the shape of the conditional distribution in addition to its central location. Researchers can also choose positions that are tailored to their specific inquiries (Hao & Naiman 2007).

Second, like the LAD, the quantile regression objective function is a weighted sum of absolute deviations, which gives a robust measure of location, so that the estimated coefficient vector is not sensitive to outlier observations on the dependent variable (Buchinsky, 1998). Buhai (2005) suggests that the estimate and the inference process have an inherent distribution-free character given that quantile estimation is influenced only by the local behavior of the conditional distribution of the response near the specified quantile. The signs of the residuals are the only thing that matters in the determination of the estimates and thus outliers in the values of the response variables influence the fit so far as their being above or below the fitted hyperplane, but how far below or above is irrelevant. Hao and Naiman (2007) suggest that quantile regression is robust to distributional assumptions because the estimator weighs the local behavior of the distribution near the specific quantile more than the remote behavior of the distribution. The quantile regression model's inferential statistics can be distribution free.

Third, OLS regression model has an assumption that the dependent variable exhibits homogeneous variance across the range of values for an independent variable. In the absence of quantile effects, the OLS model is a special case of quantile regression model (Ma & Pohlman 2008). All the regression quantile slope estimates are for a common parameter, and any deviation among the regression quantile estimates is simply due to sampling variation. But, when the predictor variables X exert both a change in means and a change in variance on the distribution of y, we have a regression model with unequal variance (heterogeneity). As a consequence, changes in the quantiles of y across X can not be the same for all quantiles. Slope estimates differ across quantiles because the parameters differ, since the variance in y changes as a function of X (Cade & Noon 2003). An advantage of using quantile regression to model heterogeneous variation in response distributions is that no specification of how variance changes are linked to the mean is required. Furthermore, changes in the shape of the distributions of y across the predictor variables can also be detected (Koenker & Machado 1999). In the condition of heterogeneity, quantile regression will provide a more complete view of the relationship between variables through the effects of independent variables across quantiles of the response distribution (Ma & Pohlman 2008). Potentially different solutions at distinct quantiles may be interpreted as differences in the response of the dependent variable to changes in the regressors at various points in the conditional distribution of the dependent variable (Buchinsky, 1998).

Fourth, quantile regression models can be easily fit by minimizing a generalized measure of distance using algorithms based on linear programming. As a result, quantile regression is now a practical tool for researchers. Software packages familiar to researchers offer readily accessed commands for fitting quantile-regression models (Hao & Naiman 2007).

In the next section, we will briefly discuss the development of quantile regression and its empirical applications in various areas.

3.4 Development and Empirical Applications

Since the first introduction of quantile regression by Koenker and Bassett (1978), the statistical theory and computational routines for estimating and making inferences on regression quantiles have been developed, especially for the linear model (Koenker & Machado 1999), but also for parametric nonlinear (Koenker and Park 1996) and nonparametric, nonlinear smoothers (Koenker et al. 1994, Yu & Jones 1998).

Empirical researchers took advantage of quantile regression's ability to examine the impact of predictor variables on the response distribution. Quantile regression has been widely employed for instance within labor or educational economics to study wage determinants, discrimination effects, transition or duration data, trends in income inequality or effects of socioeconomic characteristics and policy variables on educational attainment (Fitzenberger 1999, Machado & Mata 1999, Eide & Showalter 1998, Mueller 2000, Koenker & Billias 2001, Viscusi & Hamilton 1999). Quantile regression methods have also been used lately in micro-demand analysis (Deaton 1997) and there even seems to be a growing literature using quantile regression in empirical finance and particularly, on value at risk (Taylor 1999, Chernozhukov & Umantsev 2001). Using the census data to study of skills returns (i.e. the effect of education on wage increases), for instance, Buchinsky (1998) estimated a log regression at five interesting quantiles of the response probability, namely .10, .25, .50, .75, and .90 quantiles. He found very little increases at the lower part of the distribution while the mean of weekly earnings is consistently above the median of weekly earnings, indicating that the wage distribution is right-skewed, and more so in the latter years (Buchinsky 1998). In addition, he found that a much steeper increase in the return to education occurs at the higher quantiles of the distribution 1992 than in 1979. Quantile regression (QR) analysis reveals that the effect of education is not constant across the conditional wage distribution. They are higher for those at higher quantiles in the conditional wage distribution. Similar findings have been achieved in studies of birth weights. Clearly, it is not enough to investigate changes in the mean when the entire shape of the distribution changes dramatically.

In finance, researchers have considered regression quantile modeling in value-at-risk (VaR), portfolio returns and related problems. Chernozhukov and Umantsev (2001) model the conditional VaR in terms of the regression quantile function—the inverse of the conditional distribution function. In the empirical application of the proposed model, the study estimates and analyzes the conditional market risk of an oil producer stock price as a function of the key economic variables. The study finds that these variables impact various quantiles of the return distribution in a very differential and nontrivial manner. The study characterizes the key determinants of the extremal and intermediate conditional risk and finds that the market index is the only statistically significant determinant of the extremal risk. Basset and Chen (2001) introduce quantile regression as a complement to the least squares regression by identifying the impact of style on the conditional return distribution at places other than the expected value. The regression quantiles extract additional information from the time series of returns, allowing discrimination among portfolios that would be otherwise judged equivalent based on conditional expectations. The results of the study show how the conditional return distribution can respond to factors in different ways at alternative parts of the return distribution.

In this study, we propose to use quantile regression to model customer response to direct marketing. In the next subsection, we discuss quantile regression to customer profitability analysis, including its suitability and relevant empirical applications.

3.5 Quantile Regression for Customer Profitability Analysis

3.5.1 The Suitability

The quantile regression is particularly applicable to solving many marketing problems when data do not conform to the normality assumption, as it is the case of the distribution of customer profit data. In fact, several previous studies of segments already implicitly recognize the quantile characteristics of profit distribution, by separating customers into quartiles or quintiles (Campbell & Frei 2004; Zeithaml, Rust, & Lemon 2001). Profit data, which depart from the normality assumption, have been known to be skewed on the right with a long tail (Mulhern 1999). Such departure from normality, especially in the tails, can affect estimates of parameters by OLS regression, which may either over-estimates or under-estimates the effect of variables, in this case, on customer profitability. Thus, quantile regression has several distinctive advantages for analyzing customer profitability.

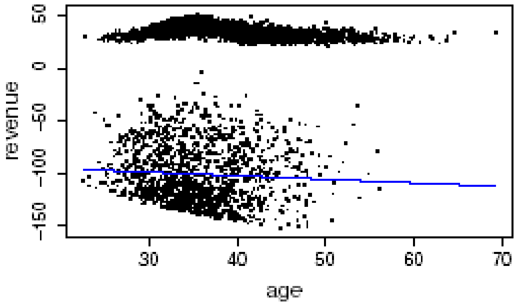
First, as profit data always exhibit long-tailed, skewed distribution with lots of outliers, OLS regression estimate is highly problematic in such situation. Quantile regression is insensitive to outliers and to the violation of normality model assumption, and thus is a robust estimation of customer profitability. Second, as profit always varies greatly across customers, indicating potential heterogeneous variance, OLS regression is not able to capture the different effects predictors exert on the response variable across different quantiles. Quantile regression naturally can model whole conditional response distribution, and thus provides a complete picture of the relationship between predictor variables and response profit. Third, direct marketers always have special interest in customers of different segments, such as low quantile or high quantile, but OLS regression only concerns the mean of profit, which is not a quite informative measure of customer profits. Quantile regression is able to model conditional quantile function of any specific quantile, such as the upper tail of a response distribution.

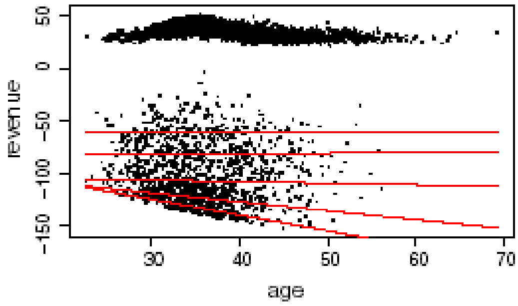
3.5.2 Relevant Applications

Through reviewing previous literature, we find that quantile regression can be very useful for marketing because the effects of covariates may be different across conditional quantiles (Chaudhuri & Loh 2002). Marketers can perform simultaneous quantile regressions to compare different customers groups, for instance, to compare the price elasticity of consumers who are very satisfied vs. those who are not (Dotson et al 2007), to estimate customer wallets (Rosset et al 2005; Perlich et al. 2007). A very recent research by Somers and Whittaker (2007) applied quantile regression in two retail credit risk assessment applications, which were to predict loss given default for secured loans and to the application of revenue modeling. In their second study, as shown in Figure 4a, the panel that displays the scatter plot of age and revenue exhibits two distinct clusters in the 5,000 accounts. The group ending the period with positive revenues has a profile that exhibits a middle aged bulge. The group with negative revenues, the defaulters, is more dispersed and show different behaviors. The accounts are segmented into the defaulters and the non-defaulters with negative and positive revenues respectively. For defaulters, the fitted values from a linear least squares regression of loss on age form the line in the panel of Figure 4b and exhibit an increasing average loss with age. However, the quantile regression on the median and higher percentiles (0.5, 0.75, 0.9), Figure 4c, shows almost no variation of loss with age. We may conclude linear least squares regression obscures some important details.

Although quantile regression has been used in marketing literature, most of them concentrate on estimating the distribution on conditional quantiles or fitting the quantile regression model to marketing data. Few researchers have dealt with forecasting or prediction using quantile regression, especially in the direct marketing context. Using quantile regression for predicting and forecasting purposes remains a significant challenge for researchers. In customer profitability analysis, forecasting both the reactions of high quantile customers to changes of predictor variables and the spending distribution of each customer is equally important. Hence, in the next section, we put much emphasis on quantile regression forecasting.

a





b **c**

Figure 4 Quantile Regression Estimation

Source: Somers, M. and Whittaker, J. (2007). “Quantile regression for modelling distributions of profit and loss”, European Journal of Operational Research. Vol. 183, No. 3, p. 1485

4 QUANTILE REGRESSION FORECASTING

4.1 Relevant Applications

Compared with ordinary least square regression, quantile regression has been found good at explaining the differences among customer groups at different quantiles (Dotson et al 2007; Rosset et al. 2005; Perlich et al. 2007; Whittaker et al. 2005; Somer & Whittaker 2007). Note that different customers have diverse reactions to the same marketing campaign. It is crucial to identify and target the most profitable customers, who will contribute most profit to the company. With the target selection information, marketing manager will be more efficient on making decision on which customers should receive a catalogue in a marketing campaign.

In order to target the most profitable customers, it is important to predict the reaction of different customers to marketing campaign. Though quantile regression has been used in the estimation of different quantiles of customers profit distribution, it has seldom been used in the prediction or forecast of customer responses to marketing campaign. Few researchers have studied the prediction accuracy of quantile regression. Through literature review, we found that quantile regression had been used for forecasting in the field of finance, weather, and wallet-estimation recently. Reviewing relevant literature gives us more insight into quantile regression forecasting.

In finance, application for quantile regression forecasting was to generate the prediction intervals for the financial market (Taylor & Bunn 1999). Taylor (2007) undertook another study to forecast daily supermarket sales by constructing interval forecasts from quantile predictions generated using exponentially weighted quantile regression. The empirical results are encouraging, with improvements over traditional methods being particularly apparent when the approach is used as the basis for robust point forecasting. Whittaker et al. (2005) applied quantile regression to the analysis of a large credit scoring database of clients who have missed a single payment. Their findings suggest that different covariates are needed to provide explanation and prediction of the target variable depending on the quantile of interest. A substantive finding was that the predictor of the median of the target variable contains different variables from those of the predictor of the 30% quantile. Somers and Whittaker (2007) applied quantile regression in two predictions. The first one is to predict loss given default for secured loans, in particular retail mortgages, based on estimation of the more relevant low tail of the house value rather than the average value. The first study tailors the prediction to the characteristics of particular account, but the calculation is still an expected value. The second study uses standard linear quantile regression to estimate revenue model and kernel smoothed quantile regression in prediction. The authors conclude that forecasting distribution using quantile regression makes a difference when the distribution is highly non-normal and understanding the changes in the dispersion around the mean is useful.

In weather prediction, a notable study by Bremnes (2004) made reliable precipitation forecasts in terms of quantiles. In this paper, the author employed a approach, which has essentially two steps: (1) estimation of probability of precipitation and (2) estimation of selected quantiles in the distribution of precipitation amounts given occurrence of precipitation. Estimates are obtained by means of probit regression and local quantile regression, respectively. By applying the laws of probability, the steps are combined to make unconditional quantile forecasts. This is similar to direct marketing, where the probability of customer response and the purchase amount of each response are both of interest to managers. Bremnes (2004) conducted another study on the probabilistic wind power forecasts using local quantile regression. It can be shown that, for some purposes, forecasts in terms of quantiles provide the type of information required to make optimal economic decisions. Sousa et al. (2008) use quantile regression to forecast next day hourly ozone concentrations, considering simultaneously the statistically significant explanatory variables. The study finds that quantile regression model was more efficient than multiple linear regression to predict high and low extreme values.

Although many studies in finance, weather, and environment have adopted quantile regression in forecasting, most of these studies apply to time-series analysis, one step ahead forecasting and use nonparametric quantile regression, such as local linear quantile regression. In direct marketing, most of the profit data is cross-sectional. In this study we attempt to use parametric linear quantile regression and propose our application of linear quantile regression to direct marketing forecasting.

4.2 Quantile Regression for Direct Marketing

Studies adopting quantile regression forecasting in marketing domain are rare. To the best of our knowledge, there are only two studies. In a study of wallet estimation, Perlich et al. (2007) first define that a good customer to IBM is one whose spending with IBM is at a high percentile of its spending distribution. Under this definition, the authors seek to build models, predict such a high quantile (0.8) of the conditional spending distribution of each customer, and make evaluation of the predictions. Their linear quantile regression analysis for three major product brands shows the most consistently good performance for a quantile of 80 percent, compared to other methods including quantile k-nearest neighbor and quantile regression tree. These three proposed quantile regression models are also applied to four different dataset, including datasets of adult income, California house value, KDD-Cup 1998 donation campaigns, and IBM wallet. Quantile modeling approaches perform better than all standard methods (linear regression, CART, and kNN). In particular, linear quantile regression performs the best on KDD98 and IBM datasets. The authors argue that it is because of the presence of an independent variable which is highly linearly correlated to the response (last year's donation, and last year's IBM sales, respectively).

The second study is by Benoit and Van den Poel (2009). The authors propose to analyze customer lifetime value (CLV) by quantile regression. The study finds that the effect of some covariates has an opposite sign for different parts of the conditional life time value distribution, while the OLS estimate shows no significant effect. The study also seeks to predict a given top percentage of customers whom managers want to target the most. Based on the combination of a lifetime value point estimate and the size of the prediction interval, the study proposes a new segmentation scheme which takes uncertainty or risk into consideration. However, this study has a common misunderstanding of quantile regression prediction. A median regression forecast can not predict the CLV of median customers but only predict the median CLV of all customers.

In our study, the main research problem is to predict customer reaction to direct marketing campaigns and seek to improve customer profitability. An essential research question is whether customer characteristics and marketing variables have the same effect on various groups of customers at different levels of profitability, or whether customers at the different levels of profitability exhibit similar characteristics in their background variables and responses to marketing activities. Such knowledge can help marketers as they attempt to augment the profitability of marketing operations by focusing on the so-called high profit customers. Since targeted marketing and CRM activities are based on the premise of the concentration of profit among a group of customers, e.g., 20% of the customers account for 80% of the profit, we can apply quantile regression to shed light on this group. Theoretically, we can choose any percentile point of the profit data distribution, such as 90%, 80% or 70%. Here, we focus on the new analytical approach to customer profitability analysis with cross-sectional data in a single period. But this method can be extended to the analysis of multiple periods.

Customer profitability analysis in marketing is primarily concerned with market segmentation and resource allocation decisions – maximize the return on the limited resources. Thus, we propose two applications for quantile regression to customer profitability analysis. First, based on the customer profitability analysis, we can perform resource allocation—to select customers from the database—to maximize the profit for the next market campaign. This can be achieved by cross-validation of the estimates parameters on an unseen dataset. A good way of accounting for the relative definition of "good" customers is to treat this as a predictive modeling problem, and model the τ-th quantile of the conditional spending distribution of y (future profit) given x (demographics and relationship variables). Given a set of estimation data, linear quantile regression identifies the linear model in x which minimizes the empirical quantile error. It should be noted that due to issues such as over-fitting and bias, it is not guaranteed that using the "correct" loss function for modeling will result in a good prediction model for future data with regard to the same loss function. It can be shown, however, that this loss function is "consistent", i.e., with infinite training data and correct model specification it will indeed result in the τ -th quantile of $P(y|x)$ as the model predictions.

The second application is segmentation or profile analysis of the profitable customers. A very important objective of customer profitability analysis is to identify the most profitable customers and distinguish them from the others. By adopting unconditional quantile regression, we can analyze the characteristics of the most profitable customers, say the top 20% based on the 20/80 split. We can also extend this analysis to multiple quantiles. After choosing the quantile of interest, researchers can perform unconditional quantile regression to determine the marginal effects that predictors exert on the response profit of the unconditional quantile of the distribution. For example, if we want to determine whether increasing marketing density would increase customers response profits for the unconditional 90% of the population, we can simply perform 90% unconditional quantile regression. In the next section, I will discuss in detail the unconditional quantile regression.

4.3 Unconditional Quantile Regression

Please note that all the details before this section refer to conditional quantile regression. Next, we will discuss the difference between conditional and unconditional quantile regression.

It is well known that regression models establish conditional relationships between a response variable *Y* and a set of explanatory variables *X*. However, many questions of economic and policy interest and business decisions concern the influence of *X* on the unconditional statistics of *Y*. For instance, one would like to know what the impact of a one-year increase in education is on earnings in a given population that contains individuals with different characteristics (unconditional effects), rather than the impact just for a subgroup with specific covariates (conditional effects). As far as the mean is concerned, the unconditional properties of *Y* can be easily obtained by averaging it over *X*. This is because linear regression models have a classical property, i.e., conditional mean model $E(Y \mid X) = X\beta$ leads immediately to $E(Y) = E(X)\beta$ .This convenience hinges on the linearity of the expectation operator and hence cannot be generalized to cases with nonlinear operators such as quantiles. Thus, conditional quantile regression models cannot answer questions about the unconditional statistical properties of the response variable *Y*. The quantile regression method developed by Koenker and Bassett (1978), which is commonly used in past research, is in fact a conditional quantile regression.

In contrast, the Recentered Influence Function—RIF method proposed by Firpo et al. (2007a) is an unconditional quantile regression, and enables investigators to obtain the unconditional quantile partial effects which are more frequently of interest in economics. Central to the RIF unconditional quantile method is an influence function. The influence function is a widely used tool in robust statistics. As its name suggests, it represents the influence of an individual observation on a distributional statistic of interest such as a quantile. The recentered influence function (RIF) is basically a linear approximation (the leading terms of a von Mises expansion) to the nonlinear function of distributional statistics of interest such as a quantile. It essentially captures the change of the distributional statistic of interest, such as a quantile, in response to a change in the underlying distribution. The RIF regression is a function, $E[RIF(Y;v) \mid X = x]$. By taking iterated expectations, one can derive the marginal effects of the covariates on the statistic of interest by averaging the RIF regression function with respect to the change in the distribution of the covariates (see Theorem 1 in Firpo et al., 2007b). Analogous to the OLS regression, the RIF regression function typically assumes a linear specification,

$$E[RIF(Y;q_\tau) \mid X] = X\beta \,, \tag{11}$$

where the coefficient β represents the marginal effect of X on the distributional statistic, quantile q_τ.

Firpo et al.(2007a, 2007b) have given the mathematical proof of the unconditional property of the RIF regression. Readers with technical interest may refer to Firpo et al. (2007a, 2007b) for details. One can simply compare the RIF regression with the OLS regression. While the RIF regression has the same unconditional property as the OLS regression, the RIF regression is more general as it applies to any distributional statistic such as a quantile, not just the mean. In fact, a simple proof shows that the RIF regression associated with the mean statistic is identical to the OLS regression (Firpo et al., 2007b).

4.4 The Proposed Study

When customers are facing with a direct marketing promotion, there are two decisions to make. The first one is whether to buy or not, and the second one is how much money to spend on this purchase. In this thesis, we propose to model the response probability by latent class analysis of a logistic regression model in the first stage. We then model the profit by quantile regression on the second stage. Conditional and unconditional quantile regressions are both used to obtain parameter estimates. Using the estimated conditional quantile regression models from training sample, we extrapolate them onto the testing sample and calculate the predicted profit values on individual level across different quantiles. We conduct 10-fold cross-validation to minimize the sample variation of prediction results. The models performance of OLS regression and quantile regression are compared in decile analysis. The cumulative profit lift is averaged to obtain more reliable results.

For direct marketing, due to budget constraints, typically only the names in top two deciles or the 80^{th} percentile (those with the highest profit) will receive a catalog (Berger & Magliozzi 1992). Thus, the cumulative profit lift in the top two deciles of the file (testing data sets) is used to compare the performance of these models. It is the ratio * 100 of the number of true positives (TPs) in a decile identified by the proposed model versus the number of TPs identified by a random model, which is the number of TPs divided by the number of deciles (10). For instance, a model with a top decile lift of 200 is said to perform twice as well as a random model.

In 10-fold cross-validation, the cases are randomly divided into 10 mutually exclusive test partitions of equal size. The cases not found in each test partition are combined as the training data (10-1), and the resulting model is tested on the test partition. The procedure is repeated for 10 times (partitions). The average cumulative profit lift over all 10 partitions is the cross-validated cumulative profit lift. Ten-fold cross-validation can be used to achieve out-of-sample repeatability for a given level of statistical accuracy.

5 DATA ANALYSIS

5.1 Data

In light of mounting competition and increasing customer saturation, research in direct marketing has been seeking more effective ways of improving customer profitability for their promotion programs. We use a direct marketing dataset to test the quantile regression approach. The data come from a U.S.-based catalog direct marketing company that sells multiple product lines of general merchandise ranging from gifts and apparel to consumer electronics. The company sends regular mailings to its list of customers. The records consist of information of 106, 284 consumers, who were selected in a recent promotion, as well as their purchase history over a twelve-year period. This study focuses on the profitability of the most recent campaign promotion, which lasted for six months and achieved a 5.4% response rate. Since this promotion is a perennial event, the database contains the lifetime variables of the customers and the information on the most recent promotion.

5.2 Variables

The dependent variable is the amount of the profit/loss before the fixed operating cost for this promotion. To achieve realistic and consistent measure of customer profitability, company fixed cost not associated with the direct marketing campaigns were excluded from the calculation of the customer profit. Since the database contains 267 variables, we include several groups of variables to analyze customer profitability. In this study, we propose to adopt the quantile regression approach to solve a direct marketing problem and examine the effects of four groups of variables: 1) RFM factors, 2) lifetime value variables, 3) marketing intensity, and 4) consumer demographic and other background factors. First, given that RFM and customer lifetime value have been an important concept in direct response marketing for many years, our study does not focus on the conceptualization or the measurement of RFM or life-time value (LTV) of customers but assessing their impact on customer profitability. Second, while most of the research efforts in customer lifetime value have been concerned with the long-term profitability of consumers, it is less known that how lifetime variables can help with marketing planning or decision making for a single marketing campaign. This study examines the effect of lifetime duration on the profitability of a single marketing campaign. Third, marketing intensity is an important variable for assessing the impact of consumer responses. Marketers often rely on repetitive promotions to increase the probability of consumer response. That is also true for direct marketers, which may also mail multiple catalogs to the same customer. This is an important variable for assessing the effect of marketing spending on customer profitability. Fourth, consumer demographics such as income, household characteristics are known to influence consumer responses. These factors account for the consumer heterogeneity and provide managerial insights for firms attempting to customize response functions at the individual level. Cui et al. (2006) has tested the explanatory power of these four groups of variables on consumer response. This thesis seeks to examine the explanatory power of these four groups of variables on customer profitability. Table 1 is the list of variables selected for data analysis.

Table 1 Variables for Data Analysis

Variables Groups	Labels	Data Description
Dependent Variable	Targpbfo	Customer Profit
RFM Variables (After Endogeneity Bias Correction)	pred_m1	Predicted Monetary Value
	resid_m1	Residual of Monetary Value
	pred_r1	Predicted Recency
	resid_r1	Residual of Recency
	pred_f1	Predicted Frequency
	resid_f1	Residual of Frequency
Lifetime Variables	prord85	Lifetime Orders Prom 85
	salflg	Dollar class of avg ord last year
	salcat	Dollar class of lifetime avg ord
Transactional Variables	hcrd	Used House Credit Card (Y/N)
	cash	Cash Order (Y/N)
	tele	Bought Telemktng Prom (Y/N)
Marketing Intensity	crcpr85	Circs Prom 85 - Lifetime Contacts
Probability (For Sample Selection Bias Correction)	pred_dep	Predicted Probability

Table 1 lists five groups of variables. The first group includes RFM variables. The original RFM variables are Recency, Frequency, and Monetary Value. This thesis corrects the potential endogeneity of RFM variables and includes the three corresponding residuals of the reduced-form RFM variables into analysis. The second group includes three lifetime variables. The third group includes three transactional variables, whether to order by cash, by house credit card, or through telemarketing promotion. The fourth group includes a variable that shows marketing intensity. The fifth group includes a probability variable which is calculated by probability model. Because the training examples are drawn from a different probability distribution than the examples to which the model is applied, sample selection bias occurs. Thus, following James J. Heckman (1979)'s procedure, this thesis uses latent class logistic regression to obtain probability estimates and added the calculated probability into profit model.

5.3 Descriptive Analysis

Before statistical analysis of our data, it is important to look at the distribution of profit. As the response rate of the recent promotional campaign is 5.47%, it indicates 95% of the customers did not respond to the mailing promotion and their profits are negative because of the loss of mailing costs. In our analysis, we will only model the customers who make a purchase so that here we only examine the profit distribution of the customers who actually responded. The total number is 5698 cases.

As Mulhern (1999) suggests, the simplest way for assessing the distribution of customer profit can be made by observing a ranking of computed profits from highest to lowest. Consider the two-dimensional plot (Figure 5) in which the vertical axis represents customers profit and the horizontal axis represents customers—in decreasing order from highest to lowest profit. A very small group of customers have high level of profit, while a large group of customers have low level of profit.

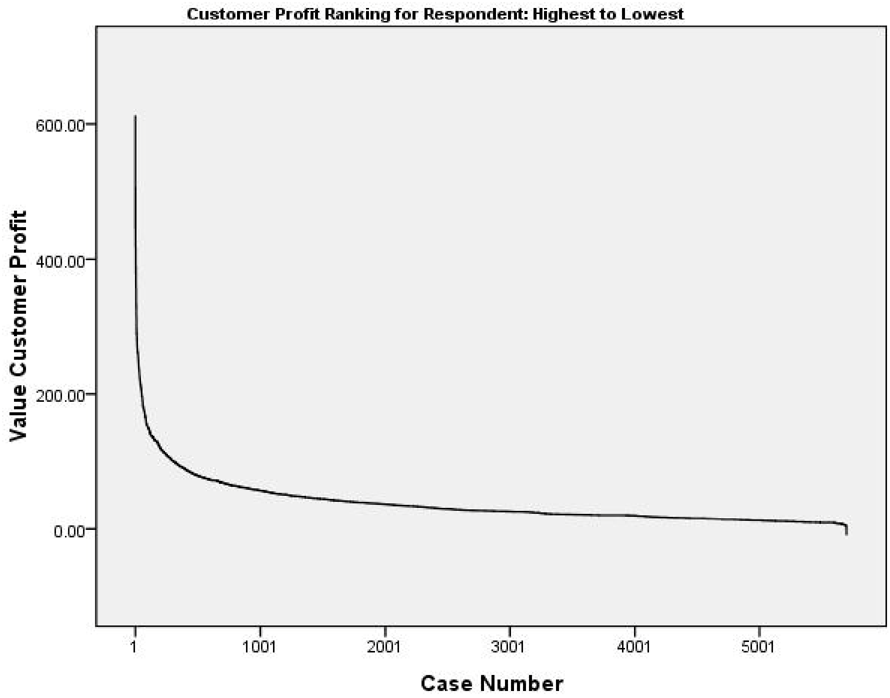


Figure 5 Customer Profit Ranking for Respondents: Highest to Lowest

For our case, looking at the cumulative distribution of profit will provide us more information (Figure 6). Profits are quite concentrated, as 20% of the customers account for 50% of the profits. As we only consider the respondents with positive profit, the profit data in our cases is not as concentrated as in Mulhern (1999).

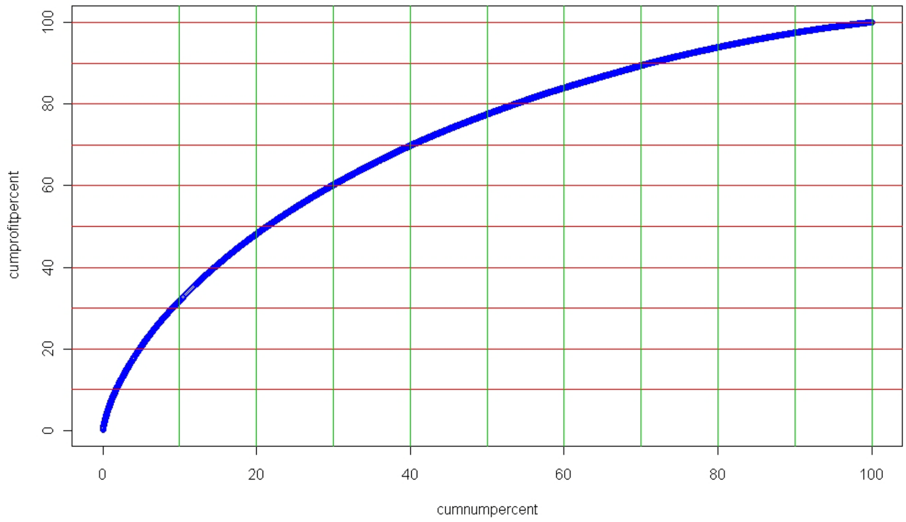


Figure 6 Inverted Lorenz Curve for Customer Profit

In addition, we also look at the histogram of customer profit with a normal curve added (Figure 7). The profit distribution is right-skewed and not normally distributed, proving that profit data do not conform to the normality assumption. A frequency analysis shows that the mean (38.3) is larger than median (27.0), and the degree of skewness equals 4.008. It also indicates that the profit distribution is right-skewed.

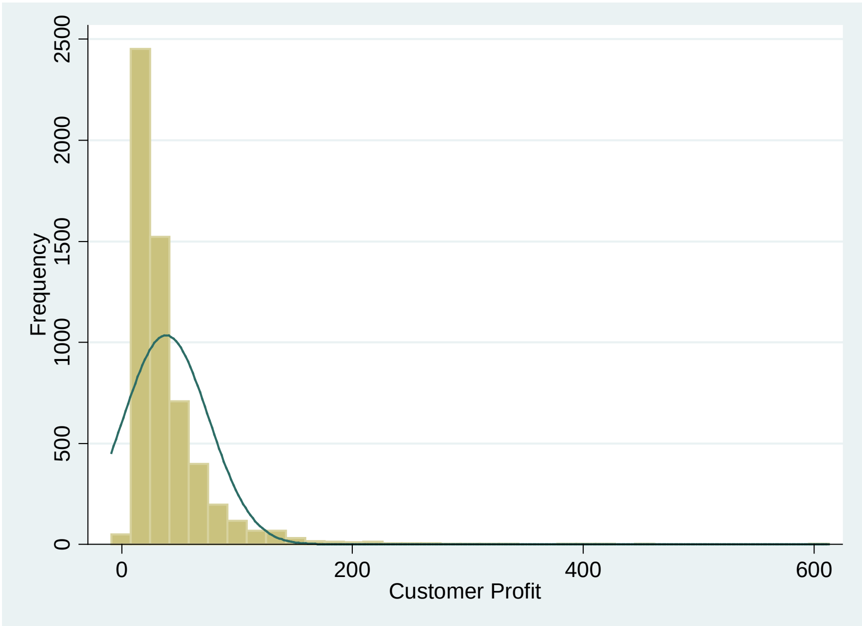


Figure 7 Frequency Distribution of Customer Profit for Respondents

Furthermore, we also examine the plot of probability density (Figure 8). The thin plot is the reference distribution, which is normal (Gaussian). The thick plot is the density probability plot, which indicating the difference from a normal distribution. It further confirms that profit data is not normally distributed. Thus, we believe while the profit data violate the normality model assumption of OLS regression, it is not appropriate to use it to customer profitability analysis. Quantile regression is a suitable method as it has no normality distributional assumption.

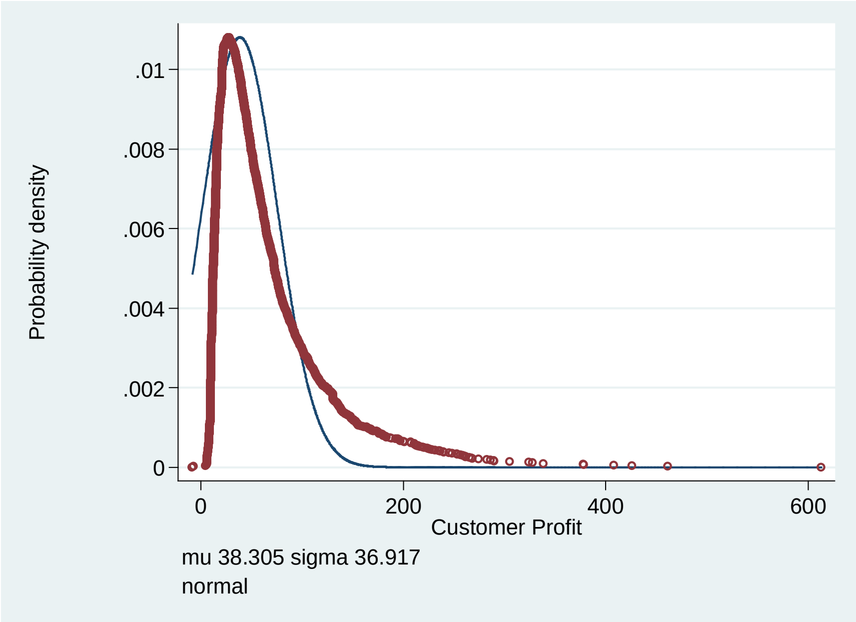


Figure 8 Probability Density Plot of Customer Profit for Respondents

Next, we examine the OLS regression's assumption of homoscedasticity. Graphically, we look at the boxplot of pairs of variables. Here, we take the boxplot of dependent variable (targpbfo) and one of the independent variables, lifetime dollar class average (salcat) (Figure 9). Each blue box shows the middle 50% of the cases for the group, indicating how spread out the group of scores is. As we can see, the heights of the boxes are different, suggesting that the variance across groups is not homogeneous. Especially, the group of customers whose lifetime dollar class average equal six is more spread out than the other groups, suggesting unequal variance. The boxplot also shows that for each group of independent variable, lifetime dollar class average, there are many outliers. The large amount of outliers indicates that OLS regression estimate will be biased and inaccurate as it is highly influenced by outliers. Quantile regression is suitable for customer profitability analysis as it is naturally robust to outliers.

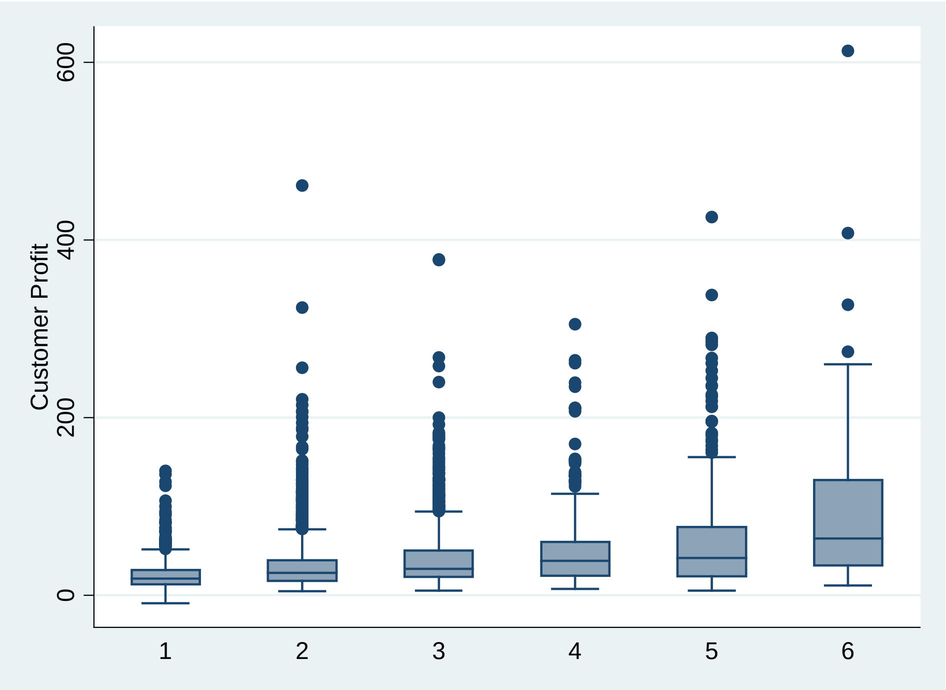


Figure 9 Boxplot of Customer Profit over Salcat

In addition, we also test the homogeneity of variance by one-way ANOVA. Here, we also make this test of targpbfo and salcat. The null hypothesis for the test of homogeneity of variance states that the variance of the dependent variable is equal across groups defined by the independent variable, i.e., the variance is homogeneous. Since the probability associated with the Levene Statistic (<0.001) is less than or equal to the level of significance, we reject the null hypothesis and conclude that the variance is not homogeneous.

Table 2 Test of Homogeneity of Variances

Customer Profit

Levene Statistic	df1	df2	Sig.
145.211	5	5692	0

The boxplot and statistical test for homogeneity both confirm that the variance of profit data is heterogeneous. Thus, OLS regression's assumption of homoscedasticity is violated, suggesting that OLS regression is not able to examine the different effects that predictors exert on the different quantile of the distribution of dependent variable. It also indicates the potential analytical advantages of quantile regression on examining the relationship between independent variables and dependent variable.

5.4 Endogeneity Bias Correction

Before we conduct any statistical data analysis, it is very important to address one critical issue, which is the possible endogeneity of the RFM variables. In a structural equation model ($y=ßx+\epsilon$), an explanatory variable must be uncorrelated with the error of the model in order to infer causality of x on y. If x is correlated the model error (ϵ), this variable is said to be endogenous and its parameter estimates may be biased. Such problems are due to the omitted variables embedded in the error, which simultaneously affect both the *y* and *x* variables. Direct marketers often use RFM variables to predict future purchases, yet the RFM variables are based on previous responses from these households, e.g., the number or frequency of past purchases. In this sense, the RFM variables may be endogenous and their parameter estimates may be biased due to the correlations between RFM variables and error of the model. This is common problem with re-occurring or simultaneous data, which arises from the lack of empirical data to control for endogeneity. In linear parametric models, researchers often adopt the instrumental variable method (Gönül et al. 2000) or the control function approach (Blundell & Powell 2004) to correct the potential bias. However, the instrumental variable approach is not applicable to the present study where instrumental variables are not available. Thus, we adopt the control function approach to test for endogeneity bias, which is accomplished by adding the residuals of the endogenous variables into the model as control variables.

Following the procedures in Blundell and Powell (2004), we first run a parametric reduced-form regression to compute the estimates of endogenous RFM variables on the whole dataset. In the second stage, the residuals of the reduced-form regressors are included as covariates in the binary response model to account for their endogeneity. In Table 3, the first three columns are the reduced-form estimates for the RFM variables with two lifetime variables as covariates. The two lifetime variables are squared because they may have nonlinear effects. Given their adjusted R-square, the explanatory power of the three reduced-form equations is fairly high. The fourth column refers to the model without any adjustment for endogeneity. Except for monetary value, recency and frequency have very small coefficient estimates, although they are statistically significant. Despite the large dataset (N = 106,280), the overall fit of the uncorrected model is statistically insignificant ($p = 0.620$), indicates a potential endogeneity bias in the RFM variables. The last column includes the residuals of the RFM variables as the control variables. The coefficient estimates change drastically, and thus the effect of correcting for endogeneity is obvious given the improvement in model fitness ($p = 0.021$). For the endogeneity tests, we employ the asymptotic t-test developed by Smith and Blundell (1986). The significant results of the t-tests reject the null hypotheses of exogeneity for the RFM variables.

Table 3 Results of the Endogeneity Test

Models and Tests /Variables	Reduced Form: Recency	Reduced Form: Frequency	Reduced Form: Monetary Value	Logit Model: Uncorrected	Logit Model: Corrected
Constant	32.650** (0.206)	0.969** (0.022)	3.427** (0.006)	–4.068** (0.097)	–11.225** (1.177)
Recency	—	—	—	–0.026** (0.002)	–0.219** (0.019)
Frequency	–6.428** (0.028)	—	0.379** (0.002)	0.025** (0.010)	–1.297** (0.056)
Monetary value	—	—	—	0.368** (0.022)	3.511** (0.217)
Lifetime contacts	–2.353** (0.118)	–0.193** (0.013)	0.0545** (0.001)	—	—
Lifetime contacts ^{2}	0.302** (0.015)	0.048** (0.002)	—	—	—
Lifetime orders	0.668** (0.011)	0.121** (0.001)	–0.0085* (0.000)	—	—
Lifetime orders ^{2}	–0.006** (0.000)	–0.001** (0.000)	—	—	—
Adjusted R^2	0.345	0.255	0.464		
Endogeneity Test: Recency					9.428** *(t)*
Endogeneity Test: Frequency					–14.486** *(t)*
Endogeneity Test: Monetary value					–15.110** *(t)*
				$p = 0.620$	$p = 0.021$

Notes: 1) The standard errors are shown in parentheses. 2) The endogeneity test is the asymptotic t-test. 3) ** = significant at the 0.001 level.

5.5 Sample Selection Bias Correction

When estimating response profit, a fundamental problem is that any estimator, for example a regression equation, must be learned based on examples of people who actually respond. But this estimator must then be applied to a different population, i.e. both respondents and non-respondents. This problem is known in general as sample selection bias. It occurs whenever the training examples are drawn from a different probability distribution than the examples to which the model is applied.

The standard method of compensating for sample selection bias in econometrics is a two-step procedure due to James J. Heckman (1979). Heckman's procedure is applicable when each example *x* belongs to one of the two cases, i.e. $j(x) = 0$ or $j(x) = 1$, and the dependent variable to be estimated $y(x)$ is observed for a training example if and only if $j(x) = 1$. The first step of the procedure is to learn a probit linear model to estimate conditional probabilities $P(j = 1 \mid x)$. A probit model is a variant of logistic regression where the cumulative Gaussian probability density function is the sigmoid function. The second step of Heckman's procedure is to estimate $y(x)$ by linear regression using only the training example *x* for which $j(x) = 1$, but including for each x a transformation of the estimated value of $P(j = 1 \mid x)$. Heckman has proved that this procedure yields estimates of $y(x)$ that are unbiased for all *x*, regardless of whether $j(x) = 0$ or $j(x) = 1$, under certain conditions (Zadrozny & Elkan 2001; Heckman 1979).

We use a latent class logit model to obtain probability estimates. In statistics, a latent class model (LCM) relates a set of observed discrete multivariate variables to a set of latent variables. It is called a latent class model because the latent variable is discrete. A class is characterized by a pattern of conditional probabilities that indicate the chance that variables take on certain values. As the standard method, probit model, is a variant of logistic regression, latent class logistic regression model is also suitable for estimating conditional probabilities in the first step.

Jedidi et al (1997) argue that using observed variables may fail to thoroughly reflect the customer heterogeneity and it is necessary to identify some underlying structures to represent the communalities of the original variables. The latent class analysis (LCA) refers to the statistical process of finding out the latent classes or latent structures from multivariate categorical data (Heinen 1996). Researchers have shown the advantages of latent class analysis over aggregate estimation in direct marketing issues for its ability of accounting for heterogeneity, as well as group size and probability (Jain, Bass & Chen, 1990). Previous studies also indicate the suitability of latent class analysis in direct marketing applications (Wedel et al. 1993).

5.6 Predictive Modeling Using Conditional Quantile Regression

As we proposed in Chapter 4, the first application of quantile regression is to predict the response profit of each customer. As to section 5.2, the dependent variable is the profit/loss before the fixed operating cost (targpbfo) for this promotion. The independent variables include four groups of variables. One group are the RFM variables after endogeneity correction, including the predicted recency of the last purchase (pred_r1), the predicted frequency of purchases in the last 36 months (pred_f1), the predicted money value of a customer's purchase in the last 36 months (pred_m1), and the corresponding residuals of these three predicted variables (resid_r1, resid_f1 and resid_m1). Second group are lifetime value variables, including lifetime orders of this kind of promotion (prord85), dollar class of average order last year (salflg) and lifetime dollar class average (salcat). Third group is a marketing intensity variable, which is the number of catalogs received previously (crcpr85). Fourth group are a number of customer transaction, demographic, and credit history variables, including purchase with the credit card from the company (hcrd), telephone orders (tele), and cash orders (cash).

After specifying variables, we begin the statistical analysis. In the next subsection, we discuss the first step of our analysis—model construction and interpretation.

5.6.1 Model Construction and Interpretation

The whole dataset contains records of 106, 284 consumer responses to a recent promotion. After correction of endogeneity of RFM variables, there are some missing values of variable--pred_m1. We eliminate the cases with missing values and the whole dataset now contains 86,191 cases. We propose a two-step sampling scheme in which the analysis is separated into a training phase and a testing phase. The training dataset contains 90% of the records of the whole dataset, while the testing dataset contains the remaining 10%. All the cases in both training and testing samples are drawn randomly but we maintain the response rate in both datasets equals to that of the whole dataset, roughly 6.18%. The model building is performed on the training dataset while the prediction is on the testing dataset.

After estimating the conditional probability by latent class logistic regression, each customer is assigned a probability score which indicates his/her likelihood of responding to this promotion. By incorporating the probability value into independent variables, we estimate the simple linear quantile regression model:

$$\begin{aligned}
Q_y(\tau \mid x_j) &= \beta_0(\tau) + \beta_1(\tau) * pred_r1 + \beta_2(\tau) * pred_f1 + \beta_3(\tau) * pred_m1 \\
&+ \beta_4(\tau) * resid_r1 + \beta_5(\tau) * resid_f1 + \beta_6(\tau) * resid_m1 + \beta_7(\tau) * prord85 \\
&+ \beta_8(\tau) * salf\lg + \beta_9(\tau) * salcat + \beta_{10}(\tau) * crcpr85 + \beta_{11}(\tau) * hcrd + \beta_{12}(\tau) * tele \\
&+ \beta_{13}(\tau) * pred_dep
\end{aligned}$$

The model is estimated at t = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9). The results of the parameter estimates are depicted in Table 4 and Figure 10 with 95% confidence band. OLS results are also plotted as bold lines in Figure 10. An overall impression suggests that quantile regression reveals significant and interesting results that would have been hidden in the traditional model. Figure 10 shows quantile regression results at τ = (0.1, -- 0.9). The gray area is the 95% confidence band. The horizontal straight line is for the visual convenience of coefficient of significance.

Figure 10 Quantile Regression Results

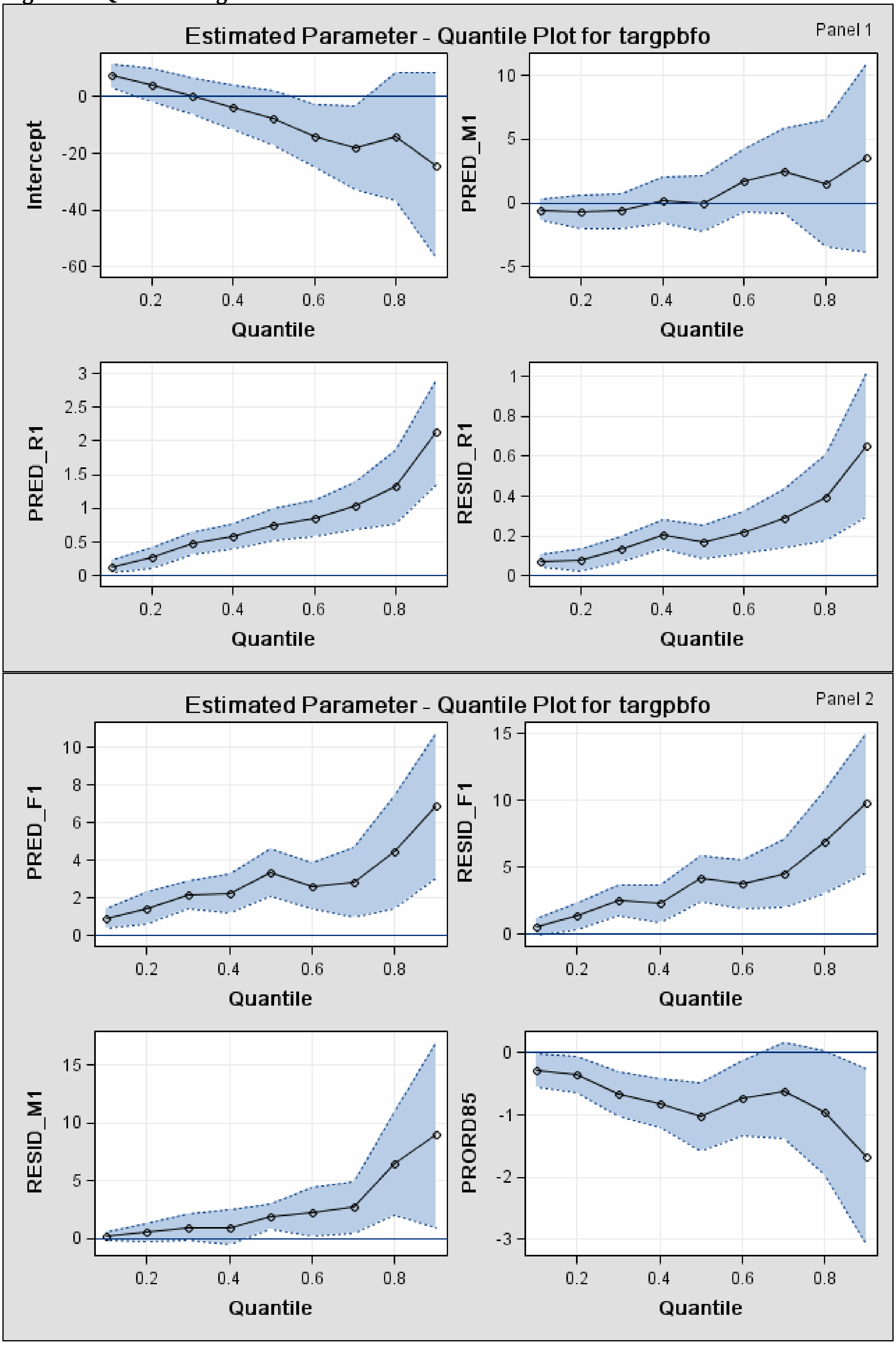
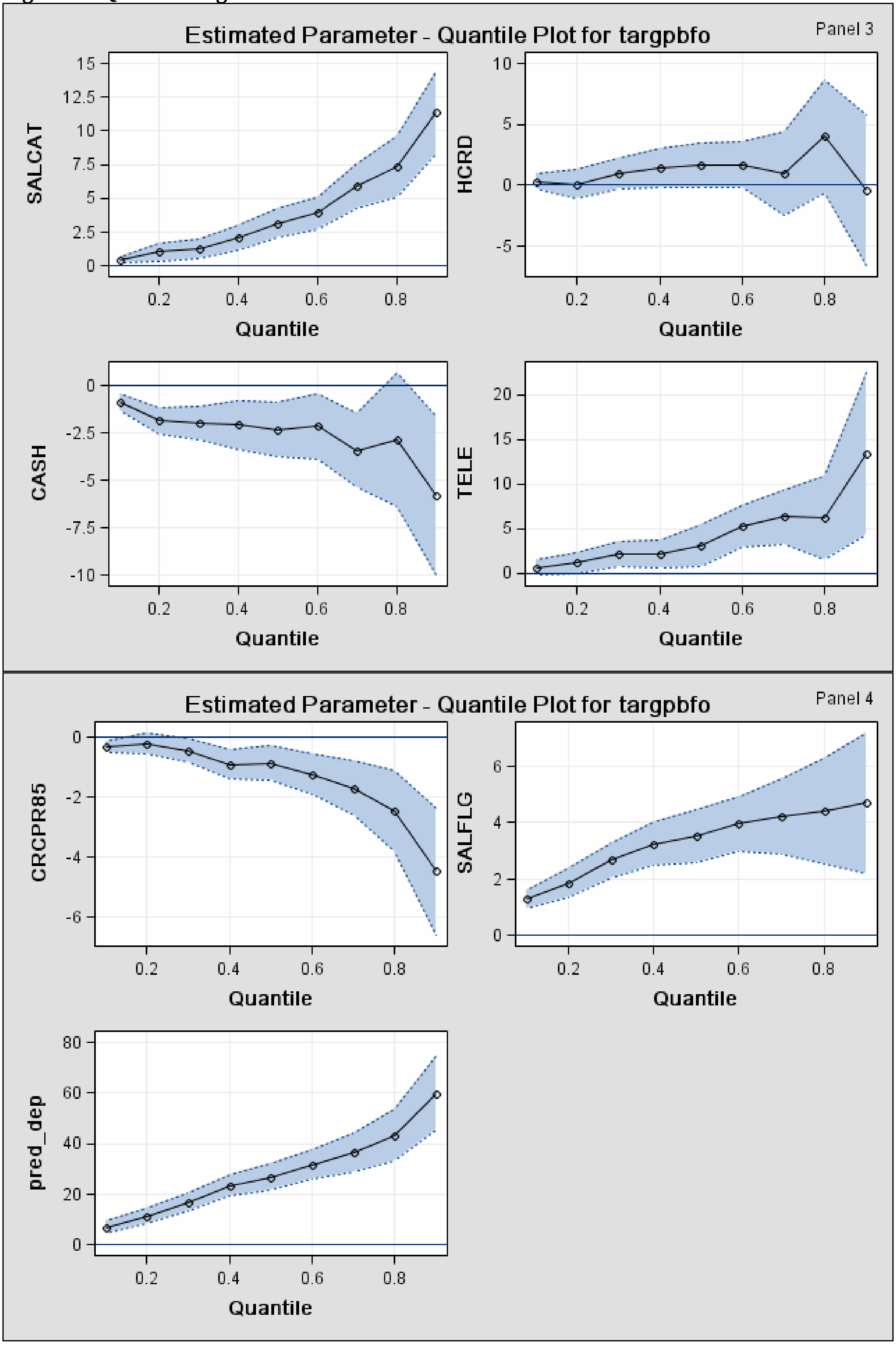


Figure 10 Quantile Regression Results—Continued



OLS regression imposes the constraint that the effect of a particular predictor on response profit is the same for the different groups of customers. When the customers are homoskedastic in terms of response profit, the slope estimates of the conditional quantile functions at each point of the distribution of the dependent variable will be equal to each other and to the slope estimates from the OLS. However, when the customers are heteroskedastic, the slope estimates of the conditional quantile functions will differ from each other as well as from the OLS slope estimates. Figure 10 gives us a direct image that the effects of predictor variables on the dependent variable which are different cross different quantiles, indicating heterogeneity of customers.

The goodness of fit of OLS regression is measured by the R-square, while the goodness of fit of quantile regression is measured by the Pseudo R-square (Table 4). In OLS-regression models, R-square is interpreted as the proportion of variation in the dependent variable explained by the predictor variables in the model. An analog of the R-square is developed for quantile-regression models. Koenker and Machado (1999) suggest measuring goodness of fit by comparing the sum of weighted distances for the model of interest with the sum in which only the intercept parameter appears. Stata provides the measure of goodness of fit and refers to it as "pseudo R-square" to distinguish it from the ordinary R-square. Interested readers can consult Koenker and Machado (1999) on the goodness of fit of quantile regression.

Estimated coefficient by OLS regression for the predicted money value of a customer's purchase in the last 36 months (Pred_M1) shows a positive impact of this variable on profit but the results by quantile regression are quite different (Table 4). Quantile regression estimated coefficients for Pred_M1 show that it has a negative effect on low and middle quantile profit (10^{th} to 50^{th}) while a positive effect on high quantile profit (60^{th} to 90^{th}). Both quantile regression and OLS regression estimated coefficients for this variable is not significant. It means that the money value of a customer's past purchase is not a significant predictor for future response profit.

The predicted recency of the last purchase has a significant positive effect on response profit estimated from both OLS and quantile regression. But the magnitude of the effects is increasing as quantile increases. For example, one unit increase in the number of months that has elapsed since the last purchase would increase the response profit by 0.134 for 10^{th} quantile customers while it will increase the response profit by 2.124 for 90^{th} quantile customers. It means that for customers who will spend low amount of money on promotion, a longer period of not buying from the company accelerates in a lower degree; while for customers who will spend high amount of money on promotion, a longer period of not buying from the company bring more profit. OLS regression coefficients suggest that one unit increase in the number of months that have elapsed since the last purchase would increase the response profit by 1.221 for every customer. We see that quantile regression provides us a more accurate estimation of the effects predictors exert on dependent variable. Managers should treat low quantile and high quantile customers differently according to different quantile regression coefficients. In this case, managers should be more concerned with high quantile customers among customers who have the same degree of recency.

The OLS regression coefficient for the predicted frequency of purchase in the last 36 months (Pred_F1) shows a marginal effect of 4.626 units on response profit for every customer (Table 4). However, quantile regression estimation shows that one unit increase in the frequency of purchase in the last 36 months would only increase response profit by a very low degree for low quantile customers, for example 0.904 for 10th quantile, 2.177 for 30th quantile and 3.354 for median customers. We can see that OLS regression gives us an incomplete picture.

Both OLS and quantile regression estimates for lifetime orders of this kind of promotion (prord85) show a negative effect on response profit (Table 4). It means that the more a customer orders from this kind of promotion, the less amount of money the customer will spend on the same promotion. But this predictor is not quite significant for high quantile customers, indicating that high quantile customers are not affected by previous orders placed on the same promotion.

Lifetime dollar class average (salcat) influences the response profit significantly and positively by both OLS and quantile regression estimation (Table 4). It may suggest that previous purchase behavior has a positive correlation with this year’s purchase, indicating the consistency of customers’ purchase behavior. The impact this variable exerts on response profit increases monotonously across quantiles. OLS regression’s estimated effect (5.664) is quite similar to 70th quantile regression estimation, indicating that the mean is close to 70th quantile conditional on this variable. We note that for 90th quantile customers, the effect (11.34) is more than twice than the mean effect, indicating the heterogeneous purchase behavior.

Dollar class of average order of last year (salflg) also has a significantly positive effect on the response profit by both OLS and quantile regression estimates (Table 4). It means last year's purchase behavior positively influences this year's purchase amount. Comparing OLS and quantile regression coefficients of dollar class of average order of last year (salflag) and lifetime dollar class average (salcat), we find that previous (for a 12-year period but not continuously) average purchase amount has greater influence on this year's purchase amount than last year's average purchase amount does. In addition, the changes of the magnitude of the effects of lifetime dollar class average are more dramatic across different quantiles than dollar class of average order last year are. It means that customers conditioning on lifetime dollar class average (salcat) are more heterogeneous than conditioning on dollar class of average order last year (salflag).

Both OLS and quantile regression indicate that the number of catalogs received previously (crcpr85) has a negative effect on response profit (Table 4). For the 20th and 30th quantile customers, the effects are not quite significant ($P<0.01$). For the 10th quantile customers, the effects are significantly negative and small (-0.329), while for the 80th and 90th quantile customers, the effects are significantly negative but large (-2.473 and -4.481). It gives suggestions to direct marketers that increased marketing intensity does decrease customers' purchase profit but low quantile customers are not quite influenced by it while high quantile customers would reduce their future purchase amount due to the overwhelming of marketing promotions. We see that the confidence band of variable—hcrd—covers the zero line, indicating that whether making purchase with the credit card from the company (hcrd) does not have any significant influence on customers' purchase amount. The effects are not significant across quantiles.

OLS regression estimated coefficient shows that whether to place cash orders does not have a significant influence on customers' purchase profit ($P<0.05$). However, quantile regression estimation shows that the effects are significant and negative across different quantiles. This contrary result further reveals that quantile regression may allow detecting significant effects from variables whose coefficients may have been dismissed for not appearing to be significantly different from zero in a mean based model such as OLS. The quantile regression estimation results suggest that high quantile customers are more unwilling to make purchase by cash. Marketing managers may encourage high quantile customers to make purchase by other forms of payment in order to increase company profit.

Whether to make telephone orders does not have a significant influence on customers' profit for 10th and 20th quantile customers but have a significant positive influence for other quantiles (Table 4). For high quantile customers, especially 90th quantile (13.419), the magnitude of effects are larger than lower quantile customers, such as 30th quantile (2.17) and 50th quantile (3.055). It suggests that encouraging telephone orders may increase customer profit more greatly for high quantile than low quantile customers.

The estimated response probability by latent class analysis has a significant positive effect on the profit amount (Table 4). It means that the more likely someone will respond to the promotion, the more money one will spend on this promotion. For high quantile customers, once their response probability increases, they will spend more money on the promotion (i.e. 90th quantile, 59.798). But for low quantile customers, the increase of profit amount is not that much (i.e. 10th quantile, 6.924). Thus, it is important for managers to increase the response probability of high quantile customers in order to increase company profit.

From the above analysis of coefficients, it is clear that with the presence of heterogeneity there are advantages of quantile regression over OLS regression. First, instead of the point estimate for the conditional mean, we have reliable estimates for the whole distribution. Second, the conditional mean results could be derived from the conditional quantile effect. In particular, if the distribution of effects is not too skewed, the conditional mean effect would be close to the median (Ma & Pohlman 2008). In our analysis, for most variables, the conditional mean effects are higher than the conditional median effects, roughly between 70th and 80th quantiles, probably because of the skewness of our profit data. However, since we are estimating many parameters, we must be cautious about our inferences. With more information, one has greater challenges in the decision-making (Ma & Pohlman 2008).

For managers, the complete information of covariates effects across different quantiles is useful for managing customer profiles and maximizing company profit. For example, the estimated covariates effects of the predicted recency of the last purchase are different across different quantiles customers. The effect is 0.134 for 10th quantile, 0.481 for 30th quantile, 0.757 for 50th quantile, 1.029 for 70th quantile, and 2.124 for 90th quantile. Managers can group the customer database into different segment, i.e. 1st tier, 2nd tier, 3rd tier, 4th tier and 5th tier customers. Basing on the budget constraint and the objective of promotion, company can allocate promotion resources to different customer segments. However, OLS regression only gives one coefficient for the predictor variable, so that each customer is assigned to one segment. Managers have insufficient knowledge on how to allocate resources in order to minimize cost and maximize profit. Thus, quantile regression can help managers manage customers more effectively.

Table 4 OLS and Quantile Regression Results

		Quantiles								
Targpbfo	OLS	10th	20th	30th	40th	50th	60th	70th	80th	90th
PRED_M1	2.069	-0.549	-0.712	-0.642	0.224	-0.061	1.754	2.498	1.519	3.531
	(1.850)	(0.436)	(0.665)	(0.719)	(0.919)	(1.096)	(1.228)	(1.686)	(2.520)	(3.753)
RESID_M1	8.743***	0.256	0.526	0.983	0.98	1.886***	2.285*	2.709*	6.454	8.971*
	(1.569)	(0.189)	(0.391)	(0.575)	(0.780)	(0.562)	(1.066)	(1.142)	(2.257)	(4.071)
PRED_R1	1.221***	0.134**	0.266***	0.481***	0.581***	0.757***	0.856***	1.029***	1.317***	2.124***
	(0.206)	(0.049)	(0.075)	(0.081)	(0.097)	(0.118)	(0.135)	(0.180)	(0.278)	(0.396)
RESID_R1	0.283***	0.075***	0.081**	0.134***	0.206***	0.169***	0.221***	0.285***	0.394***	0.653***
	(0.079)	(0.017)	(0.029)	(0.032)	(0.037)	(0.044)	(0.054)	(0.075)	(0.111)	(0.184)
PRED_F1	4.626***	0.904***	1.463***	2.177***	2.237***	3.354***	2.631***	2.818**	4.438**	6.912***
	(0.997)	(0.274)	(0.438)	(0.373)	(0.520)	(0.649)	(0.620)	(0.954)	(1.527)	(1.959)
RESID_F1	6.192***	0.534	1.377**	2.529***	2.293**	4.187***	3.759***	4.574***	6.941***	9.827***
	(1.407)	(0.314)	(0.500)	(0.582)	(0.712)	(0.865)	(0.946)	(1.316)	(1.966)	(2.664)
PRORD85	-0.946*	-0.284*	-0.343*	-0.668***	-0.81***	-1.033***	-0.728*	-0.608	-0.961	-1.676
	(0.391)	(0.140)	(0.147)	(0.180)	(0.206)	(0.285)	(0.308)	(0.396)	(0.512)	(0.724)
SALCAT	5.664***	0.429***	1.01**	1.277***	2.049***	3.155***	3.932***	5.902***	7.397***	11.34***
	(0.732)	(0.106)	(0.321)	(0.379)	(0.475)	(0.548)	(0.597)	(0.859)	(1.155)	(1.573)
HCRD	2.577	0.275	0.081	0.912	1.401	1.64	1.687	0.964	3.998	-0.423
	(1.630)	(0.317)	(0.635)	(0.652)	(0.830)	(0.941)	(0.983)	(1.731)	(2.346)	(3.171)
CASH	-2.243	-0.869***	-1.841***	-1.969***	-2.059**	-2.306**	-2.156*	-3.407***	-2.848	-5.801**
	(1.384)	(0.238)	(0.360)	(0.466)	(0.659)	(0.723)	(0.865)	(0.997)	(1.797)	(2.142)

Absolute value of t-statistics (OLS) and Bootstrap t-statistics (quantile regression) in parentheses.
Note: ***: p<0.001, **: p<0.01, *: p<0.05

Table 4 OLS and Quantile Regression Results—Continued

	Quantiles									
Targpbfo	OLS	10th	20th	30th	40th	50th	60th	70th	80th	90th
TELE	5.842***	0.627	1.166	2.17**	2.151*	3.055*	5.332***	6.34***	6.232**	13.419**
	(1.635)	(0.441)	(0.616)	(0.719)	(0.836)	(1.194)	(1.192)	(1.541)	(2.373)	(4.623)
CRCPR85	-1.439**	-0.329***	-0.214	-0.454*	-0.92***	-0.872**	-1.244***	-1.705***	-2.473***	-4.481***
	(0.484)	(0.100)	(0.180)	(0.205)	(0.251)	(0.292)	(0.338)	(0.470)	(0.687)	(1.085)
SALFLG	2.788***	1.276***	1.866***	2.66***	3.246***	3.533***	3.953***	4.212***	4.438***	4.694***
	(0.629	(0.161)	(0.271)	(0.32)	(0.392)	(0.471)	(0.491)	(0.688)	(0.959)	(1.272)
pred_dep	36.288***	6.924***	11.283***	16.913***	23.46***	26.805***	31.802***	36.385***	43.367***	59.798***
	(3.110	(1.343)	(1.600)	(1.859)	(2.197)	(2.635)	(2.818)	(3.871)	(5.079)	(7.514)
Constant	-22.94	7.384	3.975	0.037	-3.624	-7.591	-13.873	-17.957	-14.081	-24.22
	(8.714	(2.097)	(2.981)	(3.295)	(4.034)	(4.931)	(5.682)	(7.508)	(11.465)	(16.714)
Observations	4794	4794	4794	4794	4794	4794	4794	4794	4794	4794
R2/Pseudo R2	0.175	0.028	0.043	0.055	0.070	0.086	0.107	0.119	0.138	0.158

Absolute value of t-statistics (OLS) and Bootstrap t-statistics (quantile regression) in parentheses.
Note: ***: p<0.001, **: p<0.01, *: p<0.05

5.6.2 Quantile Regression Predictive Results

The usefulness of quantile regression is not limited to the informative estimation it provides on the effects of the independent variables. It can also be useful in making prediction, in our study, of the spending distribution of each customer. According to Cade and Noon (2003), the interval between the 90^{th} and 10^{th} regression quantile estimates at any specified value of X is an 80% prediction interval for a single future observation of y. Prediction intervals (for some number of future observations) based on assuming a normal error distribution, as is done in ordinary least squares regression, are sensitive to departures from this assumption (Neter et al. 1996). Quantile regression avoids this distributional assumption and gives more accurate interval predictions.

In our study, using the estimated model from the training sample, we can extrapolate it onto the testing sample. In other words, we calculate the predicted profit for each customer of different quantiles in the testing dataset from training dataset coefficients and testing dataset independent variables. For example, for a customer who has not purchased for 7 months since last purchase (Pred_R1), who has made purchase 3 times in the last 36 months (Pred_F1), whose monetary value of purchase in the last 36 months is 5 (Pred_M1), who has placed 11 lifetime orders of this kind of promotion (prord85), whose lifetime dollar class average (salcat) is 2, whose dollar class of average order last year (salflg) is 3, who received 6 catalogs previously (crcpr85), who doesn't purchase with the credit card from the company (hcrd), who does not place cash orders (cash), and who place telephone orders (tele), we predict his/her spending distribution as below. The prediction equations are as follows:

$\tau = 10^{th}$: Y = 7.384 + 0.134*Pred_R1 + 0.075*Resid_R1 + 0.904*Pred_F1 + 0.534*Resid_F1 – 0.549*Pred_M1 + 0.256*Resid_M1 – 0.284*prord85 + 0.429*salcat + 1.276*salflg – 0.329*crcpr85 + 0.275*hcrd – 0.869*cash + 0.627*tele + 6.924*pred_dep = 14.312

For other quantiles, the prediction is identical. And we have the following predictions (Table

5):

Table 5 Quantile Regression Predictive Result

Quantile	Predicted Profit	Probability of Profit
10th	14.312	P (Y<= 14.312) = .1
20th	20.320	P (Y<= 20.320) = .2
30th	24.731	P (Y<= 24.731) = .3
40th	29.954	P (Y<= 29.954) = .4
50th	35.333	P (Y<= 35.333) = .5
60th	46.244	P (Y<= 46.244) = .6
70th	56.261	P (Y<= 56.261) = .7
80th	64.898	P (Y<= 64.898) = .8
90th	89.321	P (Y<= 89.321) = .9

These predicted values can be interpreted as follows: with probability τ = .9, the profit of this customer is less or equal than 89.321 (Table 5). In other words, there are 10% chance that the profit of this customer will exceed 89.321. In summary, we can have this customer's predicted spending distributions.

In the same way, we can predict each customer's spending distribution. With such information, we understand that how much money a specific customer will spend had he/she spent his/her 10^{th} or 90^{th} quantile spending distribution. Using quantile regression prediction, we can easily make interval prediction of a customer's profit. We can choose between $\tau = 90^{th}$: Y and $\tau = 10^{th}$: Y. We can provide more informative prediction of customer profit than OLS, which only provides the predicted mean. For the example we discuss above, we can make the statement such as that: there is an 80% confidence level that this customer profit will be within the interval (14.312, 89.321).

Such information of customers' spending distribution is very useful for direct marketers (Perlich et al. 2007). For instance, one can make use of such information as they want to predict 80^{th} quantile of IBM's clients' spending distribution. They assume that 80^{th} quantile of IBM's clients spending distribution is the realistic wallet that IBM can hope for or target. In our study, we can also make use of such information in the same way. If the company can decide the amount of profit it hopes for, it can simply target those customers whose profit will exceed that amount. But one should be cautious that the probability of exceeding a high profit amount is also very low.

Once each customer is assigned a predicted profit of different quantiles from 10^{th} to 90^{th} quantiles, we rank the testing data sets by the predicted profit in descending order. We calculate and average the cumulative profit lift of 10 folds. Table 6 gives the results of 10-fold cross-validation.

The predictive performance of 30^{th}, 40^{th}, 50^{th} and 60^{th} quantile regression are better than that of OLS regression (Table 6). The other quantile regressions predictive performance are slightly worse than OLS regression. The interpretation of quantile regression prediction is different from that of OLS regression and should be done with extreme caution. For instance, if we predict the 30^{th} quantile profit, it means that if all the customers have 0.3 chance to be below the predicted profit value and 0.7 (1-0.3) chance to exceed the predicted profit, ranking the testing data sets by such predicted profit value can tell us how much profit generated by the selected customers. In our study, 40^{th} quantile regression model can identify true profit 8.02 times as well as the random model, while OLS regression model can identify true profit 7.97 times as well as the random model.

However, the limitation of the predictive power of conditional quantile regression should also be addressed carefully. For the dataset we use in this study, the predictive power of conditional quantile regression for “medium quantiles” is better than that of OLS regression, but for other datasets the situation may not be the same. In another word, managers can not simply make target selection based on the prediction of conditional quantile regression for “medium quantiles”. The reason is that managers have no idea each customer will spend at which quantile of his/her spending distribution. Although the information of the whole probabilistic spending distribution of each customer is available, direct use of such information for prediction is not useful for helping select the most profitable customers. Other applications of such information could be useful for target selection and this thesis provides a suggestion in Future Research of Chapter 6.

Tables 6 Cross-validation of Quantile Regression Predictions: Cumulative Profit Lift

Decile/ Model	OLS Regression	Quantile Regression (0.1)	Quantile Regression (0.2)	Quantile Regression (0.3)	Quantile Regression (0.4)	Quantile Regression (0.5)
1	796.6	794.8	796.5	800.0	801.9	799.8
2	453.7	440.9	443.1	447.6	450.9	451.2
3	325.0	321.8	321.4	324.0	324.3	323.9
4	252.9	250.1	250.3	251.2	251.2	252.2
5	204.1	202.5	202.8	203.6	203.8	203.9
6	171.0	169.2	169.3	170.0	170.2	170.4
7	145.0	144.9	145.0	145.2	145.3	145.4
8	127.2	127.0	127.2	127.3	127.2	127.2
9	112.3	112.2	112.3	112.2	112.4	112.3
10	100.0	100.0	100.0	100.0	100.0	100.0

Note: The reported figures are the means of the lifts of the 10 experiments

Tables 6 Continued

Decile/ Model	Quantile Regression (0.6)	Quantile Regression (0.7)	Quantile Regression (0.8)	Quantile Regression (0.9)
1	800.4	795.7	796.5	795.8
2	452.3	451.0	451.2	451.3
3	323.7	324.0	323.9	324.4
4	252.7	252.8	253.6	253.6
5	204.0	204.4	203.5	203.5
6	170.4	170.8	170.9	171.3
7	145.4	145.9	145.6	146.0
8	127.1	127.0	127.2	127.1
9	112.3	112.4	112.3	112.2
10	100.0	100.0	100.0	100.0

Note: The reported figures are the means of the lifts of the 10 experiments

5.7 Profit Analysis Using Unconditional Quantile Regression

As we discussed before in section 3.6, quantile regression proposed by Koenker and Bassett (1978) is actually a conditional quantile regression. Unlike conditional means, however, conditional quantile do not average up to their unconditional population counterparts. As a result, the estimates obtained by running a quantile regression can not be used to estimate the impact of predictor variables on the corresponding unconditional quantile. In our study, it means conditional quantile regression can not answer a question as simple as “what is the impact on 90th quantile customer profit of increasing the number of catalogs everyone received previously (crcpr85) by one unit, holding everything else constant?” (Firpo et al. 2007)

But such question could be of high interest to many database managers as they may want to know the impact of previous promotion on customer profit of this promotion at a specific quantile, i.e. 90th quantile, of a given dataset of customers. Firpo et al. (2007) proposed an unconditional quantile regression, what they call the Recentered Influence Function—RIF regression model, to solve this problem. They argue that if one is interested in the overall effect of covariates on response variable, their unconditional quantile regression should be used to obtain the effects of covariates at different quantiles of the unconditional distribution. They further argue that using conditional quantile regression to estimate the overall effect of covariates on the response variable would yield misleading results.

In our study, we view unconditional quantile regression as a very important complement to profit analysis using conditional quantile regression. While conditional quantile regression gives us a prediction of the whole spending distribution of each customer, unconditional quantile regression provides us an accurate estimation of the effects of covariates on the response variable. Customers who are in certain quantiles will not necessarily stay in the same quantile once the covariates change. Thus, unconditional quantile regression is not able to be used for forecasting. Here, although we do not use unconditional quantile regression in prediction, we run RIF-regression to the same training dataset we used in the previous section for the purpose of comparison. The estimation results are shown in Table 7.

Table 7 OLS and RIF-Regression (Unconditional Quantile Regression) Results

		Quantiles								
Targpbfo	OLS	10th	20th	30th	40th	50th	60th	70th	80th	90th
PRED_M1	2.069	-0.181	-1.052	-0.497	-0.757	-0.576	0.291	0.465	2.848	4.663
	(1.850)	(0.716)	(0.801)	(0.976)	(0.973)	(1.331)	(1.636)	(2.007)	(2.910)	(4.705)
RESID_M1	8.743***	-0.169	-0.352	-1.612*	-0.947	-1.607	-0.227	1.175	6.631*	20.449**
	(1.569)	(0.554)	(0.618)	(0.703)	(0.886)	(1.038)	(1.340)	(1.691)	(3.311)	(7.417)
PRED_R1	1.221***	0.068	0.196*	0.280**	0.316**	0.395**	0.787***	0.802***	1.554***	2.175***
	(0.206)	(0.086)	(0.098)	(0.096)	(0.111)	(0.141)	(0.223)	(0.228)	(0.381)	(0.539)
RESID_R1	0.283***	0.028	0.044	0.094*	0.144***	0.148**	0.159*	0.211*	0.269*	0.574**
	(0.079)	(0.032)	(0.034)	(0.038)	(0.042)	(0.050)	(0.074)	(0.084)	(0.144)	(0.192)
PRED_F1	4.626***	0.046	1.413*	1.438**	1.520**	1.698*	2.925**	2.766*	5.721***	8.926**
	(0.997)	(0.417)	(0.456)	(0.544)	(0.498)	(0.667)	(1.003)	(1.086)	(1.759)	(3.256)
RESID_F1	6.192***	0.244	1.150	1.151	1.280*	1.848*	3.938**	4.383**	8.815***	10.007**
	(1.407)	(0.593)	(0.636)	(0.680)	(0.623)	(0.909)	(1.510)	(1.676)	(2.644)	(3.598)
PRORD85	-0.946*	-0.096	-0.204	-0.399*	-0.523**	-0.478	-0.567	-0.661	-1.353	-2.491*
	(0.391)	(0.142)	(0.167)	(0.190)	(0.194)	(0.250)	(0.381)	(0.412)	(0.775)	(1.412)
SALCAT	5.664***	0.545*	0.875**	1.257***	1.207**	1.713***	3.080***	3.931***	6.978***	10.381***
	(0.732)	(0.268)	(0.317)	(0.374)	(0.396)	(0.495)	(0.869)	(0.862)	(1.602)	(2.185)
HCRD	2.577	-0.486	-0.336	0.099	0.304	0.841	0.357	0.670	3.114	5.309
	(1.630)	(0.744)	(0.597)	(0.836)	(0.856)	(0.962)	(1.397)	(1.633)	(2.851)	(3.616)
CASH	-2.243	-1.918***	-2.235***	-2.844***	-2.781***	-3.062***	-3.792**	-3.483*	-2.749	-3.663
	(1.384)	(0.578)	(0.699)	(0.654)	(0.736)	(0.950)	(1.194)	(1.559)	(2.146)	(2.903)

Absolute value of t-statistics (OLS) and Bootstrap t-statistics (quantile regression) in parentheses.
Note: ***: p<0.001, **: p<0.01, *: p<0.05

Table 7 OLS and RIF-Regression (Unconditional Quantile Regression) Results—Continued

		Quantiles								
Targpbfo	OLS	10th	20th	30th	40th	50th	60th	70th	80th	90th
TELE	5.842***	0.224	0.393	1.040	2.195*	3.010**	5.941***	5.621**	9.602***	15.287***
	(1.635)	(0.726)	(0.802)	(0.762)	(0.870)	(1.169)	(1.436)	(1.937)	(2.980)	(4.473)
CRCPR85	-1.439**	-0.117	-0.089	-0.244	-0.357	-0.428	-0.945**	-0.766	-2.678***	-3.979**
	(0.484)	(0.202)	(0.234)	(0.264)	(0.271)	(0.299)	(0.423)	(0.479)	(0.747)	(1.366)
SALFLG	2.788***	0.917***	1.490***	2.204***	2.449***	3.441***	4.856***	5.197***	5.432***	6.252***
	(0.629	(0.203)	(0.250)	(0.307)	(0.366)	(0.407)	(0.698)	(0.743)	(1.265)	(1.871)
pred_dep	36.288***	2.785*	6.295***	8.680***	11.773***	17.828***	27.291***	31.326***	59.621***	80.083***
	(3.110	(1.117)	(1.407)	(1.352)	(1.773)	(2.046)	(3.439)	(3.569)	(7.662)	(12.804)
Constant	-22.94	7.954	7.170	5.677	7.132	4.089	-9.946	-8.241	-31.465	-39.765
	(8.714	(3.464)	(4.276)	(4.247)	(4.884)	(6.222)	(8.672)	(9.497)	(14.292)	(23.751)
Observations	4794	4794	4794	4794	4794	4794	4794	4794	4794	4794
R2/Pseudo R2	0.175	0.030	0.055	0.076	0.082	0.107	0.127	0.133	0.137	0.121

Absolute value of t-statistics (OLS) and Bootstrap t-statistics (quantile regression) in parentheses.
Note: ***: $p<0.001$, **: $p<0.01$, *: $p<0.05$

Unconditional and conditional quantile regression estimates are close to each other but not identical (Table 7). A quick comparison suggests that there are some differences. Conditional quantile regression estimate of Pred_M1 has a positive effect of 40th quantile (0.224) but unconditional quantile regression estimate has a negative effect (-0.757). Conditional quantile regression estimate of Hcrd has positive effects of 10th (0.275) and 20th quantile (0.081) and a negative effect of 90th quantile (-0.423) but unconditional quantile regression estimate has negative effects of 10th (-0.486) and 20th quantile (-0.336) and a positive effect of 90th quantile (5.309). The magnitude of the effects between conditional and unconditional quantile regression are also very different. Table 8 compares OLS regression with the conditional quantile regression (CQR) and with the unconditional quantile regression at the 10th, 50th, and 90th quantile of the profit distribution.

A close comparison of the P-value of conditional and unconditional quantile regression reveals some interesting differences (Table 8). For the 10th quantile, unconditional quantile regression estimates only have significant ($P<0.01$) effects for two variables, cash and salflg, while conditional quantile regression estimates have significant ($P<0.01$) effects for eight variables, Pred_R1, Resid_R1, Pred_F1, salcat, cash, crcpr85, salflg, and pred_dep. It means that these predictor variables do not have a significant effect of the unconditional 10th quantile profit but have significant effect on the conditional 10th quantile profit.

For conditional quantile regression, the estimated coefficients mean that given that the other covariates constant, what the covariates effects of the predictor variables are. In another words, if the other covariates change, the covariates effects of the estimated predictor variables will change too. For unconditional quantile regression, the estimated coefficients show the covariates effects, which do not conform to any condition. Because conditional quantile regression estimate at any particular conditional quantile mixes the impact of predictor variables for different subgroups, the significance, magnitude, and signs of conditional and unconditional quantile regression estimates are quite different.

The interpretation of unconditional quantile regression estimated coefficients is quite similar as that of conditional quantile regression. The only difference is that the former examines the marginal effects of predictor variables on unconditional quantile response variable. The limitation of conditional quantile regression as a distribution tool in our study is that an unconditional or empirical median profit may well fall at different quantiles depending on the characteristics of the customers. Thus, it is the direct marketers' decision whether to use conditional or unconditional quantile regression. If marketers have an interest in the effects of covariates on the conditional quantile and want to further make forecast of the distribution of a future value, they can choose to use traditional (conditional) quantile regression. If marketers want to determine the effect of any changes of predictor variables on the unconditional customer profit, i.e. 80^{th} quantile, they can use unconditional quantile regression. In a word, conditional and unconditional quantile regressions altogether provide marketers a more informative tool to understand the relationship between customers and company profit.

Table 8 Comparing OLS, Conditional Quantile Regression (CQR) and Unconditional Quantile Regression (UQR)

	OLS	10th	quantile	50th	quantile	90th	quantile
Targpbfo		UQR	CQR	UQR	CQR	UQR	CQR
PRED_M1	2.069	-0.181	-0.549	-0.576	-0.061	4.663	3.531
	(1.85)	(0.716)	(0.436)	(1.331)	(1.096)	(4.705)	(3.753)
RESID_M1	8.743***	-0.169	0.256	-1.607	1.886***	20.449**	8.971*
	(1.569)	(0.554)	(0.189)	(1.038)	(0.562)	(7.417)	(4.071)
PRED_R1	1.221***	0.068	0.134**	0.395**	0.757***	2.175***	2.124***
	(0.206)	(0.086)	(0.049)	(0.141)	(0.118)	(0.539)	(0.396)
RESID_R1	0.283***	0.028	0.075***	0.148**	0.169***	0.574**	0.653***
	(0.079)	(0.032)	(0.017)	(0.050)	(0.044)	(0.192)	(0.184)
PRED_F1	4.626***	0.046	0.904***	1.698*	3.354***	8.926**	6.912***
	(0.997)	(0.417)	(0.274)	(0.667)	(0.649)	(3.256)	(1.959)
RESID_F1	6.192***	0.244	0.534	1.848*	4.187***	10.007**	9.827***
	(1.407)	(0.593)	(0.314)	(0.909)	(0.865)	(3.598)	(2.664)
PRORD85	-0.946*	-0.096	-0.284*	-0.478	-1.033***	-2.491*	-1.676
	(0.391)	(0.142)	(0.140)	(0.250)	(0.285)	(1.412)	(0.724)
SALCAT	5.664***	0.545*	0.429***	1.713***	3.155***	10.381***	11.34***
	(0.732)	(0.268)	(0.106)	(0.495)	(0.548)	(2.185)	(1.573)
HCRD	2.577	-0.486	0.275	0.841	1.64	5.309	-0.423
	(1.630)	(0.744)	(0.317)	(0.962)	(0.941)	(3.616)	(3.171)
CASH	-2.243	-1.918***	-0.869***	-3.062***	-2.306**	-3.663	-5.801**
	(1.384)	(0.578)	(0.238)	(0.950)	(0.723)	(2.903)	(2.142)
TELE	5.842***	0.224	0.627	3.010**	3.055*	15.287***	13.419**
	(1.635)	(0.726)	(0.441)	(1.169)	(1.194)	(4.473)	(4.623)
CRCPR85	-1.439**	-0.117	-0.329***	-0.428	-0.872**	-3.979**	-4.481***
	(0.484)	(0.202)	(0.100)	(0.299)	(0.292)	(1.366)	(1.085)
SALFLG	2.788***	0.917***	1.276***	3.441***	3.533***	6.252***	4.694***
	(0.629)	(0.203)	(0.161)	(0.407)	(0.471)	(1.871)	(1.272)
pred_dep	36.288***	2.785*	6.924***	17.828***	26.805***	80.083***	59.798***
	(3.110)	(1.117)	(1.343)	(2.046)	(2.635)	(12.804)	(7.514)
Constant	-22.94	7.954	7.384	4.089	-7.591	-39.765	-24.22
	(8.714)	(3.464)	(2.097)	(6.222)	(4.931)	(23.751)	(16.714)
Observations	4794	4794	4794	4794	4794	4794	4794
R2/Pseudo R2	0.175	0.030	0.028	0.107	0.086	0.121	0.158

Absolute value of t-statistics (OLS) and Bootstrap t-statistics (quantile regression) in parentheses.
Note: ***: p<0.001, **: p<0.01, *: p<0.05

6 CONCLUSION

6.1 Findings

Our research is one of the first empirical studies that apply quantile regression to analyze marketing problems. In this thesis, we study the customer profitability in the direct marketing context. The descriptive analysis shows that the profit distribution is heavily right-skewed with a large amount of outliers. A small group of customers have a high level of profit, while a large group of customers have a low level of profit. Graphical and statistical analysis both show the heterogeneity in consumer characteristics as well as in their responses to marketing efforts. Although OLS regression can be very powerful and easy to interpret, it is insufficient in such profitability analysis due to its mean-regression nature and its limitations of model assumptions. Quantile regression is especially useful in such a setting.

Heterogeneity from multiple sources has traditionally been an important issue in marketing. Not accounting for consumer heterogeneity may lead to biased parameter estimates and wrong conclusions, especially when targeted marketing efforts are expected. For instance, OLS regression estimates show that the variable which indicates the number of months that have elapsed since the last purchase has the same influence on profit of all customers, while quantile regression estimates suggest that the influence is different on customers of different levels of profit. This study shows that quantile regression provides insight into the effects of the covariates on the conditional profit distribution that may be missed by traditional least-squared estimates. For instance, OLS regression estimates shows that whether to place cash orders does not have any significant influence on profit but quantile regression estimates indicate significant negative influence across different quantiles of customers. The graphical demonstration explicitly shows that the slope estimates of coefficients vary across different quantiles, suggesting that the covariates effects on different groups of customers are not identical and the mean estimate (OLS) is not sufficient. In some cases, the directions of the parameters of high-profit are opposite of that of low-profit groups. Even for the parameters that have the same directions across the deciles, the effect sizes of the parameters can be very different. Thus, controlling for consumer heterogeneity may help improve the accuracy of parameter estimates.

While OLS regression only provides the mean prediction, our study shows that quantile regression provides the whole profit distribution of each individual customer. The prediction intervals given by OLS regression is based on assuming a normal error distribution, but such prediction intervals could be highly inaccurate when the normality assumption is violated. Quantile regression by nature is able to provide probabilistic forecast of future values and gives a more accurate and informative forecast of profit. For instance, OLS regression only gives one mean profit prediction but quantile regression gives prediction of any quantile profit for each observation. We also rank the database by the predicted profit and conduct decile analysis to compare model performance. The 10-fold cross-validation is used to minimize sampling variation. The results show that quantile regression performs slightly better than OLS regression for predictive purposes. However, the interpretation of quantile regression prediction should be done with caution. Managers can not make target selection based on the direct use of conditional quantile regression prediction. Other applications of conditional quantile regression prediction are suggested for the purpose of helping select the most profitable customers.

To our knowledge, this study is the first to make use of unconditional quantile regression in direct marketing context of profitability modeling. For example, when the unconditional high-profit groups of customers are of interest, unconditional quantile regression analysis should be adopted but not the traditional conditional quantile regression. We adopt the Recentered Influence Function—RIF regression proposed by Firpo et al. (2007) to conduct unconditional quantile regression. The results are somewhat different from those of conditional quantile regression. For instance, conditional quantile regression estimates show that whether to place cash orders has positive effects on low quantiles (i.e., 10^{th} and 20^{th}) and negative effects on high quantile (i.e., 90^{th}) while unconditional quantile regression estimates indicate negative effects on low quantiles and positive effects on high quantile. Our study shows that if one is interested in the overall effects of covariates on unconditional quantile of response variable, the conditional quantile regression estimation would yield misleading results.

6.2 Implications

The main features of customer profitability analysis are to estimate the relationships among marketing, customer characteristics and customer profit. Marketing programs have significant influence on customer profit, and customers with different characteristics generate different levels of profit. The main challenges of customer profitability analysis are to make estimation of such relationships, forecast and identify the most profitable customers. But the traditional OLS regression has strict model assumptions and is mean-regression in nature. The violations of OLS model assumptions in customer profit data to lead to biased estimation and even significantly misleading results. Quantile regression does not have any model assumptions and is suitable for customer profitability analysis. Quantile regression is able to make accurate estimates of customers with different characteristics. This thesis adopts quantile regression to customer profitability analysis in direct marketing context and gives reliable estimation. This thesis also makes forecast of the distribution of customer profit of each customer, which is a big challenge for direct marketing researchers. This thesis has meaningful theoretical contributions and practical managerial implications.

This study testifies the applicability of quantile regression for customer profitability analysis in direct marketing. This study is one of the few studies that incorporate customer profitability modeling into consumer response modeling. While the central interest of direct marketers is high profit aside from high response rate, our study provides a useful analytical method to argument customer profitability. Moreover, the quantile regression methods are intuitive and user-friendly and more flexible to handle different consumer groups. It can help uncover consumer heterogeneity and access the effects of marketing mix variables with greater details and accuracy. Examining consumer responses and the effects of consumer characteristics on such relationships present greater insight into the relationship between marketing and customer profitability. For instance, quantile regression can help assess the differential effects of price and promotions on different groups of consumers, such as the high-income vs. the low-income groups. Such analysis can assess the assumptions of the marketing programs and help devise more accurate and effective targeted marketing efforts.

Moreover, Customer Relationship Management (CRM) emphasizes the understanding of customers, their purchase behaviors, and ways to differentiate them. OLS regression only has one point estimate and assumes that customers with different characteristics are homogenous. Thus, OLS regression can only estimate the purchase behaviors of the "average customer" and can not separate customers with different levels of profits. Quantile regression can estimate any quantile of the profit distribution and thus can capture the characteristics and behaviors of different groups of customers. In addition, under segmentation theme, mangers want to identify the segments which will contribute the most profit to the company. For customers with the same characteristics but different profit levels, OLS regression will assign them to one segment with the same profit. Quantile regression gives different coefficients estimates and thus can develop different segments according to their characteristics and purchase behaviors. In a word, the applicability of quantile regression in marketing can improve the effectiveness of segmentation and CRM.

OLS regression has been used as an analytical method in customer profitability analysis, but this study results show that it could provide misleading results. Direct marketers who adopt OLS regression in target selection could view customers response behaviors to direct marketing promotions as homogenous and miss the unique characteristics of high-profit customers, leading to loss of profit to company. OLS regression only gives mean prediction and its prediction intervals could be inaccurate if its normality assumption is violated, leading to biased profit prediction and potential profit loss.

Quantile regression gives a more insightful profit prediction which could be of interest to marketing managers. The probabilistic profit forecast provides managers the probability that they can hope for in terms of customer profit. Such information is very valuable and managers are able to assess the risk of the prediction. The prediction intervals provided by quantile regression are obtained without conforming to any model assumptions and are especially helpful to managers to predict profit of direct marketing promotions. The Recentered Influence Function—RIF regression is proposed by Firpo et al. (2007) to conduct unconditional quantile regression. Our study supports the findings of Firpo et al. (2007) and testifies its usefulness in direct marketing.

Beside the theoretical contributions to customer profit analysis, this thesis also has practical managerial implications for marketing managers. With customer database containing large amount of customer records, marketing managers desire to understand the heterogeneous purchase behavior of customers, the influence of marketing programs on customer profit, and identify the most profitable customers for target selection. The novel method this thesis has proposed, quantile regression, is able to help managers to achieve these goals. The probabilistic quantile regression forecast of customer profit gives managers a complete picture of future profit, thus helping managers understand the degree of uncertainty associated with profit point forecast. The prediction intervals provide managers safe zones of future customer profit forecast, thus helping managers minimize the risk when conducting marketing programs. Unconditional quantile regression answers the questions of greater interest to marketing managers, which is the purchase behavior of the unconditional quantile customers. With quantile regression, marketing managers can gain more insight into customer behaviors and marketing influence and improve target selection.

6.3 Limitations and Future Research

In the meantime, researchers need to be aware of the limitations of these results and the methods. First, there are hundreds of (potential) predictor variables available. Beforehand, it is difficult to determine which of these variables should be included in the analysis. Although we have tested RFM-variables, lifetime value variables, marketing intensity variable, and consumer transactional variables, there are some other variables that may have significant effects on profit. Future research can include other variables into a quantile regression model.

Second, our study shows that some predictors have diversely significant effects on different quantiles of profit. Previous studies suggest including different variables in modeling different quantiles. For example, Whittaker et al. (2005) argue that which covariates are needed to provide explanation and prediction of the target variable depends on which quantile of this distribution is of interest. In our study, we test the same variables in modeling different quantiles. Future research can test different variables when modeling different quantiles of the response variable distribution.

Third, although decile analysis results of 10-fold cross-validation show that quantile regression models for some quantiles outperform OLS regression in predicting high-profit customers, the interpretation of quantile regression prediction is not similar to OLS regression and researchers should be very cautious. For instance, if we predict the median profit, it means if all the customers have 0.5 chance to be below the predicted profit value and (1-0.5 = 0.5) chance to exceed the predicted profit, ranking the database such predicted profit value can reveal how much profit will be generated by the selected customers.

Fourth, the key nature of direct marketing is to identify the most profitable customers. Direct use of quantile regression can examine the effects of different covariates on the response variable among customers at different levels of profit. However, quantile regression can not be used to identify who will be the high-profit customer given certain customers characteristics. And despite the interesting results from this study, forecasting using quantile regression is not straightforward and remains a significant challenge. There are several potential solutions to this problem.

In future study, researchers can use the information of the probabilistic forecast of quantile regression to identify the most profitable customers. As we know the probability each customer will exceed certain predicted quantile of profit values, we can make use of such information to predict the probability that the profit of each customers will exceed the unconditional profit of interest. It is possible to predict the unconditional counterfactual distribution in the test sample suggested by Melly (2006) by intergrating over all quantiles and over all observations by plugging in the covariates. The purchase amount at the top τ_0-percentile of the counterfactual distribution, i.e. $Q_{\tau_0}(y)$, could then be easily computed. Then we can identify which individuals are more likely to spend more than $Q_{\tau_0}(y)$ and rank the customers using the estimated probability in descending order and select customers with the highest probability. Another possible solution is to use copula model to make joint probability estimation of past purchase behavior and future purchase behavior. Such model can produce predictive densities which can be used to forecast future customer profit. The working research of Koenker and Bassett (2008) demonstrate the use of the copula in quantile regression to make forecast of 2005 UCAA basketball tournament. These potential solutions present fruitful advice for future research and can make a significant contribution to forecasting with quantile regression.

REFERENCES

Banslaben, J. (1992). “Predictive Modeling”, in E.L. Nash, editor, The Direct Marketing Handbook, McGraw-Hill, New York, 620–636.

Bassett, Jr., Gilbert W. and Chen, Hsiu-Lang. (2001). “Portfolio Style: Return-Based Attribution Using Quantile Regression”, Empirical Economics. Vol. 26, No. 1, 293-305.

Basu, Amiya K. and Atasi Basu (1995). "Modeling the Response Pattern to Direct Marketing Campaigns," Journal of Marketing Research, Vol. 32, No. 2, 204-212.

Benoit, D.F., and Van Den Poel D. (2009). “Benefits of Quantile Regression for the Analysis of Customer Lifetime Value in a Contractual Setting: an Application in Financial Services,” Forthcoming in Expert Systems with Applications.

Berger, P. D. and Nada I. N. (1998). "Customer Lifetime Value: Marketing Models and Applications", Journal of Interactive Marketing, Vol. 12, No. 1, 17-30.

Berger, P. and Magliozzi T. (1992). "The Effect of Sample Size and Proportion of Buyers in the Sample on the Performance of List Segmentation Equations Generated by Regression Analysis", Journal of Direct Marketing, Vol. 6, No. 1, 13-22.

Berry, M. J. and Gordon S. L. (2000). Mastering Data Mining: the Art and Science of Household Relationship Management, New York: Wiley.

Bitran, G. R. and Mondschein V. S. (1996). "Mailing Decisions in the Catalog Sales Industry", Management Science, Vol. 42, No. 9, 1364-1381.

Blattberg, R. and Sen, S.K. (1976). “Alternative Procedures for Modeling Buyer Behavior Using Consumer Panels” Proceedings of the American Statistical Association.

Blattberg, R., Getz, G., and Thomas, S. J. (2001). “Customer Equity: Building an Managing Relationships as Valuable Assets.” Harvard Business School Press, Boston, Mass.

Blundell, R.W., and Powell, J.L., (2004). “Endogeneity in Semiparametric Binary Response Models”. Rev. Econom. Stud. Vol. 71, No. 7, 655-679.

Bodenberg, T.M. and Roberts, M.L. (1990). “Integrating Marketing Research into the Direct-Marketing Testing Process: the Market Research Test”, Journal of Advertising Research, Vol 14, No. 4, 378-394.

Bremnes, J.B. (2004). “Probabilistic Forecasts of Precipitation in Terms of Quantiles using NWP Model Output”, Monthly Weather Review, Vol. 132, No. 1, 338-347.

Bremnes, J.B. (2004). “Probabilistic Wind Power Forecasts Using Local Quantile Regression”, Wind Energ. Vol. 7, No. 1, 47-54.

Buchinsky, M. (1998). “Recent Advances in Quantile Regression Models”, Journal of Human Resources. Vol. 33, No. 1, 88-126.

Buhai, I.S. (2005). “Quantile Regression: Overview and Selected Applications”, Ad-Astra -The Young Romanian Scientists’ Journal, Vol. 4, 1-17.

Bult J.R. (1993). “Semiparametric versus Parametric Classification Models: An. Application to Direct Marketing”, Journal of Marketing Research. Vol. 30, No. 3, 380–390.

Bult, J.R. and T.J. Wansbeek. (1995). "Optimal Selection for Direct Mail", Marketing Science. Vol 4, No. 4, 378-394.

Bult, Roelf, J. and Wittink, D.R. (1996). "Estimating and Validating Asymmetric Heterogeneous Loss Functions Applied to Health Care Fund Raising", International Journal of Research in Marketing. Vol. 13, 215-226.

Cade, B.S. and Noon, B.R. (2003). "A Gentle Introduction to Quantile Regression for Ecologists" Front Ecol Environ, Vol. 1, No. 8, 412-420.

Chaudhuri, Probal and Wei-Yin Loh (2002). "Nonparametric Estimation of Conditional Quantiles Using Quantile Regression Trees", Bernoulli Vol. 8, No. 5, 561-576.

Campbell, D. and Frei, F. (2004). "The Persistence of Customer Profitability: Empirical Evidence and Implications from a Financial Services Firm", Journal of Service Research. Vol. 7, No. 2, 107-124.

Chernozhukov, V. and. Umantsev L. (2001). "Conditional Value-At-Risk: Aspects of Modeling and Estimation", Empirical Economics. Vol. 26, No. 1, 271-92.

Cui, G., Wong, M. L. Zhang, G.C. and Lin L. (2008). "Model Selection for Direct Marketing: Performance Criteria and Validation Methods", Marketing Intelligence & Planning. Vol. 26, No. 3, 275.

Cui, G., Wong, M. L. Zhang, and Lui H.K. (2006). "Machine Learning for Direct Marketing Response Models: Bayesian Networks with Evolutionary Programming", Management Science. Vol. 52, No. 4, 597-612.

Cui G. and Wong, M. L. (2004). "Implementing Neural Networks for Decision Support in Direct Marketing", International Journal of Market Research. Vol. 46, No. 2, 235-254.

Deaton, A. (1997). The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Baltimore MD: Johns Hopkins University Press.

DeSarbo, W.S. and V. Ramaswamy (1994). "CRISP: Customer Response Based Iterative Segmentation Procedures for Response Modeling in Direct Marketing", Journal of Direct Marketing, Vol. 8, No. 3, 7–20.

Donkers B., Paap, R., Jonker, J. J., and Franses, H. P. (2006). "Deriving Target Selection Rules from Endogenously Selected Samples", Journal of Applied Econometrics. Vol. 21, No. 5, 549-562.

Dotson, J. P., Retzer, J. and Allenby G. M. (2007). "A Simultaneous Quantile Regression Model for Customer Linkage Analysis" Working paper, Ohio State University.

Eide, E. and Showalter. M. (1998). "The Effect of School Quality on Student Performance: A Quantile Regression Approach." Economics Letters. Vol. 58, No. 3, 345-50.

Firpo, S., Fortin, N.M., Lemieux, T. (2007a). "Unconditional Quantile Regression." Technical working paper No. 339, National Bureau of Economic Research, Washington, DC.

Firpo, S., Fortin, N.M., Lemieux, T. (2007b). "Decomposing Wage Distributions Using Influence Function Projections." Working paper. Department of Economics, University of British Columbia.

Fitzenberger, B. (1999). Wages and Employment across Skill Groups, Heidelberg: PhysicaVerlag.

Fitzenberger, B., Koenker, R., and Machado, J.A.F. (2002). Economic Applications of Quantile Regression. Physica-Verlag GmbH & Co.

Gönül, Füson F., and Meng Z. S. (1998). “Optimal Mailing of Catalogs: A New Methodology Using Estimable Structural Dynamic Programming Models”, Management Science, Vol. 44, No. 9, 1249-1263.

Gönül, Füson F., Byung-Do Kim and Mengzhe Shi, (2000). “Mailing Smarter to Catalog Customers,” Journal of Interactive Marketing, Vol. 14, No. 2, 2-16.

Hao, L.X. and Naiman, D.Q. (2007). “Quantile Regression” Thousand Oaks CA: Sage Publications.

Haughton, D. and Oulabi, S. (1997). “Direct Marketing Modeling with CART and CHAID” Journal of Direct Marketing, Vol. 11, No. 4, 42-52.

Heckman. J. (1979). “Sample Selection Bias as A Specification Error”, Econometrica. Vol. 47, No. 1, 153-161.

Heinen, T. (1996). “Latent Class and Discrete Latent Trait Models: Similarities and Differences”, Sage Publications, Thousand Oaks, CA.

Jain, D., Bass, F. M., and Chen, Y. M. (1990). “Estimation of Latent Class Models with Heterogeneous Choice Probabilities: An Application to Market Structuring”, Journal of Marketing Research. Vol. 27, No. 1, 94-101.

Jedidi, K., Jagpal, H. S., and Desarbo, W. S. (1997). “Finite-Mixture Structural Equation Models for Response-Based Segmentation and Unobserved Heterogeneity”. Marketing Science, Vol. 16, No. 1, 39-59.

Jonker, J.J, Franses, P.H, and Piersma, N. (2004). “Evaluating Direct Marketing Campaigns: Recent Findings and Future Research Topics” ERIM Report Series Reference No. ERS-2002-26-MKT

Kaymak, U. (2001). “Fuzzy Target Selection Using RFM Variables”. Proc. of Joint 9th IFSA World Congress and 20th NAFIPS Int. Conference, Vancouver, Canada, 1038-1043

Koenker, R. and Bassett, G. (1978). “Regression Quantiles.” Econometrica. January, Vol. 46, No. 1, 33-50.

Koenker, R. and Bassett, G. (2009) “March Madness, Quantile Regression Bracketology and the Hayek Hypothesis”, in the forthcoming, Journal of Business and Economic Statistics.

Koenker, R. and Bilias, Y. (2001). “Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments.” Empirical Economics. March, Vol. 26, No. 1, 199-220.

Koenker, R. and Hallock F. K. (2001). “Quantile Regression”. Journal of Economic Perspectives, Vol. 15, No.1, 143-156.

Koenker, R. and Machado, J.A.F. (1999). “Goodness of Fit and Related Inference Process for Quantile Regression”. Journal of American Statistics Association, Vol. 94, No. 448, 1296-1310.

Koenker, R., Ng P, and Portnoy, S. (1994). “Quantile Smoothing Splines”. Biometrika, Vol. 81, No. 4, 673-680.

Koenker, R. and Park, B.J. (1996). “An Interior Point Algorithm for Nonlinear Quantile

Regression". Journal of Econometrics, Vol. 71, No. 1-2, 256-283.

Koenker, R. (2005). "Quantile Regression" Cambridge University Press, Cambridge, UK.

Levin, N. and Zahavi, J. (2001). "Predictive Modeling Using Segmentation", Journal of Interactive Marketing, Vol. 15, No. 2, 2-22.

Libai, B., Narayandas D., and Humby C. (2002). "Toward an Individual Customer Profitability Model: A Segment-Based Approach", Journal of Service Research, Vol.5, No. 1, 69-77.

Ma, L.J. and Pohlman, L. (2008). "Return Forecasts and Optimal Portfolio Construction: A Quantile Regression Approach," The European Journal of Finance, Vol. 14, No. 5, 409-425.

Machado, J. and Mata, J. (1999). "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression", Journal of Applied Econometrics. Vol. 20, No. 4, 445-466.

Magidson, J. (1988). "Improved Statistical Techniques for Response Modeling", Journal of Direct Marketing, Vol. 2, No. 4, 6–18.

Melly, B. (2006). "Public And Private Sector Wage Distributions Controlling for Endogenous Sector Choice", Swiss Institute for International Economics and Applied Economic Research (SLAW), University of St. Gallen.

Mueller, R. (2000). "Public- and Private-Sector Wage Differentials in Canada Revisited." Industrial Relations. July, Vol. 39, No. 3, 375-400.

Mulhern, F. J. (1999). "Customer Profitability Analysis: Measurement, Concentration, and Research Directions," Journal of Interactive Marketing. Vol. 13, No. 1, 25-41.

Neter, J.M., Kutner, H., Nachtsheim, C.J., and Wasserman, W. (1996). Applied Linear Statistical Models, Chicago, IL:Irwin.

Otter W. Pieter, Hiek van der Scheer, Wansbeek Tom. (2006). "Optimal Selection of Households for Direct Marketing by Joint Modeling of the Probability and Quantity of Response". University of Groningen, CCSO Centre for Economic Research, Working Papers.

Perlich, C., Rosset, R., Lawrence, R., and Zadrozny, B. (2007). "High Quantile Modeling for Customer Wallet Estimation with Other Applications," Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 12th-15th, San Jose, CA.

Racine, J. S. (2002). "`New and Improved' Direct Marketing: A Nonparametric Approach," in A. Montgomery and P. H. Franses (Eds), Advances in Econometrics: Econometric Models in Marketing, Vol. 16, 139-162.

Rao, V.R. and Steckel, J.H. (1995), "Selecting, Evaluating, and Updating Prospects in Direct Mail Marketing", Journal of Direct Marketing, Vol. 9, No. 2, 20–31.

Reinartz, Werner J. and Kumar, V. (2003). "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration", Journal of Marketing, Vol. 67, No. 1, 77-99.

Reinartz, Werner J. and Kumar, V. (2000). "On the Profitability of Long-life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing", Journal of Marketing, Vol. 64, No. 4, 17-35.

Subom, Rhee and Gary J. Russell. (2008), "Forecasting Household Response in Database

Marketing: A Latent Trait Approach" Forthcoming in Ronald Klimberg (ed.), Advances in Business and Management Forecasting, New York: Elsevier Academic Publishers.

Roberts, M.L. and Berger, P.D. (1989). Direct Marketing Management, Englewood Cliffs, NJ: PrenticeHall Inc.

Rosset, S., Perlich, C., Zadrozny, B., Merugu, S., Weiss, S., and Lawrence, R. (2005). "Wallet estimation models", In Proc. of the Intl. Workshop on CRM: Data Mining Meets Marketing, November 18th & 19th, New York City.

Sen, A. (1997). On Economic Inequality. Oxford: Clarendon Press.

Shepard, David (1999). "Separate Models for Separate Segments?" Direct Magazine, August, htttp: www.dsadirect.com/html/article_aug_99.html, downloaded Feb. 23, 2001.

Shields, M. D. (1997). "Research in Management Accounting by North Americans in 1990s," Journal of Management Accounting Research, Vol. 9, No. 1, 3-61.

Simon, J.L. (1987), How to Start and Operate a Mail-Order Business, McGraw-Hill, New York.

Smith, R.J., and Blundell, W.R. (1986), "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply", Econometrica. Vol. 54, No. 3, 679-685.

Somers, M. and Whittaker, J. (2007), "Quantile Regression for Modelling Distributions of Profit and Loss", European Journal of Operational Research. Vol. 183, No. 3, 1477.

Sousa, S.I.V., Pires, J.C.M., Martins, F.G., Pereira, M.C., and Alvim-Ferraz, M.C.M. (2009), "Potentialities of Quantile Regression to Predict Ozone Concentrations", Environmetrics, Vol. 20, No. 2, 147-158.

Spring, P.N., Leeflang, P.S.H. and Wansbeek, T.J. (1999), "The Combination Strategy to Optimal Target Selection and Offer Segmentation in Direct Mail", Journal of Market Focused Management, Vol. 4, No. 3, 187–203.

Taylor, J.W. (2007). "Forecasting Daily Supermarket Sales Using Exponentially Weighted Quantile Regression", European Journal of Operational Research, Vol. 178, No. 26, 154-167.

Taylor, J. W. (1999). "A Quantile Regression Approach to Estimating the Distribution of Multiperiod Returns," Journal of Derivatives, Vol. 7, No. 1, 64-78.

Taylor, J. W. (1999). "Investing Improvements in the Accuracy of Prediction Intervals for Combination of Forecasts: A Simulation Study". International Journal of Forecasting Vol. 15, No. 3, 325-339.

Taylor, J. W. and Bunn, W. D. (1999). "A Quantile Regression Approach to Generating Prediction Intervals," Management Science; Vol. 45 (Feb), 225-237.

Trasher, R.P. (1991), "Cart: A recent advance in Tree-Structured List Segmentation Methodology", Journal of Direct Marketing, Vol. 5, No. 1, 35-47.

Viscusi, W. Kip and Hamilton, J. (1999). "Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions." American Economics Review. Vol. 89, No. 4, 1010-27.

Wedel, M., Desarbo, W. S., Bult, J. R., and Ramaswamy, V. (1993). "A Latent Class Poisson Regression Model for Heterogeneous Count Data". Journal of Applied Econometrics, Vol. 8, No. 4, 397-411.

Whittaker, J., Whitehead, C., and Somers, M. (2005). “The Neglog Transformation and Quantile Regression for the Analysis of a Large Credit Scoring Database”, Journal of the Royal Statistical Society: Series C (Applied Statistics) Vol. 54, No. 5, 863-878.

Yu, K. and Jones, M.C. (1998). “Local Linear Quantile Regression” Journal of American Statistical Association. Vol. 93, No. 441, 228-237.

Zadrozny, B., and Elkan, C. (2001). “Learning and Making Decisions When Costs and Probabilities Are Both Unknown”, In Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining, August 26-29, San Francisco, CA.

Zahavi, J. and Levin, N. (1997), “Applying Neural Computing to Target Marketing”, Journal of Direct Marketing, Vol. 11, No. 1, 5-22.

Zeithaml, V. A., Rust, T. R., and Lemon, N. K. (2001). “The Customer Pyramaid: Creating and Serving Profitable Customers”, California Management Review. Vol. 43, No. 4, 118-142.